SUBSAMPLING
IN INFORMATION THEORY AND DATA PROCESSING

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

Yuxin Chen
January 2015

Yuxin Chen

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 J. Goldsmith) Principal 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.

(Andrea Montanari)

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.

(Tsachy Weissman)

Approved for the Stanford University Committee on Graduate Studies.

Abstract

An ubiquitous challenge in modern data and signal acquisition arises from the ever-growing size of the object under study. Hardware and power limitations often preclude sampling with the desired rate and precision, which motivates the exploitation of signal and/or channel structures in order to enable reduced-rate sampling while preserving information integrity. This thesis is devoted to understanding the fundamental interplay between the underlying signal structures and the data acquisition paradigms, as well as developing efficient and provably effective algorithms for data reconstruction. The main contributions of this thesis are as follows.

- We investigate the effect of sub-Nyquist sampling upon the capacity of a continuous-time channel. We start by deriving the sub-Nyquist sampled channel capacity under periodic sampling systems that subsume three canonical sampling structures, and then characterize the fundamental upper limit on the capacity achievable by general time-preserving sub-Nyquist sampling methods. Our findings indicate that the optimal sampling structures extract out the set of frequencies that exhibits the highest signal-to-noise ratio and is alias-suppressing. In addition, we illuminate an intriguing connection between sampled channels and MIMO channels, as well as a new connection between sampled capacity and MMSE.

- We study the universal sub-Nyquist design when the sampler is designed to operate independent of instantaneous channel realizations, under a sparse multiband channel model. We evaluate the sampler design based on the capacity loss due to channel-independent sub-Nyquist sampling, and characterize the minimax

capacity loss. This fundamental minimax limit can be approached by random sampling in the high-SNR regime, which demonstrates the optimality of random sampling schemes.

- We explore the problem of recovering a spectrally sparse signal from a few random time-domain samples, where the underlying frequencies of the signal can assume any *continuous* values in a unit disk. To address a basis mismatch issue that arises in conventional compressed sensing methods, we develop a novel convex program by exploiting the equivalence between (off-the-grid) spectral sparsity and Hankel low-rank structure. The algorithm exploits sparsity while enforcing physically meaningful constraints. Under mild incoherence conditions, our algorithm allows perfect recovery as soon as the sample complexity exceeds the spectral sparsity level (up to a logarithmic gap).

- We consider the task of covariance estimation with limited storage and low computational complexity. We focus on a quadratic random measurement scheme in processing data streams and high-frequency signals, which is shown to impose a minimal memory requirement and low computational complexity. Three structural assumptions of covariance matrices, including low rank, Toeplitz low rank, and jointly rank-one and sparse structure, are investigated. We show that a covariance matrix with any of these structures can be universally and faithfully recovered from near-minimal sub-Gaussian quadratic measurements via efficient convex programs for the respective structure.

All in all, the central theme of this thesis is on the interplay between economical subsampling schemes and the structures of the object under investigation, from both information-theoretic and algorithmic perspectives.

Acknowledgement

There are a great number of people that I am grateful to for making my past four years at Stanford the most rewarding and enlightening journey of my life. First and foremost, I owe my deepest gratitude to my advisor, Prof. Andrea J. Goldsmith, for her continual support and unwavering encouragement throughout the course of my PhD. Her vision about wireless communications and beyond, her enthusiasm for research, her unbounded energy, and her broad and deep knowledge have always been inspirational to me. My improvements on how to identify meaningful problems and how to uncover interesting insights out of complex mathematical equations are largely attributed to Andrea. Her ability to multitask (while achieving the best possible quality in all tasks) is also something that I will always look up to.

I am grateful to the entire faculty of the Department of Electrical Engineering and Department of Statistics at Stanford for creating an inspirational and enriching environment for graduate studies. I owe my particular gratitude to Prof. Emmanuel J. Candès, who has provided me with invaluable guidance and positive encouragement despite his busy schedules. His unparalleled wisdom and passion in bringing together theory and practice, as well as his striving for the cleanest and most insightful deliveries, have been exceedingly instructive. Every time I discussed with him, I ended up learning a great deal. I would like to thank Prof. Andrea Montanari, who has been a wonderful reference and encylopedia-like resource. I have benefited greatly from my several interactions with Andrea, especially by learning the way he approached a new problem. I am also grateful to Prof. Tsachy Weissman and Prof. Abbas El Gamal for being great teachers and for serving on my oral committee. I have learned a great deal from their feedback and advice as well as their thought-provoking courses and lectures on various aspects of information theory. I would also like to thank Prof. David N.C. Tse. While I have not had much research interaction with David, his insights conveyed in his papers and lectures and his emphasis on striving for intuition have been the most eye-opening to me, which have generated my very first impulse to enter the theoretical realm.

I wish to thank my master's supervisor, Prof. Jeffrey G. Andrews, for his endless help, continual support, and always positive encouragement throughout my graduate studies. I have also been fortunate to have worked with Prof. Yonina C. Eldar. Her boundless energy and broad knowledge about signal processing have been very instructive to me. In addition, I am thankful to my colleagues and great friends Yuejie Chi, Qixing Huang, Xiaodong Li, and Yao Xie. My extensive research interactions with them in various fields have been exceedingly inspiring and enjoyable.

Pursuing my PhD at the wireless system lab at Stanford has been a memorable journey. I would like to thank my labmates for their caring and technical assistance: Jinyuan Chen, Mainak Chowdhury, Sina Firouzabadi, Pulkit Grover, Jinhua Jiang, Alon Kipnis, Alexandros Manolakos, F. Javier Lopez Martinez, Reza Mirghaderi, Yair Noam, Se-Yong Park, Milind Rao, Stefano Rini, Nima Soltani, Boon Sim Thian, Yao Xie, Yongping Zhang, Yue Zhao, Neta Zuckerman. I am also greatly indebted to Pat Oshiro and Mimi Qian for efficiently taking care of all logistics issues for me.

My life at Stanford would have been miserable without the accompany and support from my invaluable friends, particularly those from Stanford and Tsinghua. Their positive energy, love, and encouragement have brought me endless happiness, and have become my most memorable treasures.

Finally and most importantly, I am deeply indebted to my family, particularly my parents and my grandmother, for their unconditional love and support throughout all my endeavors, and for always being by my side throughout all the ups and downs of my life.

This thesis is dedicated to my parents and my grandmother.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Digital signal and data processing have revolutionized information processing and transmission, by transforming analog waveforms or time series into succinct discrete-time data samples and then performing all processing and inference over the digital domain. Conventional practice and theory in signal processing focus on the scenario where sampling is performed with rates at or above the ambient degrees of freedom of the object under study (e.g. the Nyquist rate for band-limited signals), as this is in general the necessary sampling rate limit for preserving signal integrity.

One ubiquitous challenge in modern signal and data acquisition, however, is the ever-growing size (or "dimensionality") of the signal of interest. Hardware limitations (e.g. operating frequency, storage, processing ability) impose significant cost in acquiring high-rate samples, and have stunted the development and implementation of traditional design paradigms based on perfect-rate samples (e.g. Nyquist-rate samples for bandlimited signals). Two example scenarios include:

- **Wireless Communications**. One of the underlying premises of digital communications is to exploit analog-to-digital converters (ADCs) to translate analog signals into digital data. Unfortunately, compared to the rapid growth of communication bandwidth in cutting-edge wideband systems, ADCs are often limited in their sampling frequency and precision, which precludes high-rate accurate sampling over wide bandwidths in the analog front-end. Moreover, for high-performance communications, the amount of input data that can be supported and processed by digital signal processors (DSPs) are also limited. Such limitations present a major bottleneck in transferring promising wideband receiver design paradigms from research to practice.

- **Data Stream Processing**. A high-dimensional data stream model represents *real-time* data that arrives sequentially at a high rate, where each data instance might itself be high-dimensional. However, dealing with massively generated data poses a challenge for data acquisition, storage, and information extraction. In many resource-constrained applications, the available memory and processing power at the data acquisition devices are severely limited compared with the volume and rate of the data. Consequently, it is desirable to perform information processing and inference from inputs on the fly without storing much data.

Fortunately, objects under study are often not parsimoniously representable in the raw data domain, and hence the ambient dimension defined with respect to the raw data domain might significantly exceed the underlying degrees of freedom of objects. A subset of samples acquired in this domain might contain minimal innovative information and are thereby redundant. As quoted from Donoho when he opened his landmark paper [1]:

"Why go to so much effort to acquire all the data when most of what we get will be thrown away? Can we not just directly measure the part that will not end up being thrown away? "

In order to obtain economical and information-preserving samples that will not "end up being thrown away", the key ingredient is to exploit the underlying structures (e.g. sparsity, compressibility) of signals and/or transmission mediums in the data transmission and acquisition stages. This gives rise to two fundamental questions, which we explore in this thesis:

- *What is the fundamental tradeoff between the sampling rate constraint and the data transmission rates that can be supported by a transmission medium?*

- *How to design efficient data acquisition and reconstruction paradigms by exploiting specific structural priors on the underlying object?*

In particular, we investigate efficient sampling protocols from both information theoretic and algorithmic perspectives for several different scenarios, detailed as follows.

1.1 Fundamental Limits: Capacity of Sub-Sampled Communication Channels

Shannon pioneered the field of information theory, including the fundamental capacity limits of continuous-time channels and the corresponding capacity-achieving water-filling power allocation strategies [2]. The notion of analog channel capacity, which is commensurate with the maximum achievable information rate under Nyquist-rate sampling, underlie conventional design and analysis of wireless communication systems. Since Nyquist-rate sampling is information-preserving, capacity of analog channels sampled at or above the Nyquist rate does not depend on the sampling mechanism. In wideband communication systems, however, hardware and power limitations often preclude sampling at the Nyquist rate. This motivates the investigation as to whether these hardware bottlenecks can be circumvented by designing digital communication system receivers that are sampled at sub-Nyquist rates. To this end, our investigation centers around two fundamental questions at the intersection of sampling theory and information theory: 1) how much information, in the Shannon sense, can be conveyed through undersampled communication channels; 2) under a sub-Nyquist sampling-rate constraint, which sampling structures should be employed in order to maximize the information rate.

To address these questions, we explore the performance cost as well as optimal sampling mechanisms under two different situations:

1. The class of samplers designed based on perfect channel state information (CSI);
2. The class of samplers designed independent of the instantaneous channel realization.

1.1.1 Background and Motivation

1.1.1.1 Channel-Optimized Sampling

The derivation of the capacity of LTI Gaussian channels was pioneered by Shannon [2], based on asymptotic spectral properties of Toeplitz operators [3]. This capacity result established a beautiful connection between information theoretic concepts and the signal-to-noise ratio (SNR) across the frequency domain, a measure often used in the context of data acquisition and estimation. On the other hand, the Shannon-Nyquist sampling theorem, which dictates that channel capacity is preserved when the received signal is sampled at or above the Nyquist rate, has frequently been used to transform analog channels into their discrete counterparts (e.g.[4]). For instance, this paradigm of discretization was employed by Medard [5] to bound the maximum mutual information in time-varying channels. However, the mainstream information theory works are concerned with channels sampled at or above the Nyquist rate, and do not explicitly account for the effect upon capacity of sub-Nyquist sampling.

The Nyquist rate is known to be the fundamental sampling rate limit necessary for preserving signal integrity of bandlimited signals or, more generally, the class of shift-invariant signals. Various sampling methods have been proposed for bandlimited functions [6, 7, 8], including recurrent nonuniform sampling [9] that samples the signal with blockwise periodic sample times, and generalized multi-branch sampling [10] that samples the signal through multiple linear systems. In general, for perfect signal reconstruction, these methods require sampling at an aggregate rate at or above the Nyquist rate. However, for signals with certain structure, the Nyquist rate could be excessive for signal recovery [11, 12]. For example, consider multiband signals, whose spectral contents reside within several subbands over a wide spectrum. If the spectral support is known, then the necessary sampling rate for multiband signals is their spectral occupancy, termed the *Landau rate* [13]. One popular mechanism that allows reconstruction of multiband signals under Landau-rate sampling is filterbank sampling, studied in [14, 15].

Although nonuniform sampling methods for bandlimited signals have been extensively explored, they are typically investigated either under a noiseless setting, or based on statistical measures (e.g. mean squared error (MSE)) instead of information theoretic concerns. The most relevant capacity result to our work was by Berger et. al. [16, 17, 18], which investigated joint optimization of the transmitted pulse shape and receiver prefiltering and related MSE-based optimal sampling with capacity for several special types of channels. But it remains unclear how a broader class of nonuniform sampling will affect the rates of information that can be conveyed through more general types of channels.

1.1.1.2 Channel-Independent Sampling

Practical communication systems often involve time-varying channels, e.g. wireless slow fading channels [19, 20]. Many of these channels can be modeled as a channel with state (see a detailed survey in [20, Chapter 7]), where the channel variation is captured by a state that may be fixed over a long transmission block, or, more simply, a compound channel [21] whereby the channel realization lies within a collection of possible channels [22]. One class of compound channel models concerns multiband Gaussian channels, whereby the instantaneous frequency support active for transmission resides within several continuous intervals, spread over a wide spectrum. This model naturally arises in several wideband communication systems, including time division multiple access systems and cognitive radio networks.

Nevertheless, practical receiver hardware and, in particular, sampling mechanisms are typically integrated into the hardware and hence need to be designed based on a family of possible channel realizations instead of instantaneous CSI. This has no effect if the sampling rate employed is commensurate with the maximum bandwidth (or the Nyquist rate) of the channel family. However, in the sub-Nyquist sampling rate regime, the sampler design significantly impacts the information rate achievable over different channel realizations. For channels with uncertainty, *sampled capacity loss* relative to the Nyquist-rate capacity is necessarily incurred due to channel-independent (universal) sub-Nyquist sampler design. As will be shown, the capacity-optimizing sampler for a fixed channel structure might result in very low data rate for other channel realizations.

In fact, spectrum-blind sub-Nyquist sampling schemes have recently received much attention in the signal processing community, inspired by the recent breakthrough in "compressed sensing" [23, 24, 1, 25]. These works assert that the sampling rate can be reduced significantly below the Nyquist rate without performance loss, as long as the sampling bases are sufficiently incoherent with the spectrum sparsity patterns. However, to the best of our knowledge, no prior work has investigated, from a capacity perspective, a universal (channel-independent) sub-Nyquist sampling paradigm that is robust to channel variations in the above channel models. In the discrete-time settings, Donoho *et. al.* [26] asserted that random and band-diagonal sampling systems admit perfect signal recovery from an information theoretically minimal number of samples. Nevertheless, the optimality was based on a fundamental rate-distortion limit [27] instead of channel capacity as a metric.

All in all, for both channel-optimized or channel-blind sampler designs, generalizing conventional notions of capacity to sub-Nyquist sampled channels will provide important metrics and machinery to analyze the effect of reduced-rate sampling on communication systems and to optimize the hardware structure for a given sampling rate dictated by complexity and technology constraints.

1.1.2 Main Contributions

1.1.2.1 Capacity under Channel-Dependent Sampling

We consider linear time-invariant (LTI) Gaussian channels, where perfect CSI is known at both the transmitter and receiver and the sampler is optimized based on perfect CSI. We first investigate three classes of sampling mechanisms of increasing complexity, which are the most commonly employed in practice: 1) an LTI filter followed by uniform sampling; 2) a filter bank followed by uniform sampling; 3) uniform sampling following periodic modulation and filter banks. These sampling structures admit closed-form expressions for sub-sampled channel capacity (Theorems 2.1 - 2.3) that have close bearings to multiple-input-multiple-output (MIMO) channels. The capacity-maximizing sampling structures within these classes are alias-suppressing, which coincide with those minimizing the mean squared errors (MSE) between the transmitted and reconstructed signals. One interesting fact we discover for these techniques is the non-monotonicity of capacity with sampling rate, which indicates that at certain sampling rates, channel degrees of freedom are lost.

It turns out that all these sampling mechanisms fall under the category of *periodic sampling*, namely, the sampling structures consist of a preprocessor with periodic impulse response and a recurrent sampling set. This enables a unified framework for capacity analysis (Theorem 2.4) for a broad class of sampling structures. While closed-form capacity expressions might not exist for aperiodic samplers, we are able to characterize a fundamental upper limit on sampled channel capacity (Theorem 2.5), which accommodates a very general class of time-preserving nonuniform sampling methods under a sub-Nyquist sampling rate constraint. In particular, the class of sampling systems we consider subsumes sampling structures employing irregular nonuniform sampling grids. Interestingly, this fundamental capacity limit can be achieved by filterbank sampling with varied sampling rates at different branches (Theorem 2.6), or by a single branch of modulation and filtering followed by uniform sampling (Theorem 2.7). The optimal samplers effectively extract out a spectral set of size f_s with the highest SNR, and suppress signal and noise contents outside this SNR-maximizing spectral set.

Our results reveal that irregular nonuniform sampling sets, while typically complicated to realize in hardware, do not increase channel capacity relative to analog preprocessing with regular uniform sampling sets. We also show that when optimal filterbank or modulation sampling is employed, a mild perturbation of the optimal sampling grid does not change the capacity. Our findings demonstrate that aliasing or scrambling of spectral contents does not provide capacity gain. This is primarily due to the assumptions that perfect CSI is available during the sampler design stage.

1.1.2.2 Minimax Capacity Loss under Channel-Independent Sampling

This part focuses on a multiband channel, whereby the channel bandwidth is divided into n continuous subbands and, at each timeframe, only a few subbands are active for transmission. A sampling mechanism is termed a Landau-rate sampling (resp.

super-Landau sampling) system if the sampling rate is equal to (resp. greater than) the spectral occupancy of the instantaneous channel realization. Our contributions are as follows.

1. We derive, in Theorem 3.3, a fundamental lower bound on the largest sampled capacity loss (defined in Section 2.1) incurred by any channel-independent periodic sampler, under both Landau-rate and super-Landau-rate sampling. This lower bound depends only on the band sparsity ratio and the undersampling factor, modulo a small residual term that vanishes when SNR and n increase.

2. Theorem 3.4 characterizes the sampled capacity loss under a class of periodic sampling with periodic modulation and *low-pass* filter banks, when the Fourier coefficients of the modulation waveforms are randomly generated and independent (termed *independent random sampling*). We demonstrate that with exponentially high probability, the sampled capacity loss matches the fundamental lower bound of Theorem 3.3 uniformly across all channel realizations. Our results demonstrate that independent sampling achieves the minimax capacity loss, as long as the Fourier coefficients of the modulation waveforms are independently sub-Gaussian with matching moments up to the second order. This universality phenomenon occurs due to sharp concentration of spectral measures of large random matrices [28].

3. For a large super-Landau sampling regime, we quantify the sampled capacity loss under independent random sampling when the Fourier coefficients of the modulation waveforms are i.i.d. Gaussian (termed *Gaussian random sampling*), as stated in Theorem 3.5. Encouragingly, Gaussian random sampling achieves minimax capacity loss (modulo some vanishing residual terms) with exponentially high probability.

4. Along the way we derive sharp concentration of measure of several log-determinant functions for i.i.d. random matrix ensembles, which might be of independent interest for other works involving log-determinant metrics.

1.2 Algorithms: Structured Estimation from Undersampled Data

High-dimensional statistical inference requires efficient estimation of a large number of parameters when the number of acquired samples is far smaller than the ambient degrees of freedom. A growing body of algorithmic advances have established that this is possible provided a) the object under investigation possesses a low-dimensional structure, b) the measurements are "incoherent", i.e. they do not suppress this structure, c) the recovery paradigm is capable of capturing the underlying structure of the object. Two of the key challenges when developing sampling and recovery paradigms are:

- *How to enforce appropriate structural assumptions that are consistent with the physical constraints of a problem?*
- *How to develop incoherent sampling patterns that are compatible with hardware constraints?*

In this thesis, we investigate the above challenges under two settings:

1. **Spectral compressed sensing**: reconstruction of a spectrally sparse signal from partial time-domain samples, where the spectral spikes can assume any value over a *continuous* region;
2. **Covariance estimation**: recovery of covariance matrices from quadratic sampling that naturally arise in data stream processing or high-frequency applications.

Two common themes underlying our results are the employment of certain random sampling strategies and the development of computationally efficient recovery paradigms based on convex optimization.

1.2.1 Spectral Compressed Sensing via Structured Matrix Completion

A large class of practical applications features high-dimensional signals that can be modeled or approximated by a superposition of spikes in the spectral (resp. time) domain, and involves estimation of the signal from its time (resp. frequency) domain samples. Examples include medical imaging [29, 30], target localization in radar and sonar systems [31], and channel estimation in wireless communications [32]. The data acquisition devices, however, are often limited by hardware and physical constraints, precluding sampling with the desired resolution. It is thus of paramount interest to reduce sensing complexity while retaining recovery accuracy.

Our goal is then to recover a spectrally sparse signal from a small number of low-rate time domain samples, sometimes referred to as *spectral compressed sensing*. Specifically, the signal of interest $x(\boldsymbol{t})$ with ambient dimension n is assumed to be a weighted superposition of multi-dimensional sinusoids at r distinct frequencies, where the underlying frequencies can assume any *continuous value* over a unit disk. Conventional estimation approaches either ignore structural sparsity, or impose a discrete Fourier dictionary that results in basis mismatch issues, as elaborated below.

1.2.1.1 Motivation and Background

Harmonic Retrieval. Spectral compressed sensing is closely related to the problem of *harmonic retrieval*, which seeks to extract the underlying frequencies of a signal from a collection of its time domain samples. Conventional methods for harmonic retrieval include MUSIC [33], ESPRIT [34], the matrix pencil method [35], and the finite rate of innovation approach [36, 37]. These methods routinely exploit the *shift invariance* of the harmonic structure, namely, a consecutive segment of time domain samples lies in the same subspace irrespective of the starting point of the segment. The returned spectrum is then covariant to time delays, consistent with the physical principles underlying this problem. However, one weakness of these techniques is that they require prior knowledge of the model order and the noise spectra, and are often sensitive to noise and outliers [38]. More importantly, while these techniques capture the harmonic invariance structure of the problem, they are unable to automatically enforce sparsity (or compressibility) based on the sample complexity.

Compressed Sensing and Basis Mismatch. Another line of work is concerned with Compressed Sensing (CS) [1, 23, 39] over a discrete domain, which aims to recover a sparse solution from highly incomplete random measurements. These algorithms based on convex relaxation have received much attention due to their computational efficiency and robustness against noise [40, 41]. Furthermore, they do not require prior information on the model order. Nevertheless, the success of CS relies on sparse representation of the signal over a finite discrete dictionary, while the true parameters in many applications can only be specified in a *continuous* dictionary. The *basis mismatch* between the true frequencies and the discretized grid [42] results in loss of sparsity due to spectral leakage along the Dirichlet kernel, and hence degeneration in the performance of conventional CS paradigms. That said, these paradigms do not capture sparsity in a way that is compatible with the shift-invariant structure specific to the problem.

In short, traditional harmonic retrieval methods enforce physically meaningful constraints on estimates, but ignore sparsity; conventional compressed sensing paradigms enforce sparsity, but ignore physically meaningful constraints. In this work, we aim to bridge the gap by combining sparsity with physically meaningful structural constraints.

1.2.1.2 Main Contributions

We develop an algorithm, called *Enhanced Matrix Completion (EMaC)*, that simultaneously exploits the shift invariance of harmonic structures and the spectral sparsity of signals. Inspired by the conventional matrix pencil form [43], EMaC starts by arranging the data samples into an enhanced matrix exhibiting multi-fold Hankel structures, whose rank is bounded above by the spectral sparsity r. This way we transform the spectral sparsity into low-rank structures without imposing any pre-determined grid. EMaC then invokes a nuclear norm minimization program to complete missing entries from partially observed samples.

The performance of EMaC depends on an incoherence condition that depends only on the frequency locations regardless of the amplitudes of their respective coefficients. The incoherence measure is characterized by the reciprocal of the smallest singular value of some Gram matrix, which is defined by sampling the *Dirichlet kernel* associated with all frequency pairs. The signal of interest is said to obey the incoherence condition if this Gram matrix is well conditioned, which arises over a broad class of spectrally sparse signals including (but not restricted to) signals with well-separated frequencies. We demonstrate that, under such mild incoherence conditions, EMaC enables exact recovery from $\mathcal{O}(r\log^4 n)$ random samples, which is nearly information-theoretic optimal (Theorem 4.1).

Along the way, we provide theoretical guarantees for low-rank matrix completion of Hankel matrices (Theorem 4.2), which is of great importance in control, natural language processing, and computer vision. To the best of our knowledge, our results provide the first theoretical guarantees for Hankel matrix completion that are close to the information theoretic limit.

1.2.2 Covariance Estimation from Quadratic Measurements

Accurate estimation of second-order statistics of stochastic processes and data streams is of ever-growing importance to various applications that exhibit high dimensionality. Covariance estimation is the cornerstone of modern statistical analysis and information processing, as the covariance matrix constitutes the sufficient statistics to many signal processing tasks, and is particularly crucial for extracting reduced-dimension representation of the objects under study. For signals and data streams of high dimensionality, there might be limited memory and computation power available at the data acquisition devices, which requires the covariance estimation task to be performed with minimal stored measurements and low computational complexity. This is not possible unless appropriate structural assumptions are incorporated into the high-dimensional problems. Fortunately, a broad class of high-dimensional signals indeed possesses low-dimensional structures; some of them are listed as follows.

- *Low Rank:* The covariance matrix is (approximately) low-rank, which occurs when a small number of components explain most of the variability in the data. Low-rank covariance matrices arise in applications including array signal processing, collaborative filtering, and metric learning.

- *Stationarity and Low Rank:* The covariance matrix is simultaneously low-rank and Toeplitz, which arises when the random process is generated by a few spectral spikes. Recovery of the stationary covariance matrix, often equivalent to spectral estimation, is crucial in many tasks in wireless communications (e.g. detecting spectral holes in cognitive radio networks), and array signal processing (e.g. direction of arrival analysis).

- *Joint Sparsity and Rank-One:* The covariance matrix can be approximated in a jointly sparse and rank-one matrix. This has received much attention in recent development of sparse PCA, and is closely related to sparse signal recovery from magnitude measurements (called sparse phase retrieval).

In this thesis, we explore covariance estimation from a special type of measurements $\boldsymbol{y} := \{y_i\}_{i=1}^m$ taking the form

$$y_i = \boldsymbol{a}_i^\top \boldsymbol{\Sigma} \boldsymbol{a}_i + \boldsymbol{\eta}_i, \quad i = 1, \ldots, m, \tag{1.1}$$

where $\boldsymbol{a}_i \in \mathbb{R}^n$ represents the sensing vector, $\boldsymbol{\eta} := \{\boldsymbol{\eta}_i\}_{i=1}^m$ stands for the noise term, and m denotes the number of measurements. The noise-free measurements $\boldsymbol{a}_i^\top \boldsymbol{\Sigma} \boldsymbol{a}_i$'s are quadratic in $\boldsymbol{a}_i$ and are therefore referred to as *quadratic measurements* (or *rank-one measurements*). In practice, the number of measurements one can obtain is constrained by the storage requirement at the sensors, which could be much smaller than the ambient dimension of $\boldsymbol{\Sigma}$. This sampling scheme finds applications in a wide spectrum of practical scenarios, admits optimal covariance estimation with tractable algorithms, and brings in computational and storage advantages in comparison with other types of measurements, as detailed next.

1.2.2.1 Motivation: Why Quadratic Measurements?

Quadratic measurements (1.1) are motivated by several practical scenarios including high-frequency engineering applications and data stream processing as follows.

Noncoherent Energy Measurements in Wireless Environments When communication takes place in the high-frequency regime, empirical *energy measurement* is often more accurate than the phase measurement. For instance, energy measurements are more reliable when communication systems are operating in the extremely high carrier frequency (e.g. 60GHz [44]) regime.

- **Spectral Estimation from Energy Measurements**. Many wireless communication tasks in stochastic environments rely on reliable estimation of the spectral characteristics of random processes [45]. For instance, optimal signal transmissions are often based on the Karhunen–Loeve decomposition of a random process, which requires accurate covariance estimation [3]. If we employ a sampling vector $\boldsymbol{a}_i$ and observe empirical energy measurements over N instances $\{\boldsymbol{x}_t\}_{1\leq t\leq N}$, then the energy measurement are given by

$$y_i = \frac{1}{N}\sum_{t=1}^{N}\left|\boldsymbol{a}_i^\top \boldsymbol{x}_t\right|^2 = \boldsymbol{a}_i^\top \boldsymbol{\Sigma}_N \boldsymbol{a}_i, \quad i = 1, \ldots, m \tag{1.2}$$

  where $\boldsymbol{\Sigma}_N := \frac{1}{N}\sum_{t=1}^{N}\boldsymbol{x}_t\boldsymbol{x}_t^\top$ denotes the sample covariance matrix.

- **Noncoherent Subspace Detection from Energy Measurements**. Matched subspace detection [46] finds many applications in wireless communication, radar, and pattern recognition when the transmitted signal is encoded by the subspaces. Our problem can be cast as recovering the principal subspace of a stream $\{\boldsymbol{x}_t\}_{t=1}^{N}$, with an energy detector obtaining measurements in the form of (1.2). Thus, noncoherent subspace detection is subsumed by the formulation (1.1).

Phaseless Measurements in Physics Optical imaging devices are incapable of acquiring phase measurements due to ultra-high frequencies associated with light. In many applications, measurements in the form of (1.1) arise naturally.

- **Compressive Phase Space Tomography.** Phase Space Tomography [47] is an appealing method to measure the correlation function of a wave field in physics. However, tomography becomes challenging when the correlation matrix has large dimensions. Recently, it was proposed experimentally in [48] to recover an approximately low-rank correlation matrix, which often holds in physics, by only taking a small number of measurements in the form of (1.1).

- **Phase Retrieval.** Due to the physical constraints, one can only measure the Fourier intensity of an optical object. This gives rise to the problem of recovering a signal $\boldsymbol{x}$ from magnitude measurements, referred to as *phase retrieval* [49, 50, 51, 52, 53]. Several algorithms (e.g. [54, 55, 56]) have been proposed that enable exact phase retrieval (i.e. recovers $\boldsymbol{x} \cdot \boldsymbol{x}^\top$) from random magnitude measurements. If we set $\boldsymbol{\Sigma} := \boldsymbol{x}\boldsymbol{x}^\top$, then our formulation (1.1) subsumes phase retrieval as a special case.

Covariance Sketching for Data Streams A high-rate data stream challenges information processing tasks in storage, computations, and reasoning [57]. It is desirable to extract the covariance matrix of the data instances from one pass of the inputs. Interestingly, the quadratic measurement strategy can be leveraged as an effective data stream processing method to extract the covariance information from real-time data, with limited memory and low computational complexity.

Specifically, consider an input stream $\{\boldsymbol{x}_t\}_{t=1}^{\infty}$ that arrives sequentially, where each $\boldsymbol{x}_t \in \mathbb{R}^n$ is a data instance generated at time t. The goal is to estimate the covariance matrix $\boldsymbol{\Sigma} = \mathbb{E}[\boldsymbol{x}_t\boldsymbol{x}_t^\top]$. The prohibitively high data arrival rate forces covariance extraction to function with as small a memory as possible. The scenario we consider is quite general, and we only impose that the covariance of a random substream of the original data stream converges to the same covariance as[1] $\boldsymbol{\Sigma}$.

We propose to pool the data stream $\{\boldsymbol{x}_t\}_{t=1}^{\infty}$ into a few stored measurements in an easy-to-adapt fashion with a collection of sketching vectors $\{\boldsymbol{a}_i\}_{i=1}^{m}$. Our covariance sketching method, termed *quadratic sketching*, is outlined as follows:

1. At each time t, we randomly choose a sketching vector indexed by $\ell_t \in \{1, \dots, m\}$, and obtain a single linear *sketch* $\boldsymbol{a}_{\ell_t}^{\top}\boldsymbol{x}_t$.

2. All sketches employing the same sketching vector $\boldsymbol{a}_i$ are squared, aggregated and normalized, which converge *rapidly* to a *measurement*[2]

$$y_i = \mathbb{E}[(\boldsymbol{a}_i^{\top}\boldsymbol{x}_t)^2] + \boldsymbol{\eta}_i = \boldsymbol{a}_i^{\top}\mathbb{E}[\boldsymbol{x}_t\boldsymbol{x}_t^{\top}]\boldsymbol{a}_i + \boldsymbol{\eta}_i = \boldsymbol{a}_i^{\top}\boldsymbol{\Sigma}\boldsymbol{a}_i + \boldsymbol{\eta}_i, \tag{1.3}$$

   where $\boldsymbol{\eta} := \{\boldsymbol{\eta}_i\}_{i=1}^{m}$ denotes the noise term.

There are several benefits of this covariance sketching method. 1) As will be shown, this sketching scheme allows optimal covariance estimation with theoretically minimal memory complexity m at the sensing stage. 2) The computational cost for sketching each instance is linear. 3) Unlike the uncompressed sketching methods where each instance one measures usually affects many stored measurements, our scheme allows each aggregate quadratic sketch to be composed by completely different instances, which allows sketching to be performed in a distributed and asynchronous manner. This arises since each randomized sketch is a compressive snapshot of the second-order statistics, while each uncompressed measurement itself is unable to preserve the correlation information.

In summary, these applications require faithful covariance matrix estimation from *a small number* of rank-one measurements (1.1). We aim to develop tractable algorithms that enable covariance estimation with near-optimal performance guarantees.

[1] No prior information on the correlation statistics between consecutive instances is available (e.g. they are not necessarily independent), and hence it is not feasible to exploit such correlation / independence to reduce sample complexity.

[2] Note that we might only be able to obtain measurements for empirical covariance matrices instead of $\boldsymbol{\Sigma}$, but this inaccuracy can be absorbed into the noise term $\boldsymbol{\eta}$.

1.2.2.2 Main Contributions

We develop convex optimization algorithms for covariance estimation from a set of quadratic measurements as given in (1.1) for a variety of structural assumptions including low-rank, Toeplitz low-rank, and sparse rank-one covariance matrices. The proposed algorithms exploit the assumed low-dimensional structures using semidefinite programming (SDP), specifically trace norm minimization for promoting low-rank structure, and ℓ_1 norm minimization for promoting sparsity. For a large class of sub-Gaussian sensing vectors, we derive theoretical performance guarantees (Theorems 5.1 – 5.3) concerning the following aspects:

1. **Exact and universal recovery**: once the sensing vectors are selected, then with high probability, all covariance matrices satisfying the presumed structure can be recovered;

2. **Stable recovery**: the proposed algorithms allow covariance recovery with high accuracy even under imperfect structural assumptions; additionally, if the measurements are corrupted by noise, the estimate deviates from the true covariance matrix by at most a constant multiple of the noise level;

3. **Near-minimal memory complexity**: the proposed algorithms succeed as soon as the number of measurements exceeds the fundamental sampling limits for most of the respective structure. For the special case of (sparse) rank-one matrices, our result recovers and strengthens the best-known recovery guarantees of (sparse) phase retrieval using PhaseLift [58, 59] with much simpler proofs.

Secondly, to establish some of the above theoretical guarantees (Theorems 5.1 and 5.3), we introduce a novel mixed-norm restricted isometry property, denoted by RIP-ℓ_2/ℓ_1. A measurement operator is said to satisfy the RIP-ℓ_2/ℓ_1 if the strength of the signal before and after measurements are preserved when measured in the ℓ_2 norm and in the ℓ_1 norm, respectively. While the quadratic sensing operator does not satisfy the conventional RIP-ℓ_2/ℓ_2 with respect to general low-rank structures [54], it does satisfy the RIP-ℓ_2/ℓ_1 after a small "debiasing" modification for the general low-rank, as well as jointly sparse and rank-one structural assumptions. This seemingly subtle change allows us to develop a significantly simpler approach without resorting to complicated dual construction as in [54, 58, 59]. On the other hand, we demonstrate, via the entropy method [60], that appropriate linear mixtures of quadratic measurements satisfy RIP-ℓ_2/ℓ_2 regarding *Toeplitz* low-rank structure. Along the way, we have also established a RIP-ℓ_2/ℓ_2 for bounded and near-isometric operators (Theorem 5.4), which strengthens previous work [61, 62] by offering universal and stable recovery guarantees for a broader class of operators including Fourier-type measurements.

1.2.2.3 Comparison to Prior Works

In most existing work, the covariance matrix is estimated from a collection of *full* data samples, and fundamental guarantees have been derived on how many samples are sufficient to approximate the ground truth [63, 64]. In contrast, this work is motivated by the success of CS [1, 24], which asserts that compression can be achieved at the same time as sensing without losing information. As we will show, covariance estimation from compressive measurements can be highly robust. Another work [65] proposed estimating sparse covariance matrices from measurements of the form $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^\top$, where the sketching matrix $\boldsymbol{A}$ is constructed from expander graphs. Nevertheless, this scheme cannot accommodate low-rank covariance estimation.

Our covariance estimation method is inspired by recent developments in phase retrieval [54, 58, 51, 55, 56], which is equivalent to recovering rank-one covariance matrices from quadratic measurements. In particular, our recovery algorithm coincides with *PhaseLift* [54, 58] when applied to low-rank matrices. In [58], it is shown that PhaseLift succeeds at reconstructing a signal of dimensionality n from $\Theta(n)$ phaseless *Gaussian* measurements. When specializing our result to this case, we have shown that the same type of theoretical guarantee holds for a much larger class of *sub-Gaussian* measurements, with a much simpler proof.

Finally, our problem is related to online principal component analysis (PCA), which has been an active area for decades [66], usually assuming full data samples. Inspired by CS, subspace tracking from partial observations of a data stream [67, 68] has also been explored, which is a variant of online PCA with missing data. However, existing subspace tracking methods mainly aim to recover the entire data stream, which is not necessary if one only wishes to estimate the second-order statistics.

1.3 Organization and Published Materials

The remainder of this thesis is organized as follows.

Chapter 2 This chapter explores the sampled channel capacity when the sampler is designed and optimized based on the instantaneous CSI. We start in Section 2.1 with channel and sampler models. Section 2.2 studies the sampled capacity under general periodic sampling systems, which subsume as special cases three canonical sampling mechanisms: 1) a filter followed by uniform sampling; 2) filter-bank sampling; and 3) sampling with periodic modulation and filter banks. We then characterize in Section 2.3 the fundamental upper limit on the sampled channel capacity over a very general class of time-preserving sampling systems, as well as the optimality of filter-bank sampling. These results have been published in [69, 70].

Chapter 3 This chapter studies the minimax capacity under channel-independent sampling. In Section 2.1, we introduce our system model of compound multiband channels, and the capacity loss metrics. We then determine in Section 3.2 the minimax capacity loss in terms of band sparsity ratio and undersampling factors. Specifically, we develop a lower bound on the minimax capacity loss in Section 3.2.1, and establish the optimality of random sampling for Landau-rate and super-Landau sampling in Sections 3.2.2 and 3.2.3, respectively. Along the way we derive concentration of measure of several log-determinant functions in Section 3.3.4. The results have been presented in [71, 72].

Chapter 4 This chapter addresses the basis mismatch issue in the spectral compressed sensing problem. In Section 4.1, we first describe the signal and sampling models, and then present the enhanced matrix form and the associated structured matrix completion algorithms. The main theoretical guarantees are summarized in

Section 4.2. Section 4.3 presents the numerical validation of our algorithms. The proofs of the main theorems are based on duality analysis, which are outlined in Section 4.4. These results have been published in [73, 74].

Chapter 5 This chapter studies covariance estimation algorithms under quadratic measurements. We first present the formal setup in Section 5.1, followed by convex optimization algorithms and their theoretical guarantees in Section 5.2. The analysis framework is based upon a novel mixed-norm restricted isometry property as well as conventional RIP for near-isotropic and bounded measurements, as elaborated in Sections 5.3 and 5.4. Numerical examples are provided in Sections 5.5. The results here have been presented in [75].

Chapter 6 This chapter closes the paper with a short summary of our findings and potential future directions.

Part I

Fundamental Limits: Capacity of Sub-Sampled Communication Channels

Chapter 2

Capacity under Channel-Dependent Sub-Nyquist Sampling

We consider the maximum information rate that can be conveyed through a continuous-time linear time-invariant (LTI) channel subject to a sub-Nyquist sampling rate constraint. In particular, this chapter explores the scenario where the sampling structure is designed based on perfect CSI, and characterizes the fundamental capacity limits of sampled Gaussian channels under a very general class of sampling strategies.

2.1 Problem Formulation

2.1.1 General Channel and Sampling System Models

We consider the continuous-time additive Gaussian channel, where the channel is modeled as an LTI filter with impulse response $h(t)$ and frequency response $H(f) = \int_{-\infty}^{\infty} h(t)\exp(-j2\pi ft)\mathrm{d}t$. With $x(t)$ denoting the transmitted signal, the analog channel output is given as

$$r(t) = h(t) * x(t) + \eta(t). \tag{2.1}$$

where the noise process $\eta(t)$ is assumed to be a stationary zero-mean Gaussian process with power spectral density $\mathcal{S}_\eta(f)$. We also define $s_\eta(t) := \mathcal{F}^{-1}\left(\sqrt{\mathcal{S}_\eta(f)}\right)$. Unless otherwise specified, we assume throughout that *perfect CSI* (i.e. the knowledge of both $H(f)$ and $\mathcal{S}_\eta(f)$) is available at both the transmitter and the receiver.

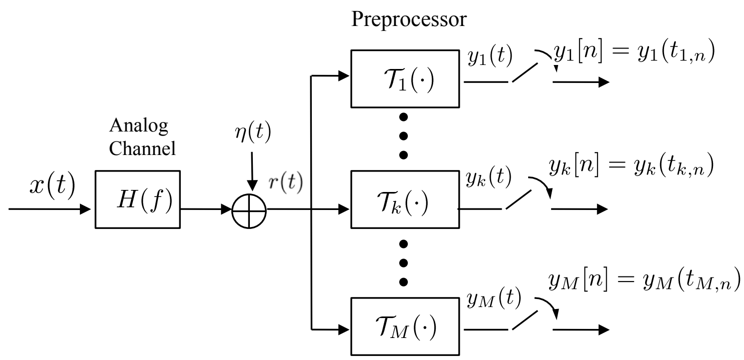


Figure 2.1: The input $x(t)$ is passed through the analog channel and contaminated by the noise $\eta(t)$. The analog channel output $r(t)$ is then passed through M linear preprocessors and sampled on the sampling sets $\Lambda_i = \{t_{i,n} \mid n \in \mathbb{Z}\}$ $(1 \leq i \leq M)$.

The analog channel output $r(t)$ is passed through M $(1 \leq M \leq \infty)$ branches of linear preprocessing systems, each followed by a pointwise sampler, as illustrated in Fig. 2.1. The preprocessed output $y_i(t)$ at the ith branch is obtained by applying a linear bounded operator $\mathcal{T}_i$ to the channel output $r(t)$:

$$y_i(t) = \mathcal{T}_i(r(t)) = \int q_i(t,\tau)\, r(\tau)\, \mathrm{d}\tau, \tag{2.2}$$

where $q_i(t,\tau)$ denotes the impulse response of the time-varying system represented by $\mathcal{T}_i$, i.e. the output seen at time t due to an impulse in the input at time τ. The linear operator $\mathcal{T}_i$ subsumes filtering and modulation as special cases. When an operator $\mathcal{T}$ is LTI, we use $q(\tau) := q(t, t-\tau)$ for brevity.

The pointwise sampler following the preprocessor can be either uniform or irregular [7]. Specifically, the preprocessed output $y_i(t)$ (at the ith branch) is sampled at times $t_{i,n}$ $(n \in \mathbb{Z})$, yielding a sample sequence $y_i[n] = y_i(t_{i,n})$. Here, we define the *sampling* *set* Λ_i at the ith branch as

$$\Lambda_i := \{t_{i,n} \mid n \in \mathbb{Z}\}\,. \tag{2.3}$$

In particular, if $t_{i,n} = nT_{i,s}$, then the sampling set at the ith branch is said to be uniform with period $T_{i,s}$.

2.1.2 Periodic Sampling Systems

While the sampling systems employed might not be time-invariant, most sampling systems applied in practice enjoy certain block-wise time invariance. These can be categorized as *periodic sampling*, defined and discussed as follows.

Definition 2.1 (**Periodic System).** *A linear preprocessing system is said to be periodic with period T_q if its impulse response $q(t,\tau)$ satisfies*

$$q(t,\tau) = q(t + T_q, \tau + T_q), \quad \forall t,\tau \in \mathbb{R}. \tag{2.4}$$

Definition 2.2 (**Periodic Sampling).** *Consider a sampling system with a preprocessor of impulse response $q(t,\tau)$ followed by a sampling set $\Lambda = \{t_k \mid k \in \mathbb{Z}\}$. A linear sampling system is said to be periodic with period T_q and sampling rate f_s ($f_sT_q \in \mathbb{Z}$) if the preprocessor is periodic with period T_q and the sampling set satisfies*

$$t_{k+f_sT_q} = t_k + T_q, \quad \forall k \in \mathbb{Z}. \tag{2.5}$$

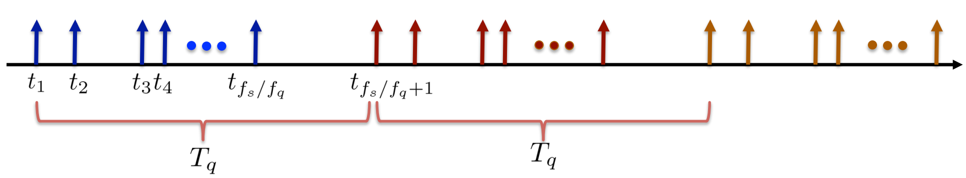


Figure 2.2: The sampling set of a periodic sampling system with period $1/f_q$ and sampling rate f_s.

In short, a periodic sampling system consists of a periodic preprocessor followed by a pointwise sampler with a periodic sampling set, as illustrated in Fig. 2.2. Since the impulse response can be arbitrary within a period, this allows us to model multibranch sampling methods with each branch using the same sampling rate. Periodic sampling schemes subsume as special cases a broad class of sampling techniques, e.g. uniform sampling with filter banks, and sampling following periodic modulation.

2.1.3 Sampling Rate Definition

Our metric of interest is the maximum information rate under a sampling rate constraint, which requires a precise sampling-rate definition. For an M-branch sampling system where each branch is uniformly sampled with interval $T_{i,s}$, the sampling rate can be naturally defined as the aggregate sampling rate, i.e. $f_s = \sum_{i=1}^{M} T_{i,s}^{-1}$; for a periodic system, the sampling rate has also been defined as the average aggregate sampling density. However, the sampling rate becomes more complicated to characterize when considering general (nonuniform) sampling sets. In order to make a formal definition of sampling rate, we adopt the notion of Beurling density [76], which is commonly used in sampling theory (e.g. [7]) for the density of a sampling set Λ.

Definition 2.3 (Beurling Density). *For a sampling set $\Lambda = \{t_k \mid k \in \mathbb{Z}\}$, the upper and lower Beurling density are given respectively as*

$$D^{+}(\Lambda) = \lim_{r\to\infty} \sup_{z\in\mathbb{R}} \frac{\mathrm{card}\,(\Lambda \cap [z, z+r])}{r}; \quad D^{-}(\Lambda) = \lim_{r\to\infty} \inf_{z\in\mathbb{R}} \frac{\mathrm{card}\,(\Lambda \cap [z, z+r])}{r}.$$

When $D^{+}(\Lambda) = D^{-}(\Lambda)$, Λ is said to have uniform Beurling density $D(\Lambda) := D^{-}(\Lambda)$.

When the sampling set is uniform with period T_s, the Beurling density reduces to $D(\Lambda) = T_s^{-1}$, which coincides with the conventional definition of sampling rate. The notion of Beurling density allows the Shannon-Nyquist sampling theorem to be extended to nonuniform sampling. Specifically, for each $\Lambda = \{t_i \mid n \in \mathbb{Z}\}$, the set of exponential functions $\{\exp(j2\pi t_n f) \mid n \in \mathbb{Z}\}$ forms a *non-harmonic* Fourier series [8]. Whether a class of signals is recoverable from the nonuniform sampled sequence is determined by the completeness of the associated non-harmonic set. For the set $\mathcal{B}_{\Omega}$ of signals supported on the spectral set Ω, Landau [13] asserts that a sampling set Λ

(without preprocessing) allows perfect signal reconstruction only if $D^{-}(\Lambda) \geq \mu(\Omega)$. The measure $\mu(\Omega)$ (i.e. the spectral occupancy) is termed the *Landau rate*.

We will therefore use Beurling density to characterize the sampling rate for a general sampling set. However, since the preprocessor might distort the time scale of the input, the resulting "sampling rate" might not capture the true sampling rate applied to the signal, as illustrated in the following example.

Example 2.1 (Compressor)**.** *Consider a preprocessor defined by the relation*

$$y(t) = \mathcal{T}(r(t)) = r(Lt)$$

with $L \geq 2$. If we apply a uniform set $\Lambda = \{t_n : t_n = n/f_s\}$ on the preprocessed output $y(t)$, the sampled sequence with "sampling rate" f_s is given by

$$y[n] = y(n/f_s) = r(nL/f_s),$$

which corresponds to sampling the system input $r(t)$ at rate f_s/L. The compressor effectively time-warps the signal, thus resulting in a mismatch of the time scales between the input and output.

The compressor example illustrates that the notion of Beurling rate may be misleading for preprocessors that result in time warping. Hence, we will focus only on sampling that preserves time scales. One class of systems that preserves time scales are modulation operators ($y(t) = p(t)x(t)$), which perform pointwise scaling of the input, and hence do not change the time scale. A more general class of systems that preserve the time scale can be generated by connecting multiple modulation and periodic subsystems in parallel or in serial, as stated below.

Definition 2.4 (Time-preserving System). *Given an index set $\mathcal{I}$, a preprocessing system $\mathcal{T} : x(t) \mapsto \{y_k(t), k \in \mathcal{I}\}$ is said to be time-preserving if*

(1) The system input is passed through $|\mathcal{I}|$ branches of linear preprocessors, yielding a set of analog outputs $\{y_k(t) \mid k \in \mathcal{I}\}$.

(2) In each branch, the preprocessor comprises a set of periodic or modulation operators connected in serial.

With a preprocessing system that preserves the time scale, we can now define the aggregate sampling rate through the Beurling density.

Definition 2.5 (Sampling Rate for Time-preserving Systems). *A sampling system is said to be time-preserving with sampling rate f_s if*

(1) Its preprocessing system $\mathcal{T}$ is time-preserving.

(2) The preprocessed output $y_k(t)$ is sampled by a sampling set $\Lambda_k = \{t_{l,k} \mid l \in \mathbb{Z}\}$ with a uniform Beurling density $f_{k,s}$, which satisfies $\sum_{k\in\mathcal{I}} f_{k,s} = f_s$.

Finally, we note that for each multibranch sampling system, one can easily generate a single-branch sampling system that yields the same set of sampled output values. This fact will allow us to simplify the analysis.

2.1.4 Capacity Definition

There are two capacity definitions that are of interest in sub-Nyquist sampled channels: (1) the sampled capacity for a given sampling system; (2) the capacity for a large class of sampling systems under a sampling rate constraint, detailed below.

Suppose that the transmit signal $x(t)$ is constrained to the time interval $[-T, T]$, and the received signal is sampled and observed over $[-T, T]$. For a *given* sampler $\mathcal{P}$ and a given time duration T, we define the information metric $C_T^{\mathcal{P}}(P)$ to be

$$C_T^{\mathcal{P}}(P) = \sup \frac{1}{2T} I\left(x\left([-T, T]\right), \{\boldsymbol{y}[t_n]\}_{[-T,T]}\right), \tag{2.6}$$

where the supremum is over all input distributions subject to a power constraint $\mathbb{E}\left[\frac{1}{2T}\int_{-T}^{T}|x(t)|^2 \mathrm{d}t\right] \leq P$. Here, $\{\boldsymbol{y}[t_n]\}_{[-T,T]}$ denotes the set of samples obtained at times within $[-T, T]$ by the sampling system $\mathcal{P}$.

Similar to [3, Chapter 8], the capacity of the undersampled channel under a given sampling system can then be studied by passing to the limit $T \to \infty$ as follows.

Definition 2.6. *$C^{\mathcal{P}}$ is said to be the information capacity of a given sampled analog channel (or* sampled channel capacity*) if $\lim_{T\to\infty} C_T^{\mathcal{P}}(P)$ exists and*

$$C^{\mathcal{P}}(P) = \lim_{T\to\infty} C_T^{\mathcal{P}}(P).$$

Note that any sampled channel can be converted to a set of independent discrete channels via a Karhunen–Loève decomposition. The metric $C^{\mathcal{P}}(P)$ then quantifies the maximum information rate that can be supported reliably by these channels.

The above capacity is defined for a given sampling mechanism. Another metric of interest is the maximum data rate achievable by all sampling schemes within a general class. This leads to the following capacity definition.

Definition 2.7 (**Sampled Capacity under a Class of Sampling Systems**). *$C_{\mathcal{A}}(f_s, P)$ is said to be the capacity of an analog channel over a class $\mathcal{A}$ of sampling systems under a given sampling rate f_s if*

$$C_{\mathcal{A}}(f_s, P) = \sup_{\mathcal{P} \in \mathcal{A}} C^{\mathcal{P}}(P).$$

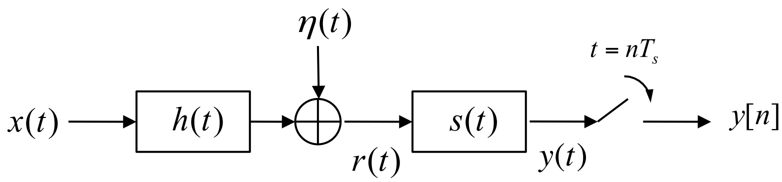
The above definition of sampled capacity characterizes the capacity of an analog channel over a large set of sampling mechanisms subject to a sampling rate constraint. This gives rise to the natural problem of jointly optimizing the input and sampling architectures to maximize capacity, a goal we aim to address in this chapter.

2.2 Capacity under Periodic Sampling Systems

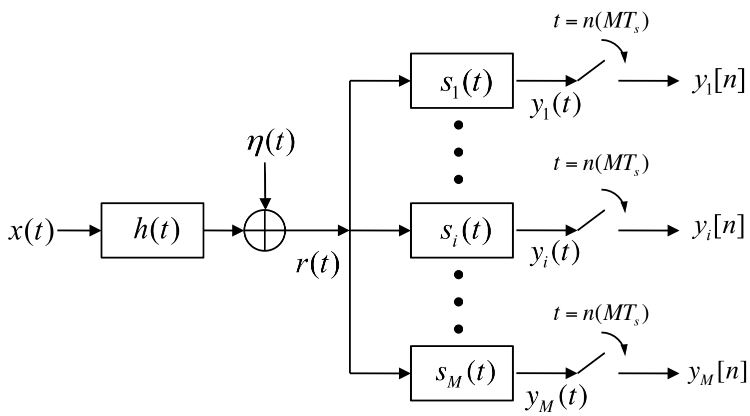
In this section, we first explore three classes of sampling strategies with increasing complexity, which are the most commonly used in practice. It turns out that all these structures are special cases of the class of general periodic sampling systems, which leads to a unified framework in studying all these sampling mechanisms.

2.2.1 Three Canonical Sampling Structures

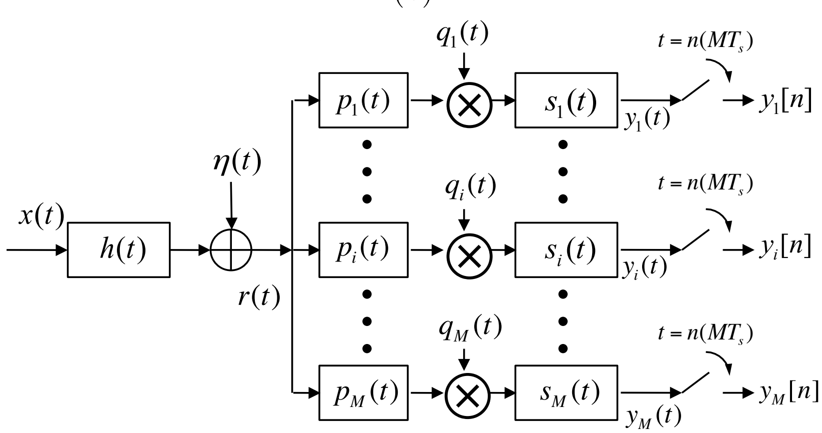
We start from sampling following a single filter, and incrementally extend our results to incorporate filter banks and modulation banks.



(a)



(b)



(c)

Figure 2.3: (a) Filtering followed by uniform sampling; (b) A filter bank followed by uniform sampling; (c) Modulation and filter banks followed by uniform sampling.

2.2.1.1 A Filter Followed by Uniform Sampling

Ideal uniform sampling is performed by sampling the analog signal uniformly at a rate $f_s = T_s^{-1}$. In order to avoid aliasing, suppress out-of-band noise, and compensate for linear distortion, a prefilter is often added prior to the sampler [77]. Our sampling process thus includes a general prefilter, as illustrated in Fig. 2.3(a). Specifically, we prefilter the received analog signal with an LTI filter having impulse response $s(t)$ and frequency response $S(f)$. The filtered output is observed over $[-T, T]$ as

$$y(t) = s(t) * (h(t) * x(t) + \eta(t)), \quad t \in [-T, T]. \tag{2.7}$$

We then sample $y(t)$ uniformly, leading to the sampled sequence $y[n] = y(nT_s)$.

The sampled channel capacity under sampling with filtering is stated as follows.

Theorem 2.1. *Consider the system shown Fig. 2.3(a). Assume that $h(t)$, $s(t)$, $S(f)\sqrt{\mathcal{S}_\eta(f)}$ are continuous, bounded and absolutely Riemann integrable, and that $\sum_{l\in\mathbb{Z}} |S(f - lf_s)|^2 \mathcal{S}_\eta(f - lf_s)$ is uniformly bounded away from 0. The capacity $C(P)$ of the sampled channel with a power constraint P is then given parametrically as*

$$C(P) = \frac{1}{2}\int_{-f_s/2}^{f_s/2} \log^+ \left(\nu\gamma^{\mathrm{s}}(f)\right) \mathrm{d}f, \tag{2.8}$$

where ν satisfies

$$\int_{-f_s/2}^{f_s/2} \left[\nu - 1/\gamma^{\mathrm{s}}(f)\right]^+ \mathrm{d}f = P. \tag{2.9}$$

Here, we denote $\gamma^{\mathrm{s}}(f) := \frac{\sum_{l\in\mathbb{Z}}|H(f-lf_s)S(f-lf_s)|^2}{\sum_{l\in\mathbb{Z}}|S(f-lf_s)|^2\mathcal{S}_\eta(f-lf_s)}$.

As expected, applying the prefilter modifies the channel gain and colors the noise accordingly. The color of the noise is reflected in the denominator term of $\gamma^{\mathrm{s}}(f)$ at each $f \in [-f_s/2, f_s/2]$ within the sampling bandwidth. The channel and prefilter response leads to an equivalent frequency-selective channel, and the ideal uniform sampling that follows generates a folded version of the non-sampled channel capacity. Water filling over $1/\gamma^{\mathrm{s}}(f)$ thus determines the optimal power allocation.

Since different prefilters lead to different channel capacities, a natural question boils down to which filter maximizes capacity. This is identified as follows.

Corollary 2.1. *Consider the system in Fig. 2.3(a), and define* $\gamma_l(f) := \frac{|H(f-lf_s)|^2}{\mathcal{S}_\eta(f-lf_s)}$ *for any* $l \in \mathbb{Z}$*. Suppose that for each* f*, there exists* k *such that* $\gamma_k(f) = \sup_{l\in\mathbb{Z}} \gamma_l(f)$*. Then the capacity in (2.8) is maximized by the filter with frequency response*

$$\forall f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right], \quad S(f-kf_s) = \begin{cases} 1, & \text{if } \gamma_k(f) = \sup_{l\in\mathbb{Z}} \gamma_l(f), \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

Proof. This is a special case of Corollary 2.2 presented later. □

The optimal prefilter puts all its mass in those frequencies with the highest SNR within each aliased set $\{f - lf_s \mid l \in \mathbb{Z}\}$. One interesting observation is that optimal prefiltering equivalently generates an *alias-free* channel. After passing through an optimal prefilter, all frequencies modulo f_s except the one with the highest SNR are removed, and hence the optimal prefilter suppresses aliasing and out-of-band noise. This alias-suppressing phenomena, while different from many sub-Nyquist works that advocate mixing instead of alias suppressing [78], arises from the fact that the sampler is designed based on perfect CSI.

2.2.1.2 A Bank of Filters Followed by Uniform Sampling

Sampling following a single filter often falls short of exploiting channel structure. In particular, although Nyquist-rate uniform sampling preserves information for bandlimited signals, for multiband signals it does not ensure perfect reconstruction at the Landau rate, since uniform sampling at sub-Nyquist rates may suppress information by collapsing subbands. This motivates the exploration of more complex sampling mechanisms. We begin with a popular class of sampling mechanisms where the received signal is preprocessed by a filter bank. Since filters may introduce delays, this approach subsumes that of a filter bank with different sampling times at each branch.

Specifically, we employ a bank of M filters each followed by uniform sampling at rate f_s/M, as illustrated in Fig. 2.3(b). We denote by $s_i(t)$ and $S_i(f)$ the impulse and frequency response of the ith LTI filter, respectively. The filtered output obeys

$$y_i(t) = (h(t) * s_i(t)) * x(t) + s_i(t) * \eta(t), \quad t \in [-T, T], \quad 1 \leq i \leq M. \tag{2.11}$$

These filtered signals are then sampled uniformly to yield $y_i[n] \overset{\Delta}{=} y_i(nMT_s)$.

For notational simplicity, we introduce a matrix $\boldsymbol{F}_s \in \mathbb{C}^{m\times\infty}$ and an infinite diagonal matrix $\boldsymbol{F}_h$ in the Fourier domain such that for each $1 \leq i \leq M$ and $l \in \mathbb{Z}$,

$$(\boldsymbol{F}_s(f))_{i,l} = S_i\left(f - \frac{lf_s}{M}\right)\sqrt{\mathcal{S}_\eta\left(f - \frac{lf_s}{M}\right)}; \qquad (\boldsymbol{F}_h(f))_{l,l} = \frac{H\left(f - \frac{lf_s}{M}\right)}{\sqrt{\mathcal{S}_\eta\left(f - \frac{lf_s}{M}\right)}}.$$

Theorem 2.2. *Consider the system shown in Fig. 2.3(b). Assume that $h(t)$ and $s_i(t)$ $(1 \leq i \leq M)$ are all continuous, bounded and absolutely Riemann integrable. Additionally, assume that $\lambda_{\min}(\boldsymbol{F}_s\boldsymbol{F}_s)$ is uniformly bounded away from 0. The capacity $C(P)$ of the sampled channel with a power constraint P is given by*

$$C(P) = \int_0^{f_s/M} \frac{1}{2}\sum_{i=1}^{M} \log^+\left(\nu\lambda_i\left(\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\boldsymbol{F}_s\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_s^*\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\right)\right)\mathrm{d}f,$$

where ν is determined by

$$\int_0^{f_s/M}\sum_{i=1}^{M}\left[\nu - \lambda_i^{-1}\left(\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\boldsymbol{F}_s\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_s^*\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\right)\right]^+\mathrm{d}f = P.$$

Remark 2.1. We can express this capacity alternatively as

$$C(P) = \max_{\boldsymbol{Q}\in\mathcal{Q}}\int_0^{f_s/M}\frac{1}{2}\log\det\left(\boldsymbol{I} + \left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\boldsymbol{F}_s\boldsymbol{F}_h\boldsymbol{Q}\boldsymbol{F}_h^*\boldsymbol{F}_s^*\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\right)\mathrm{d}f, \tag{2.12}$$

where $\mathcal{Q} = \left\{\boldsymbol{Q}(f)\ (1 \leq f \leq \frac{f_s}{M}) \mid \boldsymbol{Q}(f) \in \mathbb{S}_+; \int_0^{f_s/M}\mathrm{tr}\left(\boldsymbol{Q}(f)\right)\mathrm{d}f = P\right\}$.

The optimal function $\boldsymbol{Q}$ corresponds to a water-filling power allocation strategy based on the singular values of the equivalent channel matrix $\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\boldsymbol{F}_s\boldsymbol{F}_h$, where $\boldsymbol{F}_h$ represents the original channel and $\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\boldsymbol{F}_s$ arises from prefiltering as well as noise whitening. For each $f \in [0, f_s/M]$, the integrand in (2.12) can be interpreted as a MIMO capacity formula. The channel capacity is achieved when the transmit signals are designed to decouple the associated MIMO channel into M parallel channels, each associated with one of its singular directions.

Interestingly, the capacity-optimizing filter bank admits a simple form as follows.

Corollary 2.2. *Consider the system shown in Fig. 2.3(b). Suppose that for each aliased set* $\left\{f - \frac{if_s}{M} \mid i \in \mathbb{Z}\right\}$ *and each* k $(1 \le k \le M)$*, there exists* $l \in \mathbb{Z}$ *such that* $\frac{\left|H\left(f-\frac{lf_s}{M}\right)\right|^2}{\mathcal{S}_\eta\left(f-\frac{lf_s}{M}\right)}$ *is equal to the* k^{th} *largest element in* $\left\{\frac{\left|H\left(f-\frac{if_s}{M}\right)\right|^2}{\mathcal{S}_\eta\left(f-\frac{if_s}{M}\right)} \mid i \in \mathbb{Z}\right\}$*. The capacity (2.12) is then maximized by a bank of filters with frequency responses*

$$\forall f \in \left[0, \frac{f_s}{M}\right], \quad S_k\left(f - \frac{lf_s}{M}\right) = \begin{cases} 1, & \text{if } \frac{\left|H\left(f-\frac{lf_s}{M}\right)\right|^2}{\mathcal{S}_\eta\left(f-\frac{lf_s}{M}\right)} = \lambda_k\left(\boldsymbol{F}_h(f)\boldsymbol{F}_h^*(f)\right); \\ 0, & \text{otherwise}, \end{cases} \tag{2.13}$$

for all $l \in \mathbb{Z}$*,* $1 \le k \le M$*. The resulting maximum channel capacity is given by*

$$C(f_s) = \frac{1}{2}\int_0^{f_s/M} \sum_{k=1}^{M} \log^+\left(\nu \cdot \lambda_k\left(\boldsymbol{F}_h\boldsymbol{F}_h^*\right)\right) \mathrm{d}f, \tag{2.14}$$

where ν *is chosen such that* $\int_0^{f_s/M} \sum_{k=1}^{M} \left[\nu - \lambda_k^{-1}\left(\boldsymbol{F}_h\boldsymbol{F}_h^*\right)\right]_+ \mathrm{d}f = P$*.*

The choice of prefilters in (2.13) achieves the upper bounds on all singular values, and is hence universally optimal regardless of the water level. Since $\left(\boldsymbol{F}_s\boldsymbol{F}_s^*\right)^{-\frac{1}{2}}\boldsymbol{F}_s$ has orthonormal rows, it acts as an orthogonal projection and outputs an M-dimensional subspace. Since $\boldsymbol{F}_h$ is diagonal, the subspace closest to the channel space spanned by $\boldsymbol{F}_h$ corresponds to the M SNR-maximizing rows of $\boldsymbol{F}_h$. The maximum information rate is then achieved when the filter bank outputs M frequencies with the highest SNR among the aliased set and suppresses the contents from all other frequencies.

2.2.1.3 Modulation and Filter Banks Followed by Uniform Sampling

We generalize the filter-bank sampling strategy by adding an additional filter bank and a modulation bank of M branches, which includes as special cases a broad class of nonuniform sampling methods that are applied in both theory and practice. Specifically, in the ith branch, the received channel output $r(t)$ is prefiltered by an LTI filter with impulse response $p_i(t)$ and frequency response $P_i(f)$, modulated by a periodic waveform $q_i(t)$ of period T_q, filtered by another LTI filter with impulse response $s_i(t)$ and frequency response $S_i(f)$, and then sampled uniformly at a rate f_s/M, as illustrated in Fig. 2.3(c). The first prefilter $P_i(f)$ will be useful in removing out-of-band noise, while the periodic waveforms scramble spectral contents from different aliased sets, thus bringing in more design flexibility. More precisely, the analog filtered output at the ith branch prior to uniform sampling is

$$y_i(t) = s_i(t) * \left(q_i(t) \cdot \left[p_i(t) * \left(h(t) * x(t) + \eta(t)\right)\right]\right), \quad 1 \le i \le M, \tag{2.15}$$

resulting in the digital sequence of samples $y_i[n] = y_i(nMT_s)$.

Suppose that $\tilde{T}_s := MT_s = \frac{b}{a} T_q$ where a and b are coprime integers, and that the Fourier transform of $q_i(t)$ is given by $\sum_l c_i^l \delta(f - l f_q)$. Before stating our theorem, we introduce the following two Fourier matrices $\boldsymbol{F}^\eta \in \mathbb{C}^{aM\times\infty}$ and $\boldsymbol{F}^h$. Specifically, $\boldsymbol{F}^\eta$ contains M submatrices with the αth submatrix given by an $a \times \infty$-dimensional matrix $\boldsymbol{F}^\eta_\alpha \boldsymbol{F}^p_\alpha$. Here, for any $v \in \mathbb{Z}$, $1 \le l \le a$, and $1 \le \alpha \le M$,

$$(\boldsymbol{F}^\eta_\alpha)_{l,v} := (\boldsymbol{F}^p_\alpha)_{v,v} \left[\sum_u c_\alpha^u S_\alpha \left(-f + u f_q + v\frac{f_q}{b}\right) \exp\left(-j2\pi l M T_s \left(f - u f_q - v\frac{f_q}{b}\right)\right)\right].$$

The matrices $\boldsymbol{F}^p_\alpha$ and $\boldsymbol{F}^h$ are infinite diagonal matrices such that for each $l \in \mathbb{Z}$,

$$(\boldsymbol{F}^p_\alpha)_{l,l} := P_\alpha\left(-f + l f_q/b\right) \sqrt{\mathcal{S}_\eta\left(-f + l f_q/b\right)}, \quad \left(\boldsymbol{F}^h\right)_{l,l} := \frac{H\left(-f + l f_q/b\right)}{\sqrt{\mathcal{S}_\eta\left(-f + l f_q/b\right)}}.$$

Theorem 2.3. *Consider the system shown in Fig. 2.3(c). Assume that $h(t)$, $p_i(t)$ and $s_i(t)$ $(1 \le i \le M)$ are all continuous, bounded and absolutely Riemann integrable, $\lambda_{\min}\left(\boldsymbol{F}^\eta \boldsymbol{F}^{\eta *}\right)$ is uniformly bounded away from 0, and $\mathcal{F}\left(q_i(t)\right) = \sum_l c_i^l \delta(f - l f_q)$. Additionally, suppose that $aMT_s = bT_q$ where a and b are coprime integers. The* *capacity $C(P)$ of the sampled channel with a power constraint P is given by*

$$C(P) = \int_0^{f_s/(aM)} \frac{1}{2} \sum_{i=1}^{aM} \log^+ \left(\nu \lambda_i \left(\left(\boldsymbol{F}^\eta \boldsymbol{F}^{\eta *} \right)^{-\frac{1}{2}} \boldsymbol{F}^\eta \boldsymbol{F}^h \boldsymbol{F}^{h*} \boldsymbol{F}^{\eta *} \left(\boldsymbol{F}^\eta \boldsymbol{F}^{\eta *} \right)^{-\frac{1}{2}} \right) \right) \mathrm{d}f,$$

where ν is chosen such that

$$P = \int_0^{f_s/(aM)} \sum_{i=1}^{aM} \left[\nu - \lambda_i^{-1} \left(\left(\boldsymbol{F}^\eta \boldsymbol{F}^{\eta *} \right)^{-\frac{1}{2}} \boldsymbol{F}^\eta \boldsymbol{F}^h \boldsymbol{F}^{h*} \boldsymbol{F}^{\eta *} \left(\boldsymbol{F}^\eta \boldsymbol{F}^{\eta *} \right)^{-\frac{1}{2}} \right) \right] \mathrm{d}f.$$

The optimal ν corresponds to a water-filling power allocation strategy based on the singular values of the equivalent channel matrix $\left(\boldsymbol{F}^\eta \boldsymbol{F}^{\eta *} \right)^{-\frac{1}{2}} \boldsymbol{F}^\eta \boldsymbol{F}^h$, where $\left(\boldsymbol{F}^\eta \boldsymbol{F}^{\eta *} \right)^{-\frac{1}{2}}$ is due to noise prewhitening and $\boldsymbol{F}^\eta \boldsymbol{F}^h$ represents the equivalent channel matrix after modulation and filtering. We note that closed-form capacity expressions may be hard to obtain for general modulating sequences $q_i(t)$. This is because the multiplication operation corresponds to convolution in the frequency domain which does not preserve Toeplitz properties of the original operator associated with the channel filter. When $q_i(t)$ is periodic, however, it can be mapped to a spike train in the frequency domain, which preserves block Toeplitz properties.

2.2.1.4 Connection to MIMO Gaussian Channels

For pedagogical purposes, we illustrate in this subsection a connection between the aliased sampled channel and MISO / MIMO channels, which allows for an intuitive communication-theoretic interpretation for Theorems 2.1-2.3.

Consider first sampling with a single filter. We suppose the Fourier transform $X(f)$ of the transmitted signal exists[1]. The Fourier transform of the sampled signal at any $f \in [0, f_s]$ is then given by

$$\frac{1}{T_s} \sum_{k \in \mathbb{Z}} H\left(f - kf_s\right) S(f - kf_s) X\left(f - kf_s\right) \tag{2.16}$$

[1]Note, however, that the Fourier transform of a stationary process typically does not exist.

due to aliasing. The summing operation allows us to treat the aliased channel at each f within the sampling bandwidth as a separate MISO channel with countably many input branches and a single output branch, as illustrated in Fig. 2.4(a). As a result, the filtered noise has power spectral density $\mathcal{S}_\eta(f)|S(f)|^2$, and hence the sampled noise has power spectrum $\sum_{l\in\mathbb{Z}} \mathcal{S}_\eta(f-lf_s)\left|S(f-lf_s)\right|^2$. Since the sampling operation combines signal components at frequencies from each aliased set $\{f-lf_s \mid l\in\mathbb{Z}\}$, it is equivalent to having a set of parallel MISO channels, each indexed by some $f\in[0,f_s]$. The water-filling strategy is optimal in allocating power among these channels, which yields the parametric equation (2.9).

The above MISO interpretation can be easily extended to channels under filter-bank sampling. The only difference is that now the equivalent channel has M receive branches, each corresponding to one branch of sampling with rate f_s/M. The noise received in the ith branch has power spectrum

$$\sum\nolimits_{l\in\mathbb{Z}} \left|S_i\left(f-lf_s/M\right)\right|^2 \mathcal{S}_\eta\left(f-lf_s/M\right), \quad f\in\left[0,f_s/M\right],$$

revealing mutual noise correlation. The received noise vector can be pre-whitened through multiplying $[\cdots,Y(f),Y(f-f_s),\cdots]^\top$ by an $M\times M$ whitening matrix $\left(\boldsymbol{F}_s(f)\boldsymbol{F}_s^*(f)\right)^{-\frac{1}{2}}$. After whitening, the channel of Fig. 2.4(b) at frequency f can be represented by a channel matrix $\left(\boldsymbol{F}_s(f)\boldsymbol{F}_s^*(f)\right)^{-\frac{1}{2}}\boldsymbol{F}_s(f)\boldsymbol{F}_h(f)$. Classical MIMO capacity results [79] immediately lead to our capacity expression.

Similarly, sampling with modulation and filter banks also enjoys a simple MIMO interpretation. Note that the signal prior to modulation in the ith branch can be expressed in the Fourier domain as $P_i(f)R(f)$, where $R(f)=H(f)X(f)+N(f)$. Multiplication of this pre-modulation signal with the modulation sequence $q_i(t)=\sum_l c_i^l\delta\left(f-lf_q\right)$ corresponds to convolution in the frequency domain. Recall that $bT_q=aMT_s$ with integers a and b. We therefore divide all samples $\{y_i[k]\mid k\in\mathbb{Z}\}$ in the ith branch into a groups, where the lth $(0\leq l<a)$ group contains $\{y_i[l+ka]\mid k\in\mathbb{Z}\}$. Hence, each group is equivalent to the samples obtained by sampling at rate $f_s/Ma=f_q/b$. The sampling system, when restricted to the output on each group of the sampling set, can be treated as LTI, thus justifying its

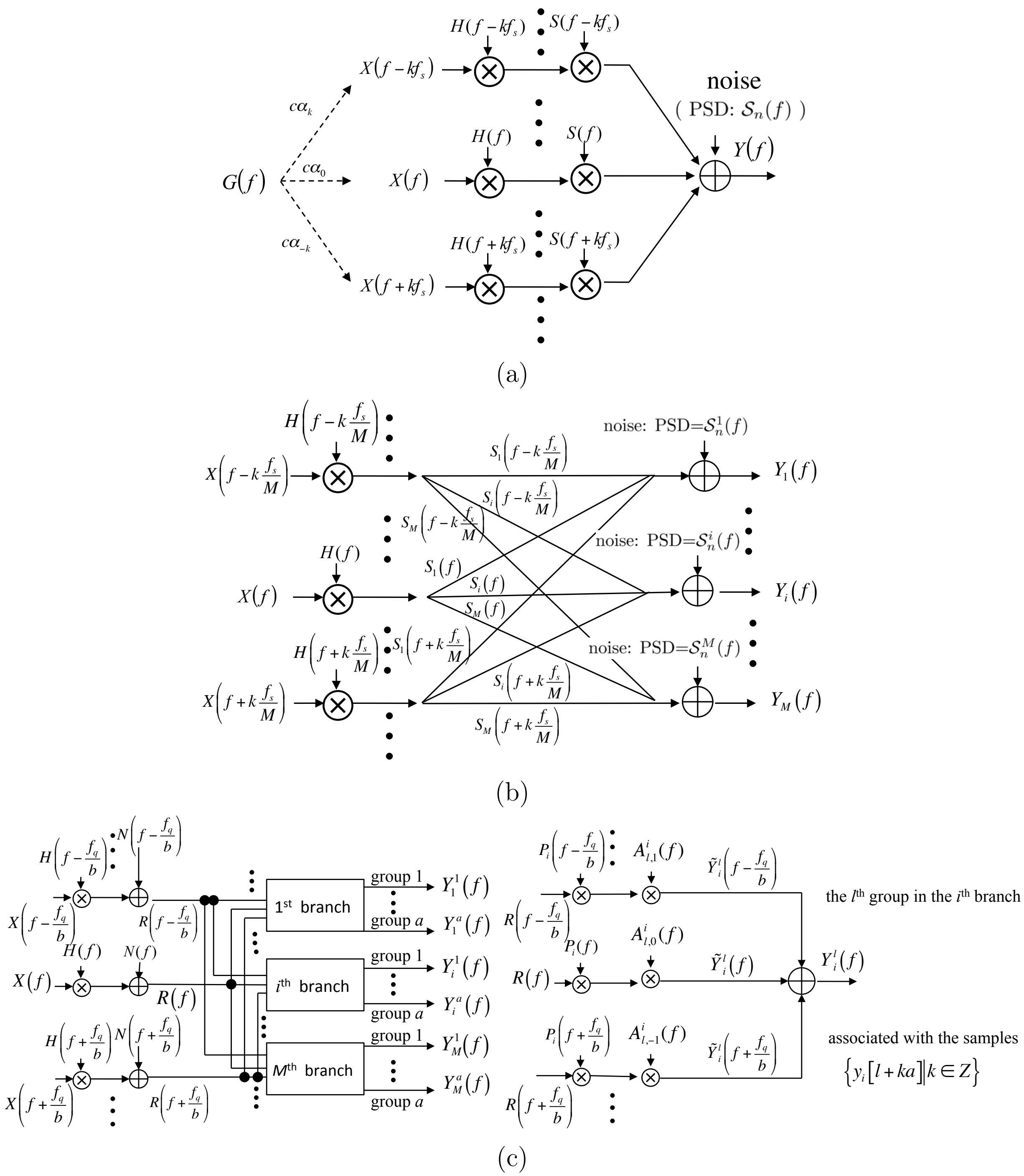


Figure 2.4: (a) Equivalent MISO Gaussian channel under filtering followed by sampling; (b) Equivalent MIMO Gaussian channel under filter-bank sampling; (c) Equivalent MIMO Gaussian channel under sampling with modulation and filter banks.

equivalent representation in the spectral domain. Since the sampled outputs of the original sampling system are equivalent to the union of samples obtained by Ma LTI systems each followed by uniform sampling at rate f_q/b, we can transform the sampled channel into a MIMO channel with infinitely many input branches and finitely many output branches, as illustrated in Fig. 2.4(c). By appropriately writing down the MIMO channel matrix, we can apply the MIMO capacity results to derive the sampled channel capacity.

2.2.1.5 Discussion and Numerical examples

Optimally Filtered Channel In general, the frequency response of the optimal prefilter is not continuous, which makes it hard to realize in practice. However, for certain classes of channel models, the prefilter enjoys a smooth frequency response. One example is a *monotone channel*, whose channel response obeys $\left|H(f_1)\right|^2/\mathcal{S}_\eta(f_1) \geq \left|H(f_2)\right|^2/\mathcal{S}_\eta(f_2)$ for any $f_1 > f_2$. Theorem 2.1 implies that the optimizing prefilter for a monotone channel reduces to a low-pass filter with cutoff frequency $f_s/2$.

For non-monotone channels, the optimal prefilter may not be a low-pass filter, as illustrated in Fig. 2.5. Fig. 2.5(b) shows the optimal filter for the channel given in Fig. 2.5(a) with $f_s = 0.4 f_{\text{NYQ}}$, which is no longer a low-pass filter.

Capacity Non-monotonicity When the channel response is not monotone, a somewhat counter-intuitive fact arises: the channel capacity $C(f_s)$ is not necessarily a non-decreasing function of the sampling rate f_s. This occurs, for example, in multi-band channels as illustrated in Fig. 2.6. Here, the Fourier transform of the channel response is concentrated in two sub-intervals within the overall channel bandwidth. Specifically, the entire channel bandwidth is contained in $[-0.5, 0.5]$ with Nyquist rate $f_{\text{NYQ}} = 1$, and that the channel response is given by

$$H(f) = \begin{cases} 1, & \text{if } |f| \in \left[\frac{1}{10}, \frac{1}{5}\right] \bigcup \left[\frac{2}{5}, \frac{1}{2}\right]; \\ 0, & \text{otherwise.} \end{cases} \tag{2.17}$$

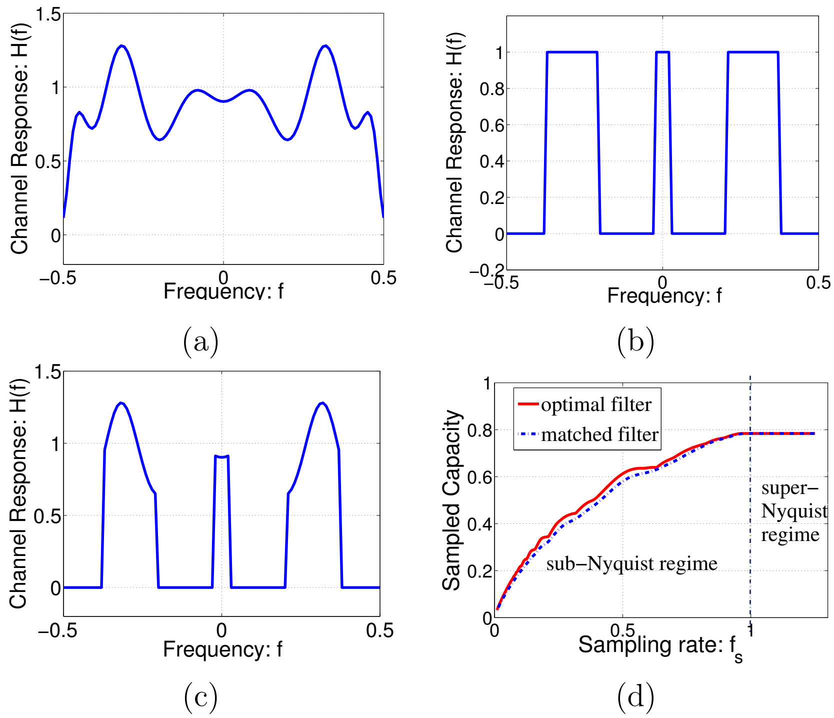


Figure 2.5: Capacity of optimally filtered channel: (a) frequency response of the original channel; (b) optimal prefilter associated with this channel for sampling rate 0.4; (c) optimally filtered channel response with sampling rate 0.4; (d) capacity vs sampling rate for the optimal prefilter and for the matched filter.

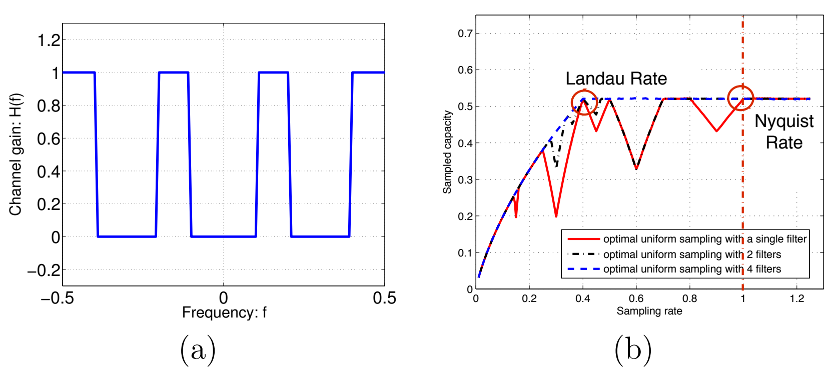


Figure 2.6: Sampled channel capacity for a multiband channel under filter-bank sampling. (a) Channel gain of the multiband channel. (b) Sampled channel capacity for a single filter followed by sampling and for a filter bank followed by sampling for a bank of two filters and of four filters.

Consider first sampling with a single filter. If this channel is sampled at a rate $f_s = \frac{3}{5} f_{\text{NYQ}}$, then aliasing occurs and leads to an aliased channel with one subband (and hence one degree of freedom). However, if sampling is performed at a rate $f_s = \frac{2}{5} f_{\text{NYQ}}$, then one can verify that the two subbands remain non-overlapping in the aliased channel, resulting in two degrees of freedom.

The tradeoff curve between capacity and sampling rate with an optimal prefilter is plotted in Fig. 2.6(b). This curve indicates that increasing the sampling rate may not necessarily increase capacity for certain channel structures. In other words, a single filter followed by sampling largely constrains our ability to exploit channel and signal structures. Similar phenomenon arises for more general filter-bank sampling. While the channel capacity with filter-bank sampling exceeds that of sampling with a single filter, the capacity is not necessarily monotone in f_s. This is shown in Fig. 2.6(b). Specifically, we see in this figure that when we apply a bank of two filters prior to sampling, the capacity curve is still non-monotonic but outperforms a single filter followed by sampling.

2.2.1.6 Single-branch Sampling with Modulation and Filtering v.s. Filter-bank Sampling

Modulation and filter bank sampling may potentially provide implementation advantages, depending on the modulation period T_q. Specifically, modulation-bank sampling may achieve a larger capacity region than that achievable by filter-bank sampling with the same number of branches.

Consider the scenario where the total channel bandwidth $W = \frac{2L}{K} f_s$ is equally divided into $2L$ subbands each of bandwidth $f_q = f_s/K$ for some integers K and L. The SNR $|H(f)|^2 / \mathcal{S}_\eta(f)$ within each subband is assumed to be flat. Algorithm 1 given below generates an alias-free sampled analog channel, which leads to maximum capacity achievable by filter-bank sampling. Specifically, for any $f \in [-f_q/2, f_q/2]$, the algorithm works as follows.

Algorithm 2.1

---

1. **Initialize.** Find the K largest elements in $\left\{ \frac{|H(f-lf_q)|^2}{\mathcal{S}_\eta(f-lf_q)} \mid -L \leq l \leq L-1 \right\}$.
Denote by $\{l_i \mid 1 \leq i \leq K\}$ the index set of these K elements such that $l_1 > l_2 > \cdots > l_K$. Set $L^* := \min \{k \mid k \geq 2L, k \bmod K = 0\}$.

2. For $i = 1 : K$

    Let $\alpha := i \cdot L^* + i - l_i$.

    Set $c^\alpha = 1$, and $S(f + \alpha f_p) = 1$.

Algorithm 1 first selects the K subbands with the highest SNR, and then moves each of the selected subbands to a new location by appropriately setting $\{c^i\}$, which guarantees that (1) the movement does not corrupt any previously designated location; (2) the contents in the newly chosen locations will be alias-free. The post-modulation filter is applied to suppress redundant spectral contents.

The performance of Algorithm 1 is equivalent to the one using an optimal filter bank followed by sampling with sampling rate f_q at each branch. Hence, single-branch sampling effectively achieves the same performance as multi-branch filter-bank sampling. This approach may be preferred since building multiple analog filters is often expensive (in terms of power consumption, size, or cost).

2.2.1.7 Connections between Capacity and MMSE

We have derived respectively the optimal prefilter / filter bank that maximize capacity. Interestingly, such choices of sampling methods coincide with the optimal prefilter / filter bank that minimize the MSE between the Gaussian channel input and the signal reconstructed from sampling the channel output.

Consider the following sampling problem. Let $x(t)$ be a zero-mean wide-sense stationary (WSS) stochastic signal whose power spectral density $\mathcal{S}_X(f)$ satisfies[2] $\int_{-\infty}^{\infty} \mathcal{S}_X(f)\mathrm{d}f = P$. This input is passed through a channel consisting of an LTI filter and contaminated by additive stationary Gaussian noise. We sample the channel output using a filter bank at a rate f_s/M in each branch, and recover a *linear*

---

[2]We restrict our attention to WSS input signals. This restriction, while falling short of generality, allows us to derive sampling results in a simple way.

MMSE estimate $\hat{x}(t)$ of $x(t)$ from its samples so as to minimize $\mathbb{E}[|x(t)-\hat{x}(t)|^2]$. We propose to jointly optimize $x(t)$ and the sampling method, i.e. which input process $x(t)$ and filter bank leads to the minimum of $\mathbb{E}[|x(t)-\hat{x}(t)|^2]$ for each sample time. The optimal input and filter bank are identified as follows.

Proposition 2.1. *Suppose the channel input* $x(t)$ *is any WSS signal satisfying* $\int_{-\infty}^{\infty}\mathcal{S}_X(f)\mathrm{d}f = P$. *For a given sampling system, let* $\hat{x}(t)$ *denote the optimal linear estimate of* $x(t)$ *from the digital sequence* $\{\boldsymbol{y}[n]\}$. *Then the capacity-optimizing filter bank given in (2.13) and its corresponding optimal input* $x(t)$ *minimize the linear MSE* $\mathbb{E}\left[|x(t)-\hat{x}(t)|^2\right]$ *over all possible LTI filter banks for any sample time.*

Proposition 2.1 implies that the input signal and the filter bank optimizing channel capacity also minimize the MSE between the original input signal and its reconstructed output. We note that if the samples $\{\boldsymbol{y}[n]\}$ and $x(t)$ are jointly Gaussian, then the MMSE estimate $\hat{x}(t)$ for a given input process $x(t)$ is linear in $\{\boldsymbol{y}[n]\}$. That said, for Gaussian inputs passed through Gaussian channels, the capacity-maximizing filter bank also minimizes the MSE even if we take into account nonlinear estimation. Thus, under sampling with filter-banks for Gaussian channels, information theory reconciles with sampling theory through the SNR metric when determining optimal systems. Intuitively, high SNR typically leads to large capacity and small MSE.

Proposition 2.1 includes the optimal prefilter under single-prefilter sampling as a special case. We note that a similar MSE minimization problem was investigated decades ago with applications in PAM [17, 18]: a given random input $x(t)$ is prefiltered, corrupted by noise, uniformly sampled, and then postfiltered to yield a linear estimate $\hat{x}(t)$. The goal in that work was to minimize the MSE between $x(t)$ and $\hat{x}(t)$ over all prefiltering (or pulse shaping) and postfiltering mechanisms. While our problem differs from this PAM design problem by optimizing directly over the random input instead of the pulse shape, the two problems are similar in spirit and result in the same alias-suppressing filter in the single-filter case.

2.2.2 Capacity under General Periodic Sampling Systems

Interestingly, all three canonical sampling structures fall under the category of periodic sampling systems, which allow us to develop a unified framework in proving Theorems 2.1-2.3.

Note that the periodicity of the sampling system renders the linear operator associated with the whole system to be block Toeplitz. The asymptotic spectral properties of block Toeplitz operators (e.g. [80]) guarantee the existence of $\lim_{T\to\infty} C_T^{\mathcal{P}}\left(f_s, P\right)$ for a given periodic sampling system $\mathcal{P}$, and allows a capacity expression to be obtained in terms of the Fourier representation. Denote by $Q_k(f)$ the Fourier transform of the impulse response $q(t_k, t_k - t)$ of the sampling system, i.e. $Q_k(f) := \int_{-\infty}^{\infty} q(t_k, t_k - t)\exp(-j2\pi f t)\mathrm{d}t$. We further introduce an $f_s T_q \times \infty$ dimensional Fourier series matrix $\boldsymbol{F}_q\left(f\right)$ and another infinite diagonal matrix $\boldsymbol{F}_h\left(f\right)$ such that for all $m, l \in \mathbb{Z}$ and $1 \leq k \leq f_s T_q$, set

$$\left(\boldsymbol{F}_q\right)_{k,l}\left(f\right) := Q_k\left(f + l f_q\right); \quad \left(\boldsymbol{F}_h\right)_{l,l}\left(f\right) := \frac{H\left(f + l f_q\right)}{\sqrt{\mathcal{S}\left(f + l f_q\right)}}.$$

We can then express the sampled capacity under $\mathcal{P}$ in closed form as follows.

Theorem 2.4. *Suppose that the sampling system $\mathcal{P}$ is periodic with period T_q and sampling rate f_s, where $f_s T_q \in \mathbb{Z}$. Let $f_q := 1/T_q$. Assume that $\left|H(f)Q_k(f)\right|^2 / \mathcal{S}_\eta(f)$ is bounded, continuous and Riemann integrable for all $1 \leq k \leq f_s T_q$. We also assume that $\lambda_{\min}\left(\boldsymbol{F}_q \boldsymbol{F}_q^*\right)$ is uniformly bounded away from 0 over all $f \in [0, f_q]$.*

(1) The sampled channel capacity under optimal power allocation is given by

$$C^{\mathcal{P}}\left(P\right) = \frac{1}{2}\int_0^{f_q} \sum_{i=1}^{f_s T_q} \log\left[\nu \cdot \lambda_i\right]^+ \mathrm{d}f, \tag{2.18}$$

where ν satisfies

$$\int_0^{f_q} \sum_{i=1}^{f_s T_q} \left[\nu - \lambda_i^{-1}\right]^+ \mathrm{d}f = P.$$

Here, λ_i denotes the ith largest eigenvalue of $\left(\boldsymbol{F}_q \boldsymbol{F}_q^*\right)^{-\frac{1}{2}} \boldsymbol{F}_q \boldsymbol{F}_h \boldsymbol{F}_h^* \boldsymbol{F}_q^* \left(\boldsymbol{F}_q \boldsymbol{F}_q^*\right)^{-\frac{1}{2}}$.

(2) Suppose further that $H(f)=0$ for any $f\notin[0,W]$, and that the transmitter employs equal power allocation over $[0,W]$. Then the sampled channel capacity is

$$C_{\mathrm{eq}}^{\mathcal{P}}(P)=\int_0^{f_q}\frac{1}{2}\log\det\left(\boldsymbol{I}+\frac{P}{W}\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\right)\mathrm{d}f. \tag{2.19}$$

In Theorem 2.4, ν is the water-level with respect to the optimal water-filling power allocation strategy over the eigenvalues of the matrix $\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}$. The capacity expression (2.18) admits a simple upper bound, as stated below.

Corollary 2.3. *(a) Consider the setup and assumptions in Lemma 2.4. Under all periodic sampling systems with period T_q and sampling rate f_s, the sampled channel capacity can be bounded above by*

$$C_{f_q}(f_s,P)=\frac{1}{2}\int_0^{f_q}\sum_{i=1}^{f_sT_q}\left[\log\left(\nu_{\mathrm{p}}\lambda_i\left(\boldsymbol{F}_h\boldsymbol{F}_h^*\right)\right)\right]^+\mathrm{d}f, \tag{2.20}$$

where ν_{p} satisfies $\int_{-f_q/2}^{f_q/2}\sum_{i=1}^{f_sT_q}\left[\nu_{\mathrm{p}}-\lambda_i^{-1}\left(\boldsymbol{F}_h\boldsymbol{F}_h^*\right)\right]^+\mathrm{d}f=P$.

(b) Suppose that there exists a frequency set B_{m} that satisfies $\mu\left(B_{\mathrm{m}}\right)=f_s$ and

$$\int_{f\in B_{\mathrm{m}}}\frac{|H(f)|^2}{\mathcal{S}_\eta(f)}\mathrm{d}f=\sup_{B:\mu(B)=f_s}\int_{f\in B}\frac{|H(f)|^2}{\mathcal{S}_\eta(f)}\mathrm{d}f.$$

Then

$$C_{f_q}(f_s,P)\leq C_{\mathrm{u}}(f_s,P), \tag{2.21}$$

where $C_{\mathrm{u}}(f_s,P)$ is given by (2.23).

Proof. (a) Following the same steps as in Proposition A.1, we can see that the ith largest eigenvalue satisfies

$$\lambda_i\left\{\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\right\}\leq\lambda_i\left(\boldsymbol{F}_h\boldsymbol{F}_h^*\right),$$

which immediately leads to (2.20).

(b) For any given f_q, the upper bound (2.20) is obtained by extracting out a certain frequency set B that has measure $\mu(B) = f_s$ and suppressing all spectral components outside B. By our definition of B_m, any choice of B with spectral size f_s will not outperform B_m. Hence, choosing $B = B_m$ leads to a universal upper bound. □

Corollary 2.3 reveals that the capacity under any periodic sampling system, regardless of its period, cannot exceed the upper bound $C_{\mathrm{u}}(f_s, P)$ in Theorem 2.5. As we will show, this is the key observation in establishing the fundamental upper bound for more general sampling systems.

2.3 Fundamental Upper Bound on Sampled Channel Capacity

We are now in a position to characterize the fundamental upper bound on sampled channel capacity for a general class of right-invertible time-preserving nonuniform sampling methods, under a sub-Nyquist sampling rate constraint. Specifically, we are concerned with the sampled channel capacity $C_{\mathcal{A}}(f_s, P)$, where

$$\mathcal{A} := \{\text{all right-invertible time-preserving sampling systems}\}.$$

Here, the right-invertibility represents some mild regularity constraints that ensure each sample contains innovation information, as will be formally defined later. Unless otherwise specified, all sampling systems mentioned in this section are assumed to be right-invertible time-preserving linear systems.

Before proceeding, we shall assume that for any frequency f,

$$\int_f \frac{H^2(f)}{\mathcal{S}_\eta(f)} \mathrm{d}f < \infty, \quad \text{and} \quad \int_f \mathcal{S}_\eta(f) \mathrm{d}f < \infty \text{ or } \mathcal{S}_\eta(f) \equiv 1. \tag{2.22}$$

2.3.1 An Upper Bound on Sampled Channel Capacity

A time-preserving sampling system preserves the time scale of the signal, and hence does not compress or expand the frequency scale. We now determine an upper bound on the sampled channel capacity for this class of general nonuniform sampling systems. Since any multibranch sampling system can be converted into a single branch sampling system without loss of information, we restrict our analysis to the class of single branch samplers, which provides exactly the same upper bound as the one accounting for multibranch systems. In addition, we constrain our attention to right-invertible sampling methods, as defined below[3].

Definition 2.8 (**Right-Invertible Sampling System**). *A sampling system with sampling set* Λ *and impulse response* $q(t_i, \tau)$ $(t_i \in \Lambda)$ *is said to be right-invertible with respect to* $\mathcal{S}_\eta(f)$ *if*

1. *for any* $k \in \mathbb{Z}$*, the frequency response* $\mathcal{F}\left(q_k\left(\tau\right)\right)$ *is bounded;*

2. *for any frequency* f *and any* T *with its associated sampling subset* $\Lambda_{[-T,T]} = [-T,T] \cap \Lambda := \{t_1, \cdots, t_{N_T}\}$*,* $\lambda_{\min}\left(\boldsymbol{F}_T \boldsymbol{F}_T^*\right)$ *is uniformly bounded away from 0 over all* f*.*

Here, $q_k\left(\tau\right) := q\left(t_k, t_k - \tau\right)$*, and* $\boldsymbol{F}_T$ *represents an* $N_T \times \infty$ *matrix obeying*

$$\left[\boldsymbol{F}_T\left(f\right)\right]_{k,l} = \mathcal{F}\left(s_\eta \cdot q_k\right)\left(f + l/T\right), \quad 1 \leq k \leq N_T,\ l \in \mathbb{Z}.$$

The right invertibility of the sampling system essentially implies that for each subset of the impulse response $\{q(t_i, \tau) \mid i \in \mathcal{I}\}$, the Fourier matrix associated with any sampled response is right-invertible. This typically implies that the set of samples is a linearly independent family – each sample provides a sufficient amount of innovative information. Our main theorem is now stated as follows.

Theorem 2.5 (**Converse**). *Assume that there exists a small constant* $\epsilon > 0$ *such that* $\mathcal{F}^{-1}\left(\frac{H(f)}{\sqrt{\mathcal{S}_\eta(f)}}\right)(t) = O\left(\frac{1}{t^{1.5+\epsilon}}\right)$*. Suppose that there exists a frequency set* B_{m} *with* $\mu\left(B_{\mathrm{m}}\right) = f_s$ *satisfying*

$$\int_{f \in B_{\mathrm{m}}} \frac{\left|H(f)\right|^2}{\mathcal{S}_\eta(f)} \mathrm{d}f = \sup_{B:\mu(B)=f_s} \int_{f \in B} \frac{\left|H(f)\right|^2}{\mathcal{S}_\eta(f)} \mathrm{d}f.$$

[3]We impose the right-invertibility constraint primarily out of mathematical convenience. We conjecture, however, that removing this constraint does not change our main result (Theorem 2.5).

Then under any time-preserving right-invertible sampling system $\mathcal{P}$ w.r.t. $\mathcal{S}_\eta(f)$ with sampling rate f_s, the sampled channel capacity is upper bounded by

$$C^{\mathcal{P}}(P) \leq C_{\mathrm{u}}(f_s, P) := \int_{f\in B_{\mathrm{m}}} \frac{1}{2}\left[\log\left(\nu \frac{|H(f)|^2}{\mathcal{S}_\eta(f)}\right)\right]^+ \mathrm{d}f, \tag{2.23}$$

where ν is given parametrically by

$$\int_{f\in B_{\mathrm{m}}} \left[\nu - \frac{\mathcal{S}_\eta(f)}{|H(f)|^2}\right]^+ \mathrm{d}f = P. \tag{2.24}$$

Remark 2.2. Note that $C_{\mathrm{u}}(f_s, P)$ is monotonically nondecreasing in f_s and P. In fact, when the sampling rate is increased from f_s to $f_s+\delta$, $C_{\mathrm{u}}(f_s, P)$ corresponds to the optimal value when considering all spectral sets of support size $f_s+\delta$, which is obtained by optimizing over a larger set of transmission / power allocation strategies than the situation with sampling rate f_s. Therefore, $C_{\mathrm{u}}(f_s, P)$ is monotone in f_s.

It can be easily shown that the upper bound $C_{\mathrm{u}}(f_s, P)$ is equivalent to the maximum capacity of a channel whose spectral occupancy is no larger than f_s. The above result basically implies that even if we employ more complex irregular sampling sets, the sampled capacity cannot exceed the one commensurate with the analog capacity when constraining all transmit signals to the interval of bandwidth f_s that experience the highest SNR. Accordingly, the optimal inputs will lie in this maximizing frequency set. This theorem also demonstrates that the capacity is attained when aliasing is suppressed by the sampling structure, as will be seen in our capacity-achieving scheme. When the optimal frequency set B_{m} is selected, a water filling power allocation strategy is performed over B_{m} with some water level ν.

2.3.1.1 Proof Sketch of Theorem 2.5

To help the reader navigate through the proof of Theorem 2.5, we now outline the key steps under white noise.

i) We start by analyzing the class of periodic sampling systems: a special type of sampling methods that allow closed-form capacity expressions. We then demonstrate that the capacity under any periodic sampling system with sampling rate f_s and transmit power P is bounded above by $C_{\rm u}(f_s, P)$.

ii) The upper bound is then derived by relating general (possibly aperiodic) sampled channels with periodic sampled channels through a *finite-duration* approximation. In fact, instead of studying the true sampled channel response directly, we truncate the channel response so that its impulse response is supported in a finite duration. The capacity for the resulting truncated channel is then bounded by the capacity of a new periodized channel we construct. As we show, the capacity of the truncated channel can be made arbitrarily close to the capacity of the true sampled channel.

The most technically involved step is Step ii), which proceeds as follows.

1. **Finite-duration** $h(t)$. Consider first channels for which $h(t)$ is of finite duration, $h(t) = 0$ for any $t \notin [-L_0, L_0]$ for some $L_0 > 0$.

   (a) Consider any given right-invertible time-preserving sampling system $\mathcal{P}$ with impulse response $q(t,\tau)$, and suppose that the input $x(t)$ is time constrained to the interval $[-T, T]$. Construct a periodic channel with period $2(T + L_0)$ based on $q(t,\tau)$. Let $C_{\rm p}^{\mathcal{P}}(P)$ denote the capacity of the periodized channel, whose sampling rate is bounded above by $f_s + \epsilon$ for any small $\epsilon > 0$.

   (b) Show that $C_T^{\mathcal{P}}(P) \leq \frac{T+L_0}{T} C_{\rm p}^{\mathcal{P}}\left(\frac{T}{T+L_0}P\right)$ holds uniformly for all $\mathcal{P}$. Since we know that $C_{\rm p}^{\mathcal{P}}(P) \leq C_{\rm u}(f_s + \epsilon, P)$ for any periodized channel (or, equivalently, any channel followed by a periodic sampler), this establishes the capacity upper bound for finite-duration channels, provided that T is sufficiently large.

2. **Infinite-duration** $h(t)$. We next extend the results to channels for which $h(t)$ is non-zero over infinite duration.

   (a) Construct a truncated channel such that

$$\tilde{h}(t) = \begin{cases} h(t), & \text{if } |t| \leq L_1, \\ 0, & \text{else,} \end{cases}$$

for some sufficiently large L_1. The capacity upper bound holds for the truncated channel, as shown in Step 1).

(b) For any given sampling system $\mathcal{P}$ and any time interval $[-T, T]$, compare the capacity of the original channel (denoted by $C_T^{\mathcal{P}}(P)$) with the capacity of the truncated channel (denoted by $\tilde{C}_T^{\mathcal{P}}(P)$), which can be completed by investigating the spectrum of the operators associated with both sampled channels. It can be shown that $C_T^{\mathcal{P}}(P)$ can be upper bounded by $\tilde{C}_T^{\mathcal{P}}(P+\xi)+\xi$ for some arbitrarily small constant $\xi > 0$, which holds uniformly over all sampling systems $\mathcal{P}$. Combining this with results shown in Step 1), we demonstrate that $\tilde{C}_T^{\mathcal{P}}(P)$ (and hence $\tilde{C}_T^{\mathcal{P}}(P)$) is arbitrarily close to $C_{\mathrm{u}}(f_s, P)$, which establishes the claim for the whole class of infinite-duration channels.

2.3.2 Achievability

It turns out that for most scenarios of interest, the capacity upper bound derived in Theorem 2.5 can be attained by filterbank sampling, as stated below.

Theorem 2.6 (**Achievability – Sampling with a Filter Bank).** *Suppose that the maximizing frequency set B_{m} introduced in Theorem 2.5 exists and is piecewise continuous or, more precisely, $B_{\mathrm{m}} = \cup_i B_{i\in\mathcal{X}}$, where $\mathcal{X}$ is an index set, and B_i's are non-overlapping continuous intervals. Consider the following filterbank sampling system $\mathcal{P}_{\mathrm{FB}}$: in the kth branch, the frequency response of the filter is given by*

$$S_k(f) = \begin{cases} 1, & \text{if } f \in B_k, \\ 0, & \text{otherwise}, \end{cases} \tag{2.25}$$

and each filter is followed by a uniform sampler with sampling rate $\mu(B_k)$. Then

$$C^{\mathcal{P}_{\mathrm{FB}}}(P) = C_{\mathrm{u}}(f_s, P),$$

where $C_{\mathrm{u}}(f_s, P)$ is the upper bound given by (2.23).

Proof. The spectral components in B_i can be perfectly reconstructed from the sequence that is obtained by first extracting out a subinterval B_i and then uniformly sampling the filtered output with sampling rate $f_{s,i}$. The capacity under $\mathcal{P}_{\mathrm{FB}}$ is commensurate to the analog capacity when constraining the transmit signal to $\cup_i B_i$, which is equivalent to $C_{\mathrm{u}}(f_s, P)$. $\square$

Note that the bandwidth of B_i may be irrational and the system may require an infinite number of filters. Theorem 2.6 indicates that filterbank sampling with *varied sampling rates* in different branches maximizes capacity.

The optimality of filterbank sampling immediately leads to another optimal sampling structure under mild conditions. As discussed above, filterbank sampling with equal rates on different branches is equivalent to a single branch of modulation. This approach attains the sampled capacity achievable by filterbank sampling if the SNRs of the analog channel are piecewise constant in frequency. Although the filterbank sampling we derive in (2.25) does not employ equal rates on different branches, for most channels of physical interest we can simply divide each branch further into a number of sub-branches to allow the rates at different branches to be reasonably close to each other. Therefore, for most channels of physical interest (say, the channels whose SNRs in frequency are Riemann-integrable), the capacity achievable through filterbank sampling can be approached arbitrarily closely by a single branch of sampling with modulation. This achievability result is formally stated as follows.

Theorem 2.7 (**Achievability – A Single Branch of Sampling with Modulation and Filtering).** *Under the assumptions of Theorem 2.6, suppose further that* $|H(f)|^2 / \mathcal{S}_\eta(f)$ *is constant within each set* B_i*. Then for any* $\epsilon > 0$*, there exists a time-preserving sampling system* $\mathcal{P}_{MF}$ *with sampling rate* f_s *using a single branch of sampling with modulation and filtering such that* $C^{\mathcal{P}_{MF}}(P) \geq C_{\mathrm{u}}(f_s, P) - \epsilon$*, where* $C_{\mathrm{u}}(f_s, P)$ *is defined in (2.23).*

Proof. It is straightforward to see that there exists a set of non-overlapping intervals $\left\{\tilde{B}_k\right\}$ each with *equal measure* that can approximate the original sets $\{B_l \mid 1 \leq l \leq n\}$ arbitrarily well. Employing the sampling method described in Algorithm 2.1 achieves a sampled capacity arbitrarily close to $C_{\mathrm{u}}(f_s, P)$. $\square$

A channel of physical interest can often be approximated as piecewise constant over frequency in this way. Given the maximizing frequency set B_{m}, the sampling structure first suppresses the frequency components outside B_{m} using an optimal LTI prefilter. A modulation module is then applied to scramble all frequency components within B_{m}. The aliasing effect can be significantly mitigated by appropriate choices of modulation weights for different spectral subbands. We then employ another band-pass filter to suppress out-of-band signals, and sample the output using a pointwise uniform sampler. Compared with filterbank sampling, a single branch of modulation and filtering might be of lower complexity to implement.

2.4 Discussion

The fundamental capacity limits we derive as well as the capacity-maximizing sampling structures provide insights into practical system design. Before concluding this chapter, we discuss some of the important implications as follows.

Monotonicity. It can be seen from Theorem 2.5 that increasing the sampling rate from f_s to $\tilde{f}_s$ results in another frequency set $\tilde{B}_{\mathrm{m}}$ of support size $\tilde{f}_s$ that has the highest SNRs. By definition, the original frequency set B_{m} must be a subset of $\tilde{B}_{\mathrm{m}}$. Therefore, the sampled capacity with rate $\tilde{f}_s$ is no lower than the sampled capacity with rate f_s.

Irregular sampling set. Sampling with irregular nonuniform sampling sets, while requiring complicated reconstruction and interpolation techniques [7], does not outperform filterbank or modulation bank sampling with regular uniform sampling sets in maximizing capacity for the channels considered herein.

Alias suppression. We have seen that aliasing does not allow a higher capacity to be achieved when perfect channel state information is known at both the transmitter and the receiver. The optimal sampling method corresponds to the optimal alias-suppression strategy. This is in contrast to the benefits obtained through random mixing of spectral components in many sub-Nyquist sampling schemes with unknown signal supports. When we are allowed to jointly optimize over both input and sampling schemes with perfect channel state information, scrambling of spectral contents does not in general maximize capacity.

Perturbation of the sampling set. If optimal filterbank or modulation sampling is employed, then mild perturbation of post-filtering uniform sampling sets does not degrade the sampled capacity. One surprisingly general example was proved by Kadec [81]. Suppose that a sampling rate $\hat{f}_s$ is used in any branch and the sampling set satisfies $\left|\hat{t}_n - n/\hat{f}_s\right| \leq \hat{f}_s/4$. Then $\left\{\exp\left(j2\pi\hat{t}_n f\right) \mid n \in \mathbb{Z}\right\}$ also forms a Riesz basis of $\mathcal{L}_2(-\hat{f}_s/2, \hat{f}_s/2)$, thereby preserving information integrity. These sampling and reconstruction schemes, while generally complicated to implement in practice, significantly broaden the class of sampling mechanisms that allow perfect reconstruction of bandlimited signals, and indicate stability of the sampling sets. Kadec's result immediately implies that the sampled capacity is invariant under mild perturbation of the sampling sets.

Hardware implementation. When the sampling rate is increased from f_{s1} to f_{s2}, we need only to insert an additional filter bank of overall sampling rate $f_{s2} - f_{s1}$ to extract out another set of spectral components with bandwidth $f_{s2} - f_{s1}$. Thus, the adjustment of the sampling hardware system for filterbank sampling is incremental with no need to rebuild the whole system from scratch.

Spectrum Blind Sampling. In this chapter we have focused on the scenario with perfect channel state information known at the transmitter, the receiver, and the sampler. This is different from the setting of compressed sensing, where the signal spectrum is unknown to the sampler and the decoder. In fact, the alias-suppressing sampler requires knowledge of the channel. If this knowledge is not available, then alias-suppressing samplers might result in low capacity. When the sampler is spectrum blind and the channel realization is uncertain, random sampling that scrambles the spectral contents [78, 37] outperforms alias-suppressing sampling in minimizing the rate loss due to channel-independent sampling design.

Chapter 3

Minimax Capacity Loss under Universal Sub-Nyquist Sampling

In this chapter, our goal is to explore universal design of a sub-Nyquist sampling system that is robust against the uncertainty and variation of instantaneous channel realizations, based on sampled capacity loss as a metric. In particular, we investigate the fundamental lower limit of sampled capacity loss in some overall sense, and design a sub-Nyquist sampling system for which the capacity loss can be uniformly optimized over all possible channel realizations.

3.1 Problem Formulation and Preliminaries

3.1.1 Compound Multiband Channel

We consider a compound multiband Gaussian channel. The channel has a total bandwidth W, and is divided into n continuous subbands each of bandwidth W/n. A state $\boldsymbol{s} \in \binom{[n]}{k}$ is generated, which dictates the channel support and realization[1]. Specifically, given a state $\boldsymbol{s}$, the channel is an LTI filter with impulse response $h_{\boldsymbol{s}}(t)$ and frequency response $H_{\boldsymbol{s}}(f) = \int_{-\infty}^{\infty} h_{\boldsymbol{s}}(t) \exp(-j2\pi f t)\mathrm{d}t$. It is assumed throughout

[1]Note that in practice, n is typically a large number. For instance, the number of subcarriers ranges from 128 to 2048 in LTE [82].

that there exists a general function $H(f, \boldsymbol{s})$ such that for every f and $\boldsymbol{s}$, $H_{\boldsymbol{s}}(f)$ can be expressed as

$$H_{\boldsymbol{s}}(f) = H(f, \boldsymbol{s})\mathbf{1}_{\boldsymbol{s}}(f), \text{ where } \mathbf{1}_{\boldsymbol{s}}(f) = \begin{cases} 1, & \text{if } f \text{ lies within subbands at indices from } \boldsymbol{s}, \\ 0, & \text{else.} \end{cases}$$

A transmit signal $x(t)$ with a power constraint P is passed through this multiband channel, which yields a channel output $r_{\boldsymbol{s}}(t) = h_{\boldsymbol{s}}(t) * x(t) + \eta(t)$, where $\eta(t)$ is stationary zero-mean Gaussian noise with power spectral density $\mathcal{S}_\eta(f)$. We assume that perfect CSI is available to both the transmitter and the receiver.

The above model subsumes as special cases the following communication scenarios.

- **Time Division Multiple Access Model.** In this setting the channel is shared by a set of different users. At each timeframe, one of the users is selected for transmission. The receiver (e.g. the base station) allocates a subset of subbands to the designated sender over that timeframe.

- **White-Space Cognitive Radio Network.** In a white-space cognitive radio network, cognitive users exploit spectrum holes unoccupied by primary users and utilize them for transmission. Since the locations of spectrum holes change over time, the spectral subbands available to cognitive users is time-varying.

3.1.2 Sampled Channel Capacity

We aim to design a sampler that works at rates below the Nyquist rate (i.e. the channel bandwidth W). In particular, we focus on periodic sampling systems defined in (2.2), which subsume the most widely used sampling mechanisms in practice.

Consider a periodic sampling system $\mathcal{P}$ with period $T_q = n/W$ and sampling rate $f_{\mathrm{s}} := mW/n$ for some integer m. A special case consists of sampling with a combination of filter banks and periodic modulation with period n/W, as illustrated in Fig. 3.1(a). Specifically, the sampling system comprises m branches, where at each branch, the channel output is passed through a pre-modulation filter, modulated by a periodic waveform of period T_q, and then filtered with a post-modulation filter followed by uniform sampling at rate f_{s}/m. Denote by $\boldsymbol{s}_i$ and f_i the ith smallest element in $\boldsymbol{s}$ and the lowest frequency of the ith subband, respectively, and define $\boldsymbol{H}_{\boldsymbol{s}}(f)$ as a $k \times k$ diagonal matrix obeying

$$(\boldsymbol{H}_{\boldsymbol{s}}(f))_{ii} = \frac{|H_{\boldsymbol{s}}(f_{\boldsymbol{s}_i} + f)|}{\sqrt{\mathcal{S}_\eta(f_{\boldsymbol{s}_i} + f)}}.$$

As will be shown later, in the high SNR regime, employing water-filling power allocation harvests only marginal capacity gain relative to equal power allocation. For this reason as well as mathematical convenience, we will start by studying the sampled channel capacity under uniform power allocation, which suffices to approach the fundamental minimax limit. Recall that if equal-power allocation is employed and perfect CSI is known at both the transmitter and receiver, then the sampled capacity given a state $\boldsymbol{s}$, denoted by $C_{\boldsymbol{s}}^{\boldsymbol{Q}}$ for brevity, obeys (Theorem 2.4)

$$C_{\boldsymbol{s}}^{\boldsymbol{Q}} = \frac{1}{2}\int_0^{W/n} \log\det\left(\boldsymbol{I}_m + \frac{P}{\beta W}\boldsymbol{Q}^{\mathrm{w}}(f)\boldsymbol{H}_{\boldsymbol{s}}^2(f)\boldsymbol{Q}^{\mathrm{w}*}(f)\right)\mathrm{d}f, \tag{3.1}$$

where $\beta := k/n$ denotes the instantaneous band sparsity ratio. Here, $\boldsymbol{Q}^{\mathrm{w}}(f) := (\boldsymbol{Q}(f)\boldsymbol{Q}^*(f))^{-1/2}\boldsymbol{Q}(f)$ for some $m \times n$ matrix $\boldsymbol{Q}(f)$ that only depends on $\mathcal{P}$. In general, $\boldsymbol{Q}(f)$ is a function that varies with f. Unless otherwise specified, we call $\boldsymbol{Q}(\cdot)$ the *sampling coefficient function* and $\boldsymbol{Q}^{\mathrm{w}}(\cdot)$ the *whitened sampling coefficient function* with respect to the sampling system $\mathcal{P}$. Note that $\boldsymbol{Q}^{\mathrm{w}}(f)\boldsymbol{Q}^{\mathrm{w}*}(f) = \boldsymbol{I}$.

3.1.2.1 Flat Sampling Coefficient Function

A special class of periodic sampling concerns the ones whose $\boldsymbol{Q}(\cdot)$ are flat over $[0, f_{\mathrm{s}}/m]$, in which case we can use an $m \times n$ matrix $\boldsymbol{Q}$ to represent the sampling coefficient function, termed a *sampling coefficient matrix*. This class of sampling systems can be realized through the m-branch sampling system illustrated in Fig. 3.1(b). In the ith branch, the channel output is modulated by a periodic waveform $q_i(t)$ of period n/W, passed through a low-pass filter with pass band $[0, f_{\mathrm{s}}/m]$, and then uniformly sampled at rate f_{s}/m, where the Fourier transform of $q_i(t)$ obeys $\mathcal{F}(q_i(t)) = \sum_{l=1}^{n} \boldsymbol{Q}_{i,l}\delta(f - lW/n)$. In this chapter, a sampling system within this

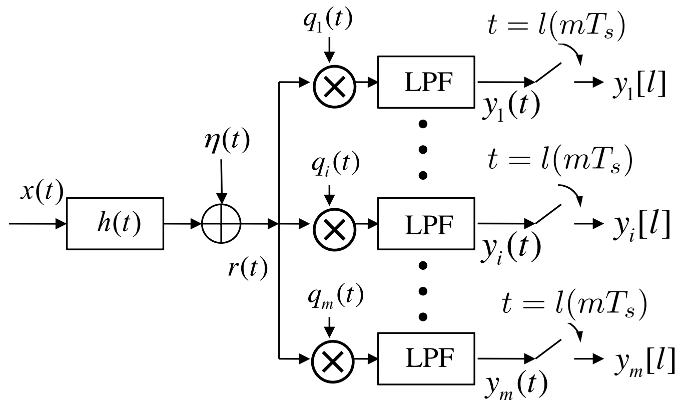


Figure 3.1: Sampling with a modulation bank and low-pass filters.

class is said to be *(independent) random sampling* if the entries of $\boldsymbol{Q}$ are randomly generated (and are independent). In addition, a sampling system is termed *Gaussian sampling* if the entries of $\boldsymbol{Q}$ are i.i.d. Gaussian.

It turns out that this simple class of sampling structures is sufficient to achieve overall robustness in terms of sampled capacity loss, provided that the entries of $\boldsymbol{Q}$ are sub-Gaussian with zero mean and unit variance, as will be detailed in Section 3.2.

3.1.3 Universal Sampling

As was shown in Section 2.3, the optimal sampling mechanism for a given LTI channel with perfect CSI extracts out the frequency set with the highest SNR. Such an alias-suppressing sampler may achieve a very low capacity for some channel realizations. In this chapter, we desire a sampler that operates independent of the instantaneous CSI, and our objective is to design a single sampler that achieves to within a minimal gap of the Nyquist-rate capacity across all possible channel realizations. Our assumptions regarding the availability of CSI to the transmitter, receiver and sampler under universal sampling is illustrated in Fig. 3.2.

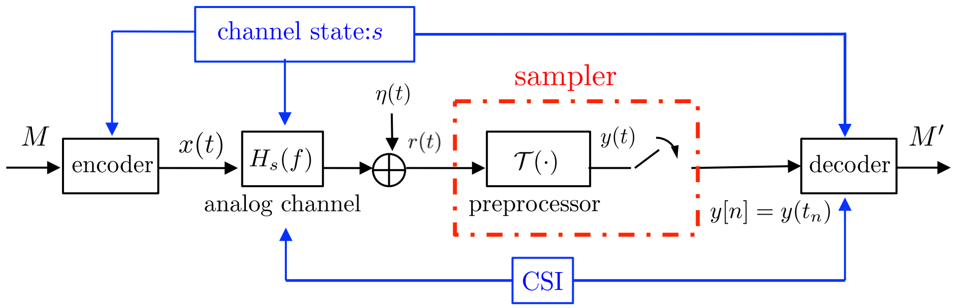


Figure 3.2: At each timeframe, a state is generated from a finite set $\mathcal{S}$, which dictates the channel realization $H_{\boldsymbol{s}}(f)$. Both the transmitter and the receiver have perfect CSI, while the sampler operates independently of $\boldsymbol{s}$.

3.1.3.1 Sampled Capacity Loss

Universal sub-Nyquist samplers suffer from information rate loss relative to Nyquist-rate capacity. In this subsection, we make formal definitions of this metric.

For any state $\boldsymbol{s}$, when equal power allocation is performed over active subbands, the Nyquist-rate capacity can be written as

$$C_{\boldsymbol{s}}^{P_{\mathrm{eq}}} = \int_0^{W/n} \frac{1}{2} \log\det\left(\boldsymbol{I}_k + \frac{P}{\beta W} \boldsymbol{H}_{\boldsymbol{s}}^2(f) \right) \mathrm{d}f, \tag{3.2}$$

which is a special case of (3.1). In contrast, if power control at the transmitter side is allowed, then the Nyquist-rate capacity is given by

$$C_{\boldsymbol{s}}^{\mathrm{opt}} = \int_0^{W/n} \frac{1}{2} \sum_{i=1}^{k} \log^+ \left(\nu \left(\boldsymbol{H}_{\boldsymbol{s}}(f) \right)_{ii}^2 \right) \mathrm{d}f, \tag{3.3}$$

where ν is determined by the equation

$$P = \int_0^{W/n} \sum_{i=1}^{k} \left(\nu - \left(\boldsymbol{H}_{\boldsymbol{s}}(f) \right)_{ii}^{-2} \right)^+ \mathrm{d}f. \tag{3.4}$$

We can then make formal definitions of sampled capacity loss as follows.

Definition 3.1 (**Sampled Capacity Loss**). *For any sampling coefficient function $\boldsymbol{Q}(\cdot)$ and any given state $\boldsymbol{s}$, we define the capacity loss without power control as*

$$L_{\boldsymbol{s}}^{\boldsymbol{Q}} := C_{\boldsymbol{s}}^{P_{\mathrm{eq}}} - C_{\boldsymbol{s}}^{\boldsymbol{Q}},$$

and define the sampled capacity loss with optimal power control as

$$L_{\boldsymbol{s}}^{\boldsymbol{Q},\mathrm{opt}} := C_{\boldsymbol{s}}^{\mathrm{opt}} - C_{\boldsymbol{s}}^{\boldsymbol{Q}}.$$

These metrics quantify the capacity gaps relative to Nyquist-rate capacity due to *universal* (channel-independent) sub-Nyquist sampling design. When sampling is performed at or above the Landau rate (which is equal to kW/n in our case) but below the Nyquist rate, these gaps capture the rate loss due to channel-independent sampling relative to channel-optimized design, either with or without power control.

For notational convenience, for an $m \times n$ matrix $\boldsymbol{M}$, we denote by $L_{\boldsymbol{s}}^{\boldsymbol{M}}$ and $L_{\boldsymbol{s}}^{\boldsymbol{M},\mathrm{opt}}$ the capacity loss with respect to a sampling coefficient function $\boldsymbol{Q}(f) \equiv \boldsymbol{M}$, which is flat across $[0, W/n]$.

3.1.3.2 Minimax sampler

Frequently used in the theory of statistics (e.g. [83]), minimaxity is a metric that seeks to minimize the loss function in some overall sense, defined as follows.

Definition 3.2 (**Minimax Sampler**). *A sampling system associated with a sampling coefficient function $\boldsymbol{Q}^m$, which minimizes the worst-case capacity loss, that is, which satisfies*

$$\max_{\boldsymbol{s} \in \binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{Q}^m} = \inf_{\boldsymbol{Q}(\cdot)} \max_{\boldsymbol{s} \in \binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{Q}},$$

is called a minimax sampler with respect to the state alphabet $\binom{[n]}{k}$.

The minimax criteria is of interest for designing a sampler robust to all possible channel states, that is, achieving to within a minimal gap relative to maximum capacity for all channel realizations. It aims to control the rate loss across all states in a uniform manner, as illustrated in Fig. 3.3. Note that the minimax sampler is in

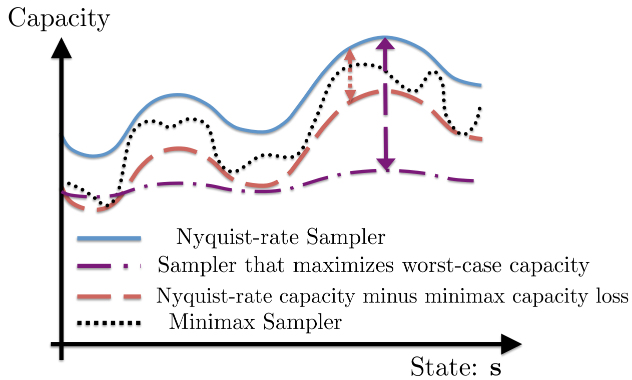


Figure 3.3: Minimax sampler v.s. the sampler that maximizes worst-case capacity, when sampling is channel-independent and performed below the Nyquist rate. The blue solid line represents the Nyquist-rate (analog) capacity, the black dotted line represents the capacity achieved by minimax sampler, the orange dashed illustrates the Nyquist-rate capacity minus the minimax capacity loss, while the purple dashed line corresponds to maximum worst-case capacity.

general different from the one that maximizes the lowest capacity among all states (worst-case capacity). While the latter guarantees an optimal worst-case capacity that can be achieved regardless of which channel is realized, it may result in significant capacity loss in many states with large Nyquist-rate capacity, as illustrated in Fig. 3.3. In contrast, a desired minimax sampler controls the capacity loss for every single state $\boldsymbol{s}$, and allows for robustness over all channel states with universal channel-independent sampling. It turns out that in the compound multiband channels,

$$\forall \boldsymbol{s}, \quad L_{\boldsymbol{s}}^{\boldsymbol{Q}^{\mathrm{m}}} = \max_{\tilde{\boldsymbol{s}} \in \binom{[n]}{k}} L_{\tilde{\boldsymbol{s}}}^{\boldsymbol{Q}^{\mathrm{m}}}$$

except for some vanishingly small residual terms, as will be shown in the next section.

3.1.4 Notation

Before proceeding, we introduce several notations used throughout this chapter. Denote by $\mathcal{H}(\beta) := -\beta \log \beta - (1-\beta)\log(1-\beta)$ the binary entropy function, and

$\mathcal{H}(\{x_1,\cdots,x_n\}) := -\sum_{i=1}^{n} x_i \log x_i$ the more general entropy function. The standard notation $f(n) = \mathcal{O}(g(n))$ means there exists a constant c (not necessarily positive) such that $|f(n)| \leq cg(n)$, $f(n) = \Theta(g(n))$ means there exist constants c_1 and c_2 such that $c_1 g(n) \leq f(n) \leq c_2 g(n)$, $f(n) = \omega(g(n))$ means that $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$, and $f(n) = o(g(n))$ indicates that $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$. For a matrix $\boldsymbol{A}$, we use $\boldsymbol{A}_{i*}$ and $\boldsymbol{A}_{*i}$ to denote the ith row and ith column of $\boldsymbol{A}$, respectively. We let $[n]$ denote the set $\{1,2,\cdots,n\}$. For any set $\mathcal{A} \subseteq [n]$, we denote by $\binom{\mathcal{A}}{k}$ the set of all k-combinations of $\mathcal{A}$. In particular, we write $\binom{[n]}{k}$ for the set of all k-element subsets of $\{1,2,\cdots,n\}$. We also use $\mathrm{card}(\mathcal{A})$ to denote the cardinality of a set $\mathcal{A}$. Let $\boldsymbol{W}$ be a $p \times p$ random matrix that can be expressed as $\boldsymbol{W} = \Sigma_{i=1}^{n} \boldsymbol{Z}_i \boldsymbol{Z}_i^{\top}$, where $\boldsymbol{Z}_i \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ are jointly independent Gaussian vectors. Then $\boldsymbol{W}$ is said to have a central Wishart distribution with n degrees of freedom and covariance matrix Σ, denoted as $\boldsymbol{W} \sim \mathcal{W}_p(n, \boldsymbol{\Sigma})$.

3.2 Minimax Sampled Capacity Loss

The minimax capacity loss problem can be cast as minimizing $\max_{\boldsymbol{s}\in\mathcal{S}} L_{\boldsymbol{s}}^{\boldsymbol{Q}}$ over all sampling coefficient functions $\boldsymbol{Q}(f)$. In general, the problem is non-convex in $\boldsymbol{Q}(f)$, and hence it is difficult to identify the optimal sampling system. Nevertheless, the minimax capacity loss can be quantified reasonably well at moderate-to-high SNR. It turns out that under both Landau-rate sampling and a large regime of super-Landau sampling, the minimax capacity loss can be well approached by random sampling.

Define the undersampling factor $\alpha := m/n$, and recall that the band sparsity ratio is $\beta := k/n$. Our main results are summarized in the following theorem.

Theorem 3.1. *Consider any sampling coefficient function $\boldsymbol{Q}(\cdot)$ with an undersampling factor α, and let the sparsity ratio be β. Define*

$$\mathsf{SNR}_{\min} := \frac{P}{\beta W} \inf_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}, \tag{3.5}$$

$$\mathsf{SNR}_{\max} := \frac{P}{\beta W} \sup_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}, \tag{3.6}$$

and suppose that $\mathsf{SNR}_{\max} \geq 1$.

(i) (Landau-rate sampling) If $\alpha = \beta$ *(or* $k = m$*), then*

$$\inf_{\boldsymbol{Q}} \max_{\boldsymbol{s} \in \binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{Q}} = \frac{W}{2} \left\{ \mathcal{H}\left(\beta\right) + \mathcal{O}\left(\min \left\{ \sqrt{\frac{\mathsf{SNR}_{\max}}{n}}, \frac{\log n + \log \mathsf{SNR}_{\max}}{n^{1/4}} \right\} \right) + \Delta_{\mathrm{L}} \right\}, \tag{3.7}$$

where

$$-\frac{2}{\sqrt{\mathsf{SNR}_{\min}}} \leq \Delta_{\mathrm{L}} \leq \frac{\beta}{\mathsf{SNR}_{\min}}.$$

(ii) (Super-Landau-rate sampling) Suppose that there is a small constant $\delta > 0$ *such that* $\alpha - \beta \geq \delta$ *and* $1 - \alpha - \beta \geq \delta$*. Then*

$$\inf_{\boldsymbol{Q}} \max_{\boldsymbol{s} \in \binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{Q}} = \frac{W}{2} \left\{ \mathcal{H}\left(\beta\right) - \alpha \mathcal{H}\left(\frac{\beta}{\alpha}\right) + \mathcal{O}\left(\frac{\log n}{n^{1/3}}\right) + \Delta_{\mathrm{SL}} \right\}, \tag{3.8}$$

where

$$-\frac{2}{\sqrt{\mathsf{SNR}_{\min}}} \leq \Delta_{\mathrm{SL}} \leq \frac{\beta}{\mathsf{SNR}_{\min}}.$$

Remark 3.1. Note that $\mathcal{H}(\cdot)$ denotes the binary entropy function. Its appearance is due to the fact that it is a tight estimate of the rate function of binomial coefficients.

Theorem 3.1 provides a tight characterization of the minimax sampled capacity loss relative to the Nyquist-rate capacity, under both Landau-rate sampling and super-Landau-rate sampling. Note that the Landau-rate sampling regime in (i) is not a special case of the super-Landau-rate regime considered in (ii). For instance, if $\beta > 1/2$, then $\alpha + \beta > 1$, which falls within a regime not accounted for by Theorem 3.1(ii).

The capacity loss expressions (3.7) and (3.8) contain residual terms that vanish for large n and high SNR. In the regime where $\frac{\mathsf{SNR}_{\max}}{n} \ll 1$ and $\mathsf{SNR}_{\min} \gg 1$, these fundamental minimax limits are approximately equal to a constant modulo a vanishing residual term. Since the Nyquist rate capacity scales as $\Theta\left(W \log \mathsf{SNR}\right)$, our results indicate that the ratio of the minimax capacity loss to the Nyquist-rate capacity vanishes at a rate $\Theta\left(1/\log \mathsf{SNR}\right)$.

Note that even if we allow power control at the transmitter side, the results are still valid at high SNR. This is summarized in the following theorem.

Theorem 3.2. *Consider the metric $L_{\boldsymbol{s}}^{\boldsymbol{Q},\mathrm{opt}}$ with power control. Under all other conditions of Theorem 3.1, the bounds (3.7) and (3.8) (with $L_{\boldsymbol{s}}^{\boldsymbol{Q}}$ replaced by $L_{\boldsymbol{s}}^{\boldsymbol{Q},\mathrm{opt}}$) continue to hold if Δ_{L} and Δ_{SL} are respectively replaced by residuals $\Delta_{\mathrm{L}}^{\mathrm{opt}}$ and $\Delta_{\mathrm{SL}}^{\mathrm{opt}}$ that obey*

$$-\frac{2}{\sqrt{\mathsf{SNR}_{\min}}} \leq \Delta_{\mathrm{L}}^{\mathrm{opt}}, \Delta_{\mathrm{SL}}^{\mathrm{opt}} \leq \frac{\beta+\overline{A}}{\mathsf{SNR}_{\min}},$$

where $\overline{A}$ is a constant defined as

$$\overline{A} := \min\left\{ \frac{\max_{\boldsymbol{s}\in\binom{[n]}{k}} \int_0^W \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)} \mathrm{d}f}{\beta W \inf_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}}, \frac{\sup_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}}{\inf_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}} \right\}.$$

Theorem 3.2 demonstrates that if the average-to-minimum ratio $\frac{\overline{\mathsf{SNR}}}{\mathsf{SNR}_{\min}}$ is bounded by a constant (where $\overline{\mathsf{SNR}} := \max_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{P}{\beta W} \int_0^W \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)} \mathrm{d}f$), then the minimax sampled capacity gap without power control remains almost the same as that with optimal power allocation within a vanishingly small gap per unit bandwidth, under the conditions of high SNR and large n. Note that the constant $\overline{A}$ given in Theorem 3.2 is fairly conservative, and can be refined with finer tuning or algebraic techniques.

Theorem 3.2 can be delivered as an immediate consequence of Theorem 3.1 if we can quantify the gap between $C_{\boldsymbol{s}}^{P_{\mathrm{eq}}}$ and $C_{\boldsymbol{s}}^{\mathrm{opt}}$. In fact, the capacity benefits of using power control at high SNR regime is no larger than $\mathcal{O}(\mathsf{SNR}^{-1})$ per unit bandwidth. See Appendix B.1 for details. For this reason, our analysis is mainly devoted to $L_{\boldsymbol{s}}^{\boldsymbol{Q}}$, which corresponds to the capacity loss under uniform power allocation.

The proof of Theorem 3.1 involves the verification of two parts: a converse part that provides a fundamental lower bound on the minimax sampled capacity loss, and an achievability part that provides a sampling scheme to approach this bound. As we show, the class of sampling systems with periodic modulation followed by low-pass filters, as illustrated in Fig. 3.1(b), is sufficient to approach the minimax loss.

Throughout the remainder of the chapter, we suppose that the noise is of unit power spectral density $\mathcal{S}_\eta(f) \equiv 1$ unless otherwise specified. Note that this incurs no loss of generality since we can always include a noise-whitening LTI filter at the first stage of the sampling system.

3.2.1 The Converse

We need to show that the minimax sampled capacity loss under any channel-independent sampler cannot be lower than (3.7) and (3.8). This is given by the following theorem, which takes into account the entire regime including the situation where $\alpha + \beta > 1$.

Theorem 3.3. *Consider any Riemann-integrable sampling coefficient function* $\boldsymbol{Q}(\cdot)$ *with an undersampling factor* $\alpha := m/n$*. Suppose the sparsity ratio* $\beta := k/n$ *satisfies* $\beta \leq \alpha \leq 1$*. The minimax capacity loss can be lower bounded by*

$$\inf_{\boldsymbol{Q}} \max_{\boldsymbol{s} \in \binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{Q}} \geq \frac{W}{2} \left\{ \mathcal{H}\left(\beta\right) - \alpha \mathcal{H}\left(\frac{\beta}{\alpha}\right) - \frac{2}{\sqrt{\mathsf{SNR}_{\min}}} - \frac{\log\left(n+1\right)}{n} \right\}. \tag{3.9}$$

For a given β, the bound is decreasing in α. While the active channel bandwidth is smaller than the total bandwidth, the noise (even though the SNR is large) is scattered over the entire bandwidth. Thus, none of the universal sub-Nyquist sampling strategies are information preserving, and increasing the sampling rate can always harvest capacity gain.

3.2.2 Achievability with Landau-rate Sampling ($\alpha = \beta$)

Consider the achievability part when the sampling rate equals the active frequency bandwidth ($\beta = \alpha$). In general, it is computationally intractable to find a deterministic solution to approach the lower bound (3.9). A special instance of sampling methods that we can analyze concerns the case in which the sampling coefficient functions are flat over $[0, W/n]$ and whose coefficients are generated in a random fashion. It turns out that as n grows large, the capacity loss achievable by random sampling approaches the lower bound (3.9) uniformly across all realizations. The results are stated in Theorem 3.4 after introducing a class of sub-Gaussian measure below.

Definition 3.3. *A measure ν on $\mathbb{R}$ satisfies the logarithmic Sobolev inequality (LSI) with constant c_{LS} if, for any differentiable function g,*

$$\int g^2 \log \frac{g^2}{\int g^2 \mathrm{d}\nu} \mathrm{d}\nu \leq 2c_{LS} \int |g'|^2 \, \mathrm{d}\nu.$$

Remark 3.2. A probability measure obeying the LSI possesses sub-Gaussian tails, and a large class of sub-Gaussian measures satisfies this inequality for some constant. See [28, 84] for examples. In particular, the standard Gaussian measure satisfies this inequality with constant $c_{\mathrm{LS}} = 1$ (e.g. [85]).

Theorem 3.4. *Let $\boldsymbol{M} \in \mathbb{R}^{k \times n}$ be a random matrix such that $\boldsymbol{M}_{ij}$'s are jointly independent symmetric random variables with zero mean and unit variance. In addition, suppose that $\boldsymbol{M}_{ij}$ satisfies either of the following conditions:*

(a) $\boldsymbol{M}_{ij}$ is bounded in magnitude by an absolute constant D;

(b) The probability measure of $\boldsymbol{M}_{ij}$ satisfies the LSI with a bounded constant c_{LS}.

If $\mathsf{SNR}_{\max} \geq 1$, then there exist absolute constants $c_0, c_1, C > 0$ such that

$$\max_{\boldsymbol{s} \in \binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{M}} \leq \frac{W}{2} \left(\mathcal{H}(\beta) + c_1 \min \left\{ \frac{\sqrt{\mathsf{SNR}_{\max}}}{\sqrt{n}}, \frac{\log n + \log \mathsf{SNR}_{\max}}{n^{1/4}} \right\} + \frac{\beta}{\mathsf{SNR}_{\min}} \right) \tag{3.10}$$

with probability exceeding $1 - C \exp(-c_0 n)$.

Theorem 3.4 demonstrates that independent random sampling achieves minimax sampled capacity loss, which is $\frac{1}{2}\mathcal{H}(\beta)$ per unit bandwidth modulo some vanishing residual term. In fact, our analysis demonstrates that the sampled capacity loss approaches the minimax limit uniformly over all states $\boldsymbol{s}$. Another interesting observation is the universality phenomenon, i.e. a broad class of sub-Gaussian ensembles, as long as the entries are jointly independent with matching moments, suffices to generate minimax samplers.

3.2.3 Achievability with Super-Landau-Rate Sampling ($\alpha > \beta, \alpha + \beta < 1$)

So far we have considered the case in which the sampling rate is equal to the spectral support. While the active bandwidth for transmission is smaller than the total bandwidth, the noise (even though the SNR is large) is scattered over the entire bandwidth. This indicates that none of the sub-Nyquist sampling strategies preserves all information content conveyed through the noisy channel, unless they know the channel support. One may thus hope that increasing the sampling rate would improve the achievable information rate. The achievability result for super-Landau-rate sampling is stated in the following theorem.

Theorem 3.5. *Let $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ be a Gaussian matrix such that $\boldsymbol{M}_{ij}$'s are independently drawn from $\mathcal{N}(0,1)$. Suppose that $1-\alpha-\beta \geq \varepsilon$ and $\alpha - \beta \geq \varepsilon$ for some small constant $\varepsilon > 0$. Then there exist universal constants $C, c > 0$ such that*

$$\max_{\boldsymbol{s}\in\binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{M}} \leq \frac{W}{2}\left[\mathcal{H}(\beta) - \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) + \frac{c\log n}{n^{1/3}} + \frac{\beta}{\mathsf{SNR}_{\min}}\right]$$

with probability exceeding $1 - C\exp(-n)$.

Theorem 3.5 indicates that i.i.d. Gaussian sampling approaches the minimax capacity loss (which is about $\frac{1}{2}\mathcal{H}(\beta) - \frac{1}{2}\alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right)$ per Hertz) to within a vanishingly small gap. As will be shown in our proof, with exponentially high probability, the sampled capacity loss for all states are equivalent and coincide with the fundamental minimax limit. In contrast to Theorem 3.4, we restrict our attention to Gaussian sampling, which suffices for the proof of Theorem 3.1.

3.3 Equivalent Algebraic Problems

Our main results in Section 3.2 are established by investigating three equivalent algebraic problems. Recall that $\frac{P}{\beta W}\boldsymbol{H}_{\boldsymbol{s}}^2 \succeq \mathrm{SNR}_{\min}\boldsymbol{I}_k$. Define $\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}} := (\boldsymbol{Q}\boldsymbol{Q}^*)^{-\frac{1}{2}}\boldsymbol{Q}_{\boldsymbol{s}}$. Simple manipulations yield

$$
\begin{aligned}
L_{\boldsymbol{s}}^{\boldsymbol{Q}} &= -\frac{1}{2}\int_{0}^{W/n} \log\det\left(\boldsymbol{I}_m + \frac{P}{\beta W}\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}(f)\boldsymbol{H}_{\boldsymbol{s}}^{2}(f)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}(f)\right)\mathrm{d}f \\
&\quad + \frac{1}{2}\int_{0}^{W/n} \log\det\left(\boldsymbol{I}_k + \frac{P}{\beta W}\boldsymbol{H}_{\boldsymbol{s}}^{2}(f)\right)\mathrm{d}f \\
&= -\frac{1}{2}\int_{0}^{W/n} \log\det\left(\boldsymbol{I}_k + \frac{P}{\beta W}\boldsymbol{H}_{\boldsymbol{s}}(f)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}(f)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}(f)\boldsymbol{H}_{\boldsymbol{s}}(f)\right)\mathrm{d}f \qquad (3.11) \\
&\quad + \frac{1}{2}\int_{0}^{W/n} \log\det\left(\frac{P}{\beta W}\boldsymbol{H}_{\boldsymbol{s}}^{2}(f)\right)\mathrm{d}f + \frac{\beta W}{2}\Delta_{\boldsymbol{s}} \qquad (3.12) \\
&= -\frac{1}{2}\int_{0}^{W/n} \log\det\left(\frac{\beta W}{P}\boldsymbol{H}_{\boldsymbol{s}}^{-2}(f) + \boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}(f)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}(f)\right)\mathrm{d}f + \frac{\beta W}{2}\Delta_{\boldsymbol{s}}, \qquad (3.13)
\end{aligned}
$$

where $\Delta_{\boldsymbol{s}}$ denotes some residual term. In particular, $\Delta_{\boldsymbol{s}}$ can be bounded as

$$0 \leq \Delta_{\boldsymbol{s}} \leq \frac{1}{\mathsf{SNR}_{\min}}. \qquad (3.14)$$

This is an immediate consequence of the following observation: for any $k \times k$ positive semidefinite matrix $\boldsymbol{A}$,

$$0 \leq \frac{1}{k}\log\det\left(\boldsymbol{I}_k + \boldsymbol{A}\right) - \frac{1}{k}\log\det\left(\boldsymbol{A}\right) = \frac{1}{k}\sum_{i=1}^{k}\log\left(1 + \frac{1}{\lambda_i\left(\boldsymbol{A}\right)}\right) \leq \frac{1}{\lambda_{\min}(\boldsymbol{A})}. \qquad (3.15)$$

Recall that $\mathsf{SNR}_{\min} := \frac{P}{\beta W}\inf_{0\leq f\leq W}\left|H(f)\right|^2$ and $\mathsf{SNR}_{\max} := \frac{P}{\beta W}\sup_{0\leq f\leq W}\left|H(f)\right|^2$. Therefore, $\frac{\beta W}{P}\boldsymbol{H}_{\boldsymbol{s}}^{-2}$ can be bounded as

$$\frac{1}{\mathsf{SNR}_{\max}}\boldsymbol{I}_k \preceq \frac{\beta W}{P}\boldsymbol{H}_{\boldsymbol{s}}^{-2} \preceq \frac{1}{\mathsf{SNR}_{\min}}\boldsymbol{I}_k. \qquad (3.16)$$

This bound together with (3.13) makes $\det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}\right)$ a quantity of interest (for some small ϵ). In the sequel, we provide tight bounds on several algebraic problems, which in turn establish Theorems 3.3-3.5. The proofs of these results rely heavily on *non-asymptotic* (random) matrix theory. In particular, the proofs for achievability bounds are established based on measure concentration of log-determinant functions, which we provide as well in this section.

3.3.1 The Converse

Note that $\boldsymbol{Q}^{\mathrm{w}}(f)$ has orthonormal rows. The following theorem investigates the properties of $\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{B}_{\boldsymbol{s}}^{*}\boldsymbol{B}_{\boldsymbol{s}}\right)$ for any $m \times n$ matrix $\boldsymbol{B}$ that has orthonormal rows. This establishes Theorem 3.3.

Theorem 3.6. *(1) Consider any $m \times n$ matrix $\boldsymbol{B}$ $(n \geq m \geq k)$ that satisfies $\boldsymbol{B}\boldsymbol{B}^{*} = \boldsymbol{I}_m$, and denote by $\boldsymbol{B}_{\boldsymbol{s}}$ the $m \times k$ submatrix of $\boldsymbol{B}$ with columns coming from the index set $\boldsymbol{s}$. Then for any $\epsilon > 0$, one has*

$$\binom{m}{k} \leq \sum_{\boldsymbol{s}\in\binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I}_m + \boldsymbol{B}_{\boldsymbol{s}}^{*}\boldsymbol{B}_{\boldsymbol{s}}\right) = \sum_{l=0}^{k} \binom{n-l}{k-l}\binom{m}{l}\epsilon^{k-l} \tag{3.17}$$

$$\leq \binom{m}{k}\left(1+\sqrt{\epsilon}\right)^{n+k}. \tag{3.18}$$

(2) For any positive integer p, suppose that $\boldsymbol{B}_1, \cdots, \boldsymbol{B}_p$ are all $m \times n$ matrices such that $\boldsymbol{B}_i\boldsymbol{B}_i^{*} = \boldsymbol{I}_m$. Then,

$$\min_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{1}{np}\sum_{i=1}^{p} \log\det\left(\epsilon \boldsymbol{I}_k + \left(\boldsymbol{B}_i\right)_{\boldsymbol{s}}^{*}\left(\boldsymbol{B}_i\right)_{\boldsymbol{s}}\right) \leq \frac{1}{n}\log\binom{m}{k} - \frac{1}{n}\log\binom{n}{k} + 2\sqrt{\epsilon} \tag{3.19}$$

$$\leq \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) - \mathcal{H}\left(\beta\right) + 2\sqrt{\epsilon} + \frac{\log\left(n+1\right)}{n}. \tag{3.20}$$

Note that $\boldsymbol{Q}^{\mathrm{w}}(f)$ has orthonormal rows for any f, and $\boldsymbol{Q}^{\mathrm{w}}(f)$ is assumed to be Riemann integrable. For any $\delta > 0$, we can find a sufficiently large p such that

$$\int_0^{W/n} \log\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}(f)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}(f)\right)\mathrm{d}f \leq \delta + \frac{W}{pn}\sum_{i=1}^{p} \log\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}\left(\frac{iW}{pn}\right)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}\left(\frac{iW}{pn}\right)\right).$$

Since δ can be arbitrarily small, applying Theorem 3.6 immediately yields that for any $\boldsymbol{Q}(\cdot)$:

$$\min_{\boldsymbol{s}\in\binom{[n]}{k}} \int_0^{W/n} \log\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}(f)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}(f)\right)\mathrm{d}f \leq W\left\{\alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) - \mathcal{H}\left(\beta\right) + 2\sqrt{\epsilon} + \frac{\log\left(n+1\right)}{n}\right\}.$$

This together with (3.13), (3.14) and (3.16) leads to

$$\inf_{\boldsymbol{Q}} \max_{\boldsymbol{s}\in\binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{Q}} \geq \frac{W}{2}\left\{\mathcal{H}(\beta) - \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) - \frac{2}{\sqrt{\mathsf{SNR}_{\min}}} - \frac{\log(n+1)}{n}\right\},$$

which completes the proof of Theorem 3.3.

3.3.2 Achievability (Landau-rate Sampling)

When it comes to the achievability part, the major step is to quantify $\det(\epsilon\boldsymbol{I}_k + (\boldsymbol{M}\boldsymbol{M}^\top)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top)$ for every $\boldsymbol{s} \in \binom{[n]}{k}$. Interestingly, this quantity can be uniformly controlled due to the concentration of spectral measure phenomenon [28]. This is stated in the following theorem.

Theorem 3.7. *Let $\boldsymbol{M} \in \mathbb{R}^{k\times n}$ be a random matrix of independent entries, and let $0 < \epsilon < 1$ denote some constant. Under the conditions of Theorem 3.3, there exist absolute constants c_0, c_1, $C > 0$ such that*

$$\begin{aligned}-\mathcal{H}(\beta) - c_1 \min\left\{\frac{1}{\sqrt{n\epsilon}}, \frac{\log n + \log\frac{1}{\epsilon}}{n^{1/4}}\right\} &\leq \min_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{1}{n}\log\det\left(\epsilon\boldsymbol{I}_k + \left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right)\\ &\leq -\mathcal{H}(\beta) + 2\sqrt{\epsilon} + \frac{\log(n+1)}{n} \qquad (3.21)\end{aligned}$$

with probability exceeding $1 - C\exp(-c_0 n)$.

Putting Theorem 3.7 and equations (3.13), (3.14) and (3.16) together gives

$$\max_{\boldsymbol{s}\in\binom{[n]}{k}} L_{\boldsymbol{s}}^{\boldsymbol{M}} \leq \frac{W}{2}\left[\mathcal{H}(\beta) + \mathcal{O}\left(\min\left\{\sqrt{\frac{\mathsf{SNR}_{\max}}{n}}, \frac{\log n + \log\mathsf{SNR}_{\max}}{n^{1/4}}\right\}\right) + \frac{\beta}{\mathsf{SNR}_{\min}}\right]$$

with exponentially high probability, as claimed in Theorem 3.4.

3.3.3 Achievability (Super-Landau-rate Sampling)

Instead of studying a large class of sub-Gaussian random ensembles[2], the following theorem focuses on i.i.d. Gaussian matrices, which establishes the optimality of Gaussian random sampling for the super-Landau regime.

Theorem 3.8. *Let $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ be an i.i.d. random matrix satisfying $\boldsymbol{M}_{ij} \sim \mathcal{N}(0,1)$. Suppose that $1-\alpha-\beta \geq \zeta$ and $\alpha-\beta \geq \zeta$ for some small constant $\zeta > 0$. Then there exist absolute constants c, $C > 0$ such that*

$$\begin{aligned} -\mathcal{H}(\beta) + \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) + \frac{c\log n}{n^{1/3}} &\leq \min_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{1}{n}\log\det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{M}_{\boldsymbol{s}}^{\top}\left(\boldsymbol{M}\boldsymbol{M}^{\top}\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right) \\ &\leq -\mathcal{H}\left(\beta\right) + \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) + 2\sqrt{\epsilon} + \frac{\log\left(n+1\right)}{n} \end{aligned}$$

with probability at least $1 - C\exp\left(-n\right)$.

Combining Theorem 3.8 and equations (3.13), (3.14) and (3.16) implies that

$$L_{\boldsymbol{s}}^{\boldsymbol{M}} \leq \frac{W}{2}\left[\mathcal{H}(\beta) - \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) + \mathcal{O}\left(\frac{\log n}{n^{1/3}}\right) + \frac{\beta}{\mathsf{SNR}_{\min}}\right]$$

with exponentially high probability.

3.3.4 Measure Concentration of Log-Determinant for Random Matrices

It has been shown above that the achievability bounds can be established by demonstrating measure concentration of certain log-determinant functions. In fact, the concentration of log-determinant has been studied in the random matrix literature as a key ingredient in demonstrating universality laws for spectral statistics (e.g. [86, Proposition 48]). However, these bounds are only shown to hold with overwhelming probability (i.e. with probability at least $1 - e^{-\omega(\log n)}$), which are not sharp enough

[2]The proof argument for Landau-rate sampling cannot be readily carried over to super-Landau regime since $\boldsymbol{M}_{\boldsymbol{s}}$ is now a tall matrix, and hence we cannot separate $\boldsymbol{M}_{\boldsymbol{s}}$ and $\boldsymbol{M}\boldsymbol{M}^*$ easily.

for our purpose. As a result, we provide sharper measure concentration results of log-determinant in this subsection.

Lemma 3.1. *Suppose that $k/p \in (0,1]$, and that $\epsilon > 0$. Consider a random matrix $\boldsymbol{A} = \mathbb{R}^{k\times p}$ where $\boldsymbol{A}_{ij}$ are symmetric and jointly independent with zero mean and unit variance.*

(a) If $\boldsymbol{A}_{ij}$'s are bounded in magnitude by a constant D, then for any $\delta > \frac{8D\sqrt{\pi}}{\sqrt{\epsilon}}$:

$$\mathbb{P}\left(\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right) < \mathbb{E}\left[\log\det\left(\frac{2\epsilon}{e}\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] - \delta\right) \leq 4\exp\left(-\frac{\epsilon\delta^2}{16D^2}\right); \tag{3.22}$$

$$\mathbb{P}\left(\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right) > \mathbb{E}\left[\log\det\left(\frac{e\epsilon}{2}\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] + \delta\right) \leq 4\exp\left(-\frac{\epsilon e\delta^2}{32D^2}\right). \tag{3.23}$$

(b) If $\boldsymbol{A}_{ij}$'s satisfy the LSI with uniformly bounded constant c_{LS}, then for any $\delta > 0$:

$$\mathbb{P}\left(\left|\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right) - \mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]\right| > \delta\right) \leq 2\exp\left(-\frac{\epsilon\delta^2}{2c_{\mathrm{LS}}}\right). \tag{3.24}$$

Proof. See Appendix B.5. □

The concentration results for $\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)$ will be useful if we can quantify $\mathbb{E}\left[\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]$, as established in the following lemma.

Lemma 3.2. *Let $\boldsymbol{A} \in \mathbb{R}^{k\times k}$ be a random matrix such that all entries are jointly independent with zero mean and unit variance. For any $\epsilon \in \left(\frac{4}{k}, \frac{1}{e^2}\right)$, we have*

$$\begin{aligned}\mathbb{E}\left[\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] &\leq \frac{1}{k}\log\mathbb{E}\left[\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] \\ &\leq -1 + \frac{1.5\log(ek)}{k} + 2\sqrt{\epsilon}\log\frac{1}{\epsilon}.\end{aligned} \tag{3.25}$$

Additionally, under Condition (a) or (b) of Lemma 3.1, one has, for any $\frac{1}{k} < \epsilon \leq 0.8$,

$$\mathbb{E}\left[\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] \geq -1+\frac{\log k}{2k}-\mathcal{O}\left(\frac{1}{k\epsilon}\right). \tag{3.26}$$

In particular, if $\boldsymbol{A}_{ij} \sim \mathcal{N}(0,1)$, *then for any* $0 < \epsilon \leq 0.8$,

$$\mathbb{E}\left[\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] \geq -1+\frac{\log k}{2k}-\frac{2}{k\epsilon}. \tag{3.27}$$

Proof. See Appendix B.6. □

One issue concerning Lemmas 3.1 and 3.2 is that the bounds might become useless when ϵ is extremely small. To mitigate this issue, we obtain another measure concentration lower bound in the following lemma.

Lemma 3.3. *Let* $\boldsymbol{A} \in \mathbb{R}^{k\times k}$ *be a random matrix satisfying the assumptions of Lemma 3.1. There exists an absolute constant* $c_{14} > 0$ *such that for any* $\frac{1}{\sqrt{k}} \leq \tau < k$ *and* $0 < \epsilon < 1$,

$$\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \geq -1-\frac{c_{14}\tau^{1/4}}{k^{1/4}}\left(\log\frac{1}{\epsilon}+\log k\right)$$

with probability at least $1-4\exp(-\tau k)$.

Proof. See Appendix B.7. □

Lemmas 3.1, 3.2, and 3.3 taken collectively allow us to derive the measure concentration of $\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right)$ uniformly over all $\boldsymbol{s} \in \binom{[n]}{k}$ as follows.

Lemma 3.4. *Let* $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ *be a random matrix satisfying the conditions of Theorem 3.3. Then for any constant* $0 < \epsilon < 1$,

$$\forall \boldsymbol{s}: \ \frac{1}{n}\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right) \geq -1-\mathcal{O}\left(\min\left\{\frac{1}{\sqrt{n\epsilon}},\frac{\log\frac{1}{\epsilon}+\log n}{n^{1/4}}\right\}\right) \tag{3.28}$$

holds with probability exceeding $1-2\left(\frac{2}{e}\right)^{n}$.

In particular, if $\boldsymbol{M}_{ij} \sim \mathcal{N}(0,1)$ are jointly independent, then for any $0 < \epsilon \leq 0.8$,

$$\forall \boldsymbol{s}: \ \frac{1}{k} \log\det\left(\epsilon \boldsymbol{I} + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right) \geq -1 - \frac{2}{k\epsilon} - \sqrt{\frac{2}{\beta k \epsilon}} \tag{3.29}$$

with probability at least $1 - \left(\frac{2}{e}\right)^n$.

Proof. Fix $\delta = c_5\sqrt{k/\epsilon}$ for some numerical constant $c_5 > 0$ in Lemma 3.1. Combining Lemmas 3.1, 3.2, and 3.3 yields the following: under the assumptions of Lemma 3.1, for any small constant $\epsilon > 0$ and some sufficiently large c_5 we have

$$\frac{1}{k} \log\det\left(\epsilon \boldsymbol{I} + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right) \geq -1 + \frac{\log k}{2k} - \mathcal{O}\left(\min\left\{\frac{1}{\sqrt{n\epsilon}}, \frac{\log\frac{1}{\epsilon} + \log n}{n^{1/4}}\right\}\right) \tag{3.30}$$

with probability exceeding $1 - 4e^{-n}$. Since there are at most $\binom{n}{k} \leq \frac{1}{2} \cdot 2^n$ different $\boldsymbol{s}$, applying the union bound establishes (3.28). Similarly, for the ensemble $\boldsymbol{M}_{ij} \sim \mathcal{N}(0,1)$, Lemmas 3.1 and 3.2 indicate that

$$\frac{1}{k} \log\det\left(\epsilon \boldsymbol{I} + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right) \geq -1 + \frac{\log k}{2k} - \frac{2}{k\epsilon} - \sqrt{\frac{2}{\beta k \epsilon}}$$

with probability at least $1-2e^{-n}$. The proof is then complete via the union bound. □

Another important class of log-determinant function takes the form of $\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)$. The concentration of such functions for i.i.d. rectangular random matrices is characterized in the following lemmas.

Lemma 3.5. *Let $\alpha := m/n$ be a fixed constant independent from (m,n), and assume that $\alpha \in [\delta, 1-\delta]$ for some small constant $\delta > 0$. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ be a random matrix such that all entries are symmetric and jointly independent with zero mean and unit variance. Under Condition (a) or (b) of Lemma 3.1, there exist universal constants $c_7, C_7 > 0$ independent of n such that*

$$\left|\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - (1-\alpha)\log\frac{1}{1-\alpha} + \alpha\right| \leq \frac{1}{\sqrt{n}} \tag{3.31}$$

with probability exceeding $1 - C_7\exp\left(-c_7 n\right)$.

Proof. See Appendix B.8. □

This result can be made more explicit for Gaussian ensembles as follows.

Lemma 3.6. *Suppose that $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ is a random matrix with independent entries satisfying $\boldsymbol{A}_{ij} \sim \mathcal{N}(0,1)$. Suppose that there exists a small constant $\delta > 0$ such that $\alpha := \frac{m}{n} \in [\delta, 1-\delta]$.*

(1) For any $\epsilon > 0$ and any $\tau > 0$,

$$\frac{\mathrm{card}\left\{i \mid \lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) < \epsilon\right\}}{n} < \frac{\alpha}{1-\alpha-\frac{1}{n}}\epsilon + \frac{4\sqrt{\alpha\tau}}{\sqrt{n\epsilon}} \tag{3.32}$$

with probability exceeding $1 - 2\exp(-\tau n)$.

(2) For any $n > \max\left\{\frac{2}{1-\sqrt{\alpha}}, 7\right\}$ and any $\tau > 0$,

$$\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) \le (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{3\log n}{n} + \frac{5\sqrt{\alpha}}{\left(1-\sqrt{\alpha}-\frac{2}{n}\right)}\frac{\tau}{n} \tag{3.33}$$

with probability exceeding $1 - 2\exp(-2\tau^2)$.

(3) For any $n > \max\left\{\frac{6.414}{1-\alpha}\cdot e^{\frac{\tau}{1-\alpha}}, \left(\frac{\alpha}{1-\alpha-\frac{1}{n}} + 4\sqrt{\alpha\tau}\right)^3\right\}$ and any $\tau > 0$,

$$\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) > (1-\alpha)\log\frac{1}{1-\alpha} - \alpha - \left(\frac{2\alpha}{1-\alpha-\frac{1}{n}} + 11\sqrt{\alpha\tau}\right)\frac{\log n}{n^{1/3}} - \frac{6\alpha}{n^{2/3}} \tag{3.34}$$

with probability exceeding $1 - 5\exp(-\tau n)$.

Proof. See Appendix B.9. □

The last log-determinant function considered here takes the form of $\log\det\left(\epsilon\boldsymbol{I} + \boldsymbol{A}^\top\boldsymbol{B}^{-1}\boldsymbol{A}\right)$ for some independent random matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, as stated in the following lemma.

Lemma 3.7. *Suppose that $m > k$ and m/k is bounded away from[3] 1. Let matrix $\boldsymbol{A} = \mathbb{R}^{m\times k}$ and $\boldsymbol{B} \in \mathbb{R}^{m\times m}$ be two independent random matrices such that $\boldsymbol{A}_{ij} \sim$*

[3]In other words, there exists a small constant $\delta > 0$ such that $m/k \ge 1 + \delta$.

$\mathcal{N}(0,1)$ *are jointly independent, and* $\boldsymbol{B} \sim \mathcal{W}_m(n-k, \boldsymbol{I}_m)$. *Then for any* $\tau > \frac{1}{n^{1/3}}$,

$$\begin{aligned}\frac{1}{n}\log\det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{A}^\top\boldsymbol{B}^{-1}\boldsymbol{A}\right) \geq & -(\alpha-\beta)\log(\alpha-\beta) + \alpha\log\alpha \\ & + (1-\alpha-\beta)\log\left(1-\frac{\beta}{1-\alpha}\right) - \beta\log(1-\alpha) - \frac{c_8\tau\log n}{n^{1/3}}\end{aligned}$$

with probability exceeding $1 - C_8\exp(-\tau^2 n)$ *for some absolute constants* $c_8, C_8 > 0$.

Proof. See Appendix B.10. □

3.4 Discussion

3.4.1 The Converse

Our analysis demonstrates that at high SNR, the loss $L_{\boldsymbol{s}}^{\boldsymbol{Q}}$ depends almost solely on the quantity

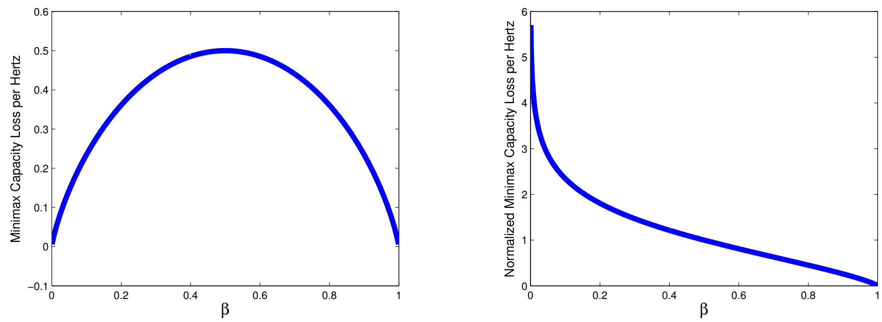
$$d(\boldsymbol{Q}(f), \boldsymbol{s}, \epsilon) := \det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}*}(f)\boldsymbol{Q}_{\boldsymbol{s}}^{\mathrm{w}}(f)\right)$$

for small $\epsilon > 0$, which is approximately the exponential of capacity loss at a given pair $(\boldsymbol{s}, f)$. The key observation underlying the proof is that for any f, the sum

$$\sum_{\boldsymbol{s}\in\binom{[n]}{k}} d(\boldsymbol{Q}(f), \boldsymbol{s}, \epsilon)$$

is a *constant* independent of the sampling coefficient function $\boldsymbol{Q}$. In other words, at any given f, the exponential sum of capacity loss over all states $\boldsymbol{s}$ is invariable regardless of what samplers we employ.

This invariant quantity is critical in identifying the minimax sampling method. In fact, it motivates us to seek a sampling method that achieves equivalent performance over all states $\boldsymbol{s}$. Large random matrices exhibit sharp concentration of spectral measure, and hence become a natural candidate to attain minimaxity.



(a) (b)

Figure 3.4: Plot (a) illustrates $\mathcal{H}(\beta)/2$ v.s. the sparsity ratio β, which characterizes the fundamental minimax capacity loss pert Hertz within a vanishing gap. Plot (b) illustrates $\mathcal{H}(\beta)/(2\beta)$ v.s. β, which corresponds approximately to the normalized capacity loss per Hertz.

3.4.2 Landau-rate Sampling

When sampling is performed at the Landau rate, the minimax capacity loss per unit Hertz is almost solely determined by the entropy function $\mathcal{H}(\beta)$. Specifically, when n and k are sufficiently large, the minimax limit depends only on the sparsity ratio $\beta = \frac{k}{n}$ rather than (n, k). Some implications from Theorems 3.3 - 3.4 are as follows.

1. The capacity loss per unit Hertz is illustrated in Fig. 3.4(a). The capacity loss vanishes when $\beta \to 1$, since Nyquist-rate sampling is information preserving. The capacity loss divided by β is plotted in Fig. 3.4(b), which provides a normalized view of the capacity loss. It can be seen that the normalized loss decreases monotonically with β, indicating that the loss is more severe in sparse channels. Note that this is different from an LTI channel whereby sampling at the Landau rate is sufficient to preserve all information. When the channel state is uncertain, increasing the sampling rate above the Landau rate (but below the Nyquist rate) effectively increases the SNR, and hence allows more information to be harvested from the noisy sampled output.

2. The capacity loss incurred by independent random sampling meets the fundamental minimax limit for Landau-rate sampling uniformly across all states $\boldsymbol{s}$,

which reveals that with exponentially high probability, random sampling is optimal in terms of universal sampling design. The capacity achievable by random sampling exhibits very sharp concentration around the minimax limit uniformly across all states $\boldsymbol{s} \in \binom{[n]}{k}$.

3. A *universality* phenomenon that arises in large random matrices (e.g. [87]) leads to the fact that the minimaxity of random sampling matrices does not depend on the particular distribution of the coefficients. For a large class of sub-Gaussian measure, as long as all entries are jointly independent with matching moments up to the second order, the sampling mechanism it generates is minimax with exponentially high probability.

3.4.3 Super-Landau-Rate Sampling

The random sampling analyzed in Theorem 3.5 only involves Gaussian random sampling, and we have not shown universality results. Some implications under super-Landau-rate sampling are as follows.

1. Similar to the case with Landau-rate sampling, Gaussian sampling achieves minimax capacity loss uniformly across all states $\boldsymbol{s}$ within a large super-Landau-rate regime. The capacity gap is illustrated in Fig. 3.5. It can be observed from the plot that increasing the α/β ratio improves the capacity gap, shrinks the locus and shifts it leftwards.

2. Theorem 3.5 only concerns i.i.d. Gaussian random sampling instead of more general independent random sampling. While we conjecture that the universality phenomenon continues to hold for other jointly independent random ensembles with sub-Gaussian tails and appropriate matching moments, the mathematical analysis turns out to be more tricky than for Landau-rate sampling.

3. The capacity gain by sampling above the Landau rate depends on the undersampling factor α as well. Specifically, the capacity benefit per unit bandwidth due to super-Landau sampling is captured by the term $\frac{1}{2}\alpha\mathcal{H}\left(\beta/\alpha\right)$. When $\alpha \to 1$,

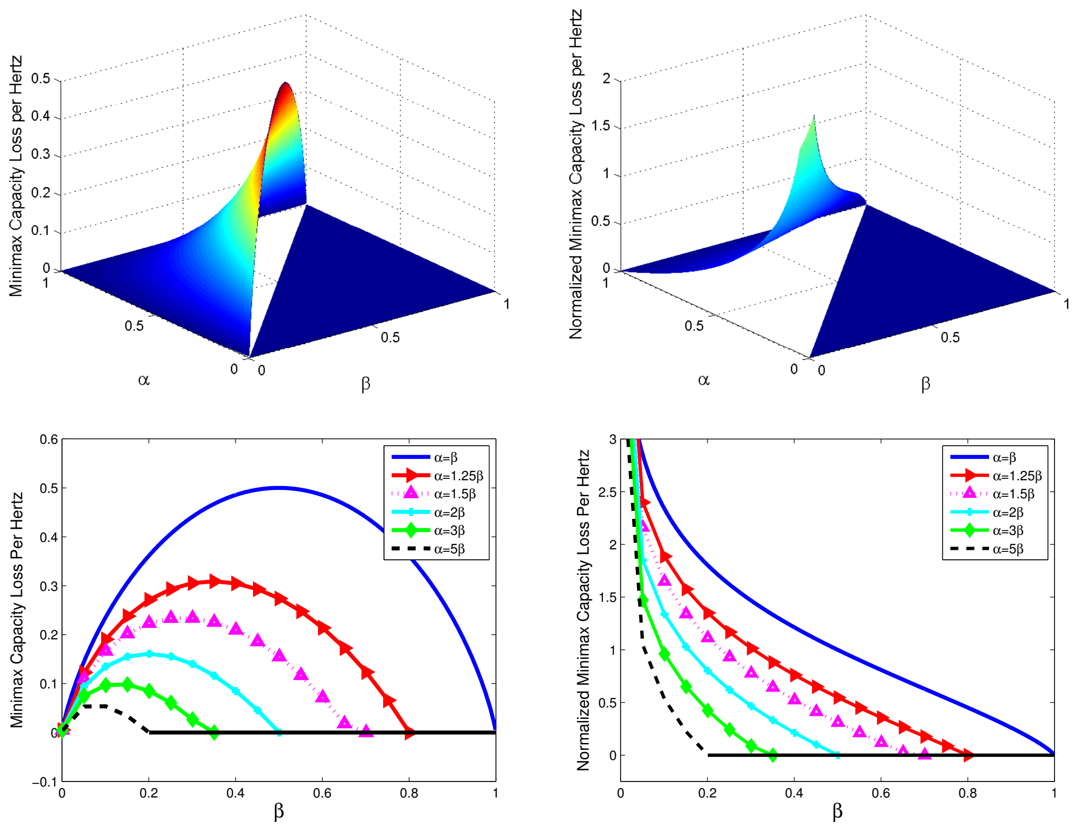


Figure 3.5: The function $\frac{1}{2}\mathcal{H}(\beta) - \frac{\alpha}{2}\mathcal{H}(\beta/\alpha)$ v.s. the sparsity ratio β and the undersampling factor α. Here, $\frac{1}{2}\mathcal{H}(\beta) - \frac{\alpha}{2}\mathcal{H}(\beta/\alpha)$ characterizes the fundamental minimax capacity loss pert Hertz within a vanishing gap.

the capacity loss per Hertz reduces to

$$\frac{1}{2}\mathcal{H}\left(\beta\right)-\frac{1}{2}\alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right)=0,$$

meaning that there is effectively no capacity loss under Nyquist-rate sampling. This agrees with the fact that Nyquist-rate sampling is information preserving.

Part II

Algorithms: Structured Estimation from Randomly Undersampled Data

Chapter 4

Spectral Compressed Sensing via Structured Matrix Completion

This chapter investigates recovery of spectrally sparse signals from time-domain samples, where the underlying spectral spikes of the signals under study can assume any set of *continuous* values over a unit disk. Conventional harmonic-retrieval methods enforce constraints that are consistent with the physics of the problem (e.g. shift invariance) but ignore sparsity. In contrast, compressed sensing paradigms seek sparsity but ignore the shift invariance properties. Our goal is to develop an efficient recovery algorithm that can appropriately exploit spectral sparsity while enforcing physically-meaningful constraints.

4.1 Models and Algorithms

Assume that the signal of interest $x(\boldsymbol{t})$ can be modeled as a weighted superposition of K-dimensional sinusoids at r distinct frequencies $\boldsymbol{f}_i \in [0,1)^K$, $1 \leq i \leq r$, i.e.

$$x(\boldsymbol{t}) = \sum_{i=1}^{r} d_i e^{j2\pi\langle \boldsymbol{t}, \boldsymbol{f}_i \rangle}, \; \boldsymbol{t} \in \mathbb{Z}^K. \tag{4.1}$$

It is assumed throughout that the frequencies $\{\boldsymbol{f}_i\}_{i=1}^{r}$ are normalized relative to Nyquist rate and the time domain measurements are sampled at integer values. We denote by $\{d_i\}_{i=1}^r$ the complex amplitudes of the associated coefficients. For concreteness, our discussion is mainly devoted to a 2-D frequency model such that $K = 2$. This subsumes line spectral estimation as a special case, and indicates how to address multi-dimensional models.

4.1.1 2-D Frequency Model

Consider a data matrix $\boldsymbol{X} = [X_{k,l}]_{0\le k<n_1, 0\le l<n_2}$ of ambient dimension $n := n_1 n_2$, which is obtained by sampling the signal (4.1) over a uniform grid. From (4.1) each entry $X_{k,l}$ can be expressed as

$$X_{k,l} = x(k,l) = \sum_{i=1}^{r} d_i y_i^k z_i^l, \tag{4.2}$$

where for any i $(1 \le i \le r)$, $y_i := e^{j2\pi f_{1i}}$ and $z_i := e^{j2\pi f_{2i}}$ for some frequency pairs $\{\boldsymbol{f}_i = (f_{1i}, f_{2i}) \mid 1 \le i \le r\}$. We can then express $\boldsymbol{X}$ in a matrix form as follows

$$\boldsymbol{X} = \boldsymbol{Y}\boldsymbol{D}\boldsymbol{Z}^\top, \tag{4.3}$$

where the above matrices $\boldsymbol{Y}$, $\boldsymbol{Z}$ and $\boldsymbol{D}$ are defined as

$$\boldsymbol{Y} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ y_1 & y_2 & \cdots & y_r \\ \vdots & \vdots & \vdots & \vdots \\ y_1^{n_1-1} & y_2^{n_1-1} & \cdots & y_r^{n_1-1} \end{bmatrix}, \quad \boldsymbol{Z} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_r \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{n_2-1} & z_2^{n_2-1} & \cdots & z_r^{n_2-1} \end{bmatrix}, \tag{4.4}$$

and

$$\boldsymbol{D} := \operatorname{diag}\left[d_1, d_2, \cdots, d_r\right]. \tag{4.5}$$

The above form (4.3) is referred to as the Vandemonde decomposition of $\boldsymbol{X}$.

Suppose that we take measurements over a (random) location set Ω of size m, i.e. $X_{k,l}$ is observed if and only if $(k,l) \in \Omega$. We aim at recovering $\boldsymbol{X}$ from partial entries revealed over the index set Ω.

4.1.2 Matrix Enhancement

One might naturally attempt recovery by applying the low-rank MC algorithms [88], arguing that when r is small, perfect recovery of $\boldsymbol{X}$ is possible from partial measurements since $\boldsymbol{X}$ is low rank if $r \ll \min\{n_1, n_2\}$. Unfortunately, conventional MC algorithms (e.g. nuclear-norm minimization [88, 61]) require at least $\Theta\left(r \max\left(n_1, n_2\right) \log\left(n_1 n_2\right)\right)$ samples in order to achieve perfect recovery, which far exceeds the degrees of freedom (which is $\Theta\left(r\right)$) in this problem. This motivates us to seek other formulations that better capture the harmonic structure.

In this chapter, we adopt one effective enhanced form of $\boldsymbol{X}$ based on the following two-fold Hankel structure. The enhanced matrix $\boldsymbol{X}_{\mathrm{e}}$ with respect to $\boldsymbol{X}$ is defined as a $k_1 \times (n_1 - k_1 + 1)$ block Hankel matrix

$$\boldsymbol{X}_{\mathrm{e}} := \begin{bmatrix} \boldsymbol{X}_0 & \boldsymbol{X}_1 & \cdots & \boldsymbol{X}_{n_1-k_1} \\ \boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_{n_1-k_1+1} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{X}_{k_1-1} & \boldsymbol{X}_{k_1} & \cdots & \boldsymbol{X}_{n_1-1} \end{bmatrix}, \tag{4.6}$$

where k_1 ($1 \leq k_1 \leq n_1$) is called a pencil parameter. Each block is a $k_2 \times (n_2 - k_2 + 1)$ Hankel matrix defined such that for every ℓ ($0 \leq \ell < n_1$):

$$\boldsymbol{X}_{\ell} := \begin{bmatrix} X_{\ell,0} & X_{\ell,1} & \cdots & X_{\ell,n_2-k_2} \\ X_{\ell,1} & X_{\ell,2} & \cdots & X_{\ell,n_2-k_2+1} \\ \vdots & \vdots & \vdots & \vdots \\ X_{\ell,k_2-1} & X_{\ell,k_2} & \cdots & X_{\ell,n_2-1} \end{bmatrix}, \tag{4.7}$$

where $1 \leq k_2 \leq n_2$ is another pencil parameter. This enhanced form allows us to express each block as

$$\boldsymbol{X}_{\ell} = \boldsymbol{Z}_{\mathrm{L}} \boldsymbol{Y}_{\mathrm{d}}^{\ell} \boldsymbol{D} \boldsymbol{Z}_{\mathrm{R}}, \tag{4.8}$$

where $\boldsymbol{Z}_{\mathrm{L}}$, $\boldsymbol{Z}_{\mathrm{R}}$ and $\boldsymbol{Y}_{\mathrm{d}}$ are defined respectively as

$$\boldsymbol{Z}_{\mathrm{L}} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_r \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{k_2-1} & z_2^{k_2-1} & \cdots & z_r^{k_2-1} \end{bmatrix}, \quad \boldsymbol{Z}_{\mathrm{R}} := \begin{bmatrix} 1 & z_1 & \cdots & z_1^{n_2-k_2} \\ 1 & z_2 & \cdots & z_2^{n_2-k_2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & z_r & \cdots & z_r^{n_2-k_2} \end{bmatrix},$$

and

$$\boldsymbol{Y}_{\mathrm{d}} := \mathrm{diag}\left[y_1, y_2, \cdots, y_r\right].$$

Substituting (4.8) into (4.6) yields the following:

$$\boldsymbol{X}_{\mathrm{e}} = \underbrace{\begin{bmatrix} \boldsymbol{Z}_{\mathrm{L}} \\ \boldsymbol{Z}_{\mathrm{L}}\boldsymbol{Y}_{\mathrm{d}} \\ \vdots \\ \boldsymbol{Z}_{\mathrm{L}}\boldsymbol{Y}_{\mathrm{d}}^{k_1-1} \end{bmatrix}}_{\sqrt{k_1k_2}\boldsymbol{E}_{\mathrm{L}}} \boldsymbol{D} \underbrace{\left[\boldsymbol{Z}_{\mathrm{R}}, \quad \boldsymbol{Y}_{\mathrm{d}}\boldsymbol{Z}_{\mathrm{R}}, \quad \cdots, \quad \boldsymbol{Y}_{\mathrm{d}}^{n_1-k_1}\boldsymbol{Z}_{\mathrm{R}}\right]}_{\sqrt{(n_1-k_1+1)(n_2-k_2+1)}\boldsymbol{E}_{\mathrm{R}}}, \tag{4.9}$$

where $\boldsymbol{E}_{\mathrm{L}}$ and $\boldsymbol{E}_{\mathrm{R}}$ span the column and row space of $\boldsymbol{X}_{\mathrm{e}}$, respectively. This immediately implies that $\boldsymbol{X}_{\mathrm{e}}$ is *low-rank*, i.e. $\mathrm{rank}\left(\boldsymbol{X}_{\mathrm{e}}\right) \leq r$. This form is inspired by the matrix pencil form proposed in [35, 43]. One can extract all underlying frequencies of $\boldsymbol{X}$ using methods proposed in [43], as long as $\boldsymbol{X}$ can be faithfully recovered.

4.1.3 The EMaC Algorithm in the Absence of Noise

We then attempt recovery through the following *Enhancement Matrix Completion (EMaC)* algorithm:

$$\begin{aligned} \text{(EMaC)} \quad \underset{\boldsymbol{M}\in\mathbb{C}^{n_1\times n_2}}{\text{minimize}} \quad & \left\|\boldsymbol{M}_{\mathrm{e}}\right\|_* \\ \text{subject to} \quad & \mathcal{P}_{\Omega}\left(\boldsymbol{M}\right) = \mathcal{P}_{\Omega}\left(\boldsymbol{X}\right), \end{aligned} \tag{4.10}$$

where $\boldsymbol{M}_{\mathrm{e}}$ denotes the enhanced form of $\boldsymbol{M}$, $\left\|\cdot\right\|_*$ the nuclear norm (i.e. sum of singular values of a matrix), and $\mathcal{P}_{\Omega}(\boldsymbol{X})$ the orthogonal projection of $\boldsymbol{X}$ onto the subspace of matrices that vanish outside Ω. In other words, EMaC minimizes the nuclear norm of the enhanced form over all matrices compatible with the samples. This convex program can be tractably solved using off-the-shelf SDP solvers (see, e.g., [89]).

One natural question regarding this technique is that the performance of EMaC depends on the choices of the pencil parameters k_1 and k_2. In fact, if we define a quantity

$$c_{\mathrm{s}} := \max\left\{\frac{n_1 n_2}{k_1 k_2}, \frac{n_1 n_2}{(n_1 - k_1 + 1)(n_2 - k_2 + 1)}\right\} \tag{4.11}$$

that measures how close $\boldsymbol{X}_{\mathrm{e}}$ is to a square matrix, then it will be shown later that the required sample complexity for faithful recovery is an increasing function of c_{s}. In fact, both our theory and empirical experiments are in favor of a small c_{s}, as will be shown later.

4.1.4 Connection and Comparison to Prior Work

The multi-fold Hankel structure, which plays a central role in the EMaC algorithm, roots from the traditional spectral estimation technique named Matrix Enhancement Matrix Pencil (MEMP) [43] for multi-dimensional harmonic retrieval. The conventional MEMP algorithm assumes fully observed equi-spaced time domain samples for estimation, and requires prior knowledge on the model order. Cadzow's denoising method [90] also exploits the low-rank structure of the matrix pencil form for denoising line spectrum, but the method is non-convex and lacks performance guarantees.

When the frequencies of the signal indeed reside on a grid, CS algorithms based on ℓ_1 minimization [23, 1] can be used to recover the spectrally sparse signal from $\mathcal{O}(r \log n)$ random time-domain samples. These algorithms admit faithful recovery even when the samples are contaminated by bounded noise [40, 91] or sparse outliers [41]. When the inevitable *basis mismatch* issue [42] is present, several modifications for CS algorithms have been proposed to mitigate the effect of mismatch [92, 93] under random linear measurements, although theoretical guarantees are in general lacking.

Recently, Candès and Fernandez-Granda [94] proposed a total-variation norm minimization algorithm to super-resolve a sparse signal from frequency samples at the *low end* of the spectrum. This algorithm allows accurate super-resolution when the point sources are sufficiently separated, and is stable against noise [95]. Inspired by this approach, Tang et. al. [96, 97] then developed an atomic norm minimization algorithm for line spectral estimation from $\mathcal{O}(r \log r \log n)$ random time domain samples, which enables exact recovery when the frequencies are separated by at least $4/n$ with random amplitude phases. This approach is later extended to multi-dimensional frequencies [98, 99] and multiple measurement vectors [100]. However, these results are established under a random signal model, i.e. the complex signs of the frequency spikes were assumed to be i.i.d. and drawn from a uniform distribution. The robustness of the method against noise and outliers was not established. In contrast, our approach yields deterministic conditions for multi-dimensional frequency models that guarantee perfect recovery with noiseless samples and are provably robust against noise and sparse corruptions (see [73]).

More broadly, our algorithm is inspired by recent advances of Matrix Completion (MC) [88, 101, 102, 103], which aims at recovering a low-rank matrix from partial entries. It has been shown [61, 104] that exact recovery is possible via nuclear norm minimization, as soon as the sample complexity exceeds the order of the information theoretic limit. This line of algorithms is also robust against noise and outliers [105, 106], and allows exact recovery even in the presence of a constant portion of adversarially corrupted entries [107, 108, 109], which have found numerous applications in collaborative filtering [110], medical imaging [111, 112], etc. Nevertheless, the theoretical guarantees of these algorithms do not apply to the more structured observation models associated with the proposed multi-fold Hankel structure. Consequently, direct application of existing MC results delivers pessimistic sample complexity, which far exceeds the degrees of freedom underlying the signal.

4.1.5 Notations

Before continuing, we introduce a few notations that will be used throughout this chapter. Let the singular value decomposition (SVD) of $\boldsymbol{X}_{\mathrm{e}}$ be $\boldsymbol{X}_{\mathrm{e}} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^*$. Let

$$T := \left\{ \boldsymbol{U}\boldsymbol{M}^* + \tilde{\boldsymbol{M}}\boldsymbol{V}^* : \boldsymbol{M} \in \mathbb{C}^{(n_1-k_1+1)(n_2-k_1+1)\times r}, \tilde{\boldsymbol{M}} \in \mathbb{C}^{k_1k_2\times r} \right\} \tag{4.12}$$

be the tangent space with respect to $\boldsymbol{X}_{\mathrm{e}}$, and $T^{\perp}$ the orthogonal complement of T. Denote by $\mathcal{P}_U$ (resp. $\mathcal{P}_V$, $\mathcal{P}_T$) the orthogonal projection onto the column (resp. row, tangent) space of $\boldsymbol{X}_{\mathrm{e}}$, i.e. for any $\boldsymbol{M}$,

$$\mathcal{P}_U(\boldsymbol{M}) = \boldsymbol{U}\boldsymbol{U}^*\boldsymbol{M}, \quad \mathcal{P}_V(\boldsymbol{M}) = \boldsymbol{M}\boldsymbol{V}\boldsymbol{V}^*, \quad \text{and} \quad \mathcal{P}_T = \mathcal{P}_U + \mathcal{P}_V - \mathcal{P}_U\mathcal{P}_V.$$

We let $\mathcal{P}_{T^\perp} = \mathcal{I} - \mathcal{P}_T$ be the orthogonal complement of $\mathcal{P}_T$, where $\mathcal{I}$ denotes the identity operator. Denote by $\|\boldsymbol{M}\|$, $\|\boldsymbol{M}\|_{\mathrm{F}}$ and $\|\boldsymbol{M}\|_*$ the spectral norm (operator norm), Frobenius norm, and nuclear norm of $\boldsymbol{M}$, respectively. We denote by $\boldsymbol{e}_i$ the i^{th} standard basis vector.

Additionally, we denote by $\Omega_{\mathrm{e}}(k,l)$ the set of locations of the enhanced matrix $\boldsymbol{X}_{\mathrm{e}}$ containing copies of $X_{k,l}$. Due to the Hankel or multi-fold Hankel structures, one can easily verify the following: each location set $\Omega_{\mathrm{e}}(k,l)$ contains at most one index in any given row of the enhanced form, and at most one index in any given column. For each $(k,l) \in [n_1] \times [n_2]$, we use $\boldsymbol{A}_{(k,l)}$ to denote a basis matrix that extracts the average of all entries in $\Omega_{\mathrm{e}}(k,l)$. Specifically,

$$\left(\boldsymbol{A}_{(k,l)}\right)_{\alpha,\beta} := \begin{cases} \frac{1}{\sqrt{|\Omega_{\mathrm{e}}(k,l)|}}, & \text{if } (\alpha,\beta) \in \Omega_{\mathrm{e}}(k,l), \\ 0, & \text{else.} \end{cases} \tag{4.13}$$

We will use $\omega_{k,l} := |\Omega_{\mathrm{e}}(k,l)|$ throughout for brevity.

4.2 Main Results

This section delivers encouraging news: under mild incoherence conditions, EMaC enables faithful recovery of the possibly off-the-grid spectral spikes from an order-wise minimal number of random time-domain samples.

4.2.1 Incoherence Measure

In general, MC from a few entries is hopeless unless the underlying structure is sufficiently uncorrelated with the observation basis. This inspires us to introduce certain incoherence measures. To this end, we define the 2-D Dirichlet kernel as

$$\mathcal{D}(k_1,k_2,\boldsymbol{f}) := \frac{1}{k_1k_2}\left(\frac{1-e^{-j2\pi k_1 f_1}}{1-e^{-j2\pi f_1}}\right)\left(\frac{1-e^{-j2\pi k_2 f_2}}{1-e^{-j2\pi f_2}}\right), \tag{4.14}$$

where $\boldsymbol{f}=(f_1,f_2)\in[0,1)^2$. Fig. 4.1 (a) illustrates the amplitude of $\mathcal{D}(k_1,k_2,\boldsymbol{f})$ when $k_1=k_2=6$. The value of $|\mathcal{D}(k_1,k_2,\boldsymbol{f})|$ decays inverse proportionally with respect to the frequency $\boldsymbol{f}$. Set $\boldsymbol{G}_{\mathrm{L}}$ and $\boldsymbol{G}_{\mathrm{R}}$ to be two $r\times r$ Gram matrices such that their entries are specified respectively by

$$(\boldsymbol{G}_{\mathrm{L}})_{i,l} := \mathcal{D}(k_1,k_2,\boldsymbol{f}_i-\boldsymbol{f}_l), \quad (\boldsymbol{G}_{\mathrm{R}})_{i,l} := \mathcal{D}(n_1-k_1+1,n_2-k_2+1,\boldsymbol{f}_i-\boldsymbol{f}_l),$$

where the difference $\boldsymbol{f}_i-\boldsymbol{f}_l$ is understood as the wrap-around distance in the interval $[-1/2,1/2)^2$. Simple manipulation reveals that $\boldsymbol{G}_{\mathrm{L}}=\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}$ and $\boldsymbol{G}_{\mathrm{R}}=(\boldsymbol{E}_{\mathrm{R}}\boldsymbol{E}_{\mathrm{R}}^*)^\top$, where $\boldsymbol{E}_{\mathrm{L}}$ and $\boldsymbol{E}_{\mathrm{R}}$ are defined in (4.9).

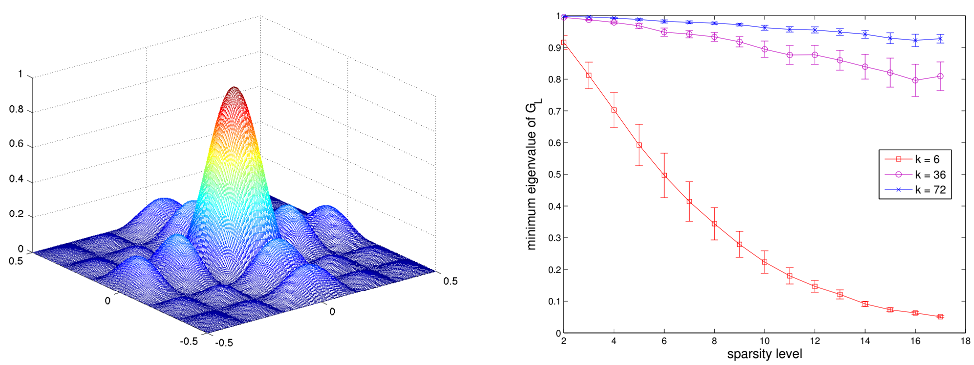
Our incoherence measure is then defined as follows.

Definition 4.1 (Incoherence). *A matrix $\boldsymbol{X}$ is said to obey the incoherence property with parameter μ_1 if*

$$\sigma_{\min}(\boldsymbol{G}_{\mathrm{L}})\geq 1/\mu_1 \quad \textit{and} \quad \sigma_{\min}(\boldsymbol{G}_{\mathrm{R}})\geq 1/\mu_1, \tag{4.15}$$

where $\sigma_{\min}(\boldsymbol{G}_{\mathrm{L}})$ (resp. $\sigma_{\min}(\boldsymbol{G}_{\mathrm{R}})$) is the least singular value of $\boldsymbol{G}_{\mathrm{L}}$ (resp. $\boldsymbol{G}_{\mathrm{R}}$).

The incoherence measure μ_1 depends only on the locations of the frequency spikes, irrespective of the amplitudes of their respective coefficients. The signal is said to satisfy the incoherence condition if μ_1 is bounded away from 0, which occurs when both $\boldsymbol{G}_{\mathrm{L}}$ and $\boldsymbol{G}_{\mathrm{R}}$ are well-conditioned. Our incoherence condition naturally requires certain separation among all frequency pairs, as when two frequency spikes are closely located, μ_1 gets undesirably large. As recently demonstrated in [113, Theorem 2], a separation of about $2/n$ for line spectrum is sufficient to guarantee incoherence. However, it is worth emphasizing that in contrast to [96], strict separation is not necessary for exhibiting incoherence, and our incoherence condition is applicable to a broader class of spectrally sparse signals.



(a) (b)

Figure 4.1: (a) The 2-D Dirichlet kernel when $k = k_1 = k_2 = 6$; (b) The empirical distribution of $\sigma_{\min}(\boldsymbol{G}_{\mathrm{L}})$ for various choices of k, when the spikes are randomly generated.

To give a flavor of the incoherence condition, we list two examples below. For ease of presentation, we assume in the following examples 2-D frequency models with $n_1 = n_2$. Note, however, that the asymmetric cases and general K-dimensional frequency models can be analyzed in the same manner.

- *Random frequency locations*: suppose that the r frequencies are generated uniformly at random, then the minimum pairwise separation can be crudely bounded by $\Theta\left(\frac{1}{r^2 \log n_1}\right)$. If $n_1 \gg r^2 \log n_1$, then one can see via [113, Theorem 2] that both $\sigma_{\min}\left(\boldsymbol{G}_{\mathrm{L}}\right)$ and $\sigma_{\min}\left(\boldsymbol{G}_{\mathrm{R}}\right)$ are bounded away from 0.

- *Small perturbation off the grid*: suppose that all frequencies are within a distance at most $\frac{1}{n_1 r^{1/4}}$ from some grid points $\left(\frac{l_1}{k_1}, \frac{l_2}{k_2}\right)$ $(0 \le l_1 < k_1, 0 \le l_2 < k_2)$. One can verify that $\forall i_1 \neq i_2$,

$$\max\left\{\frac{1}{k_1}\frac{1-\left(y_{i_1}^* y_{i_2}\right)^{k_1}}{1-y_{i_1}^* y_{i_2}}, \frac{1}{k_2}\frac{1-\left(z_{i_1}^* z_{i_2}\right)^{k_2}}{1-z_{i_1}^* z_{i_2}}\right\} < \frac{1}{2\sqrt{r}},$$

  and hence the magnitude of all off-diagonal entries of $\boldsymbol{G}_{\mathrm{L}}$ and $\boldsymbol{G}_{\mathrm{R}}$ are no larger than $1/(4r)$. This immediately suggests that $\sigma_{\min}(\boldsymbol{G}_{\mathrm{L}})$ and $\sigma_{\min}(\boldsymbol{G}_{\mathrm{R}})$ are lower bounded by $3/4$.

Note that the class of incoherent signals are far beyond the ones discussed above.

4.2.2 Theoretical Guarantees

Exact recovery is possible from a minimal number of noise-free samples, as asserted in the following theorem.

Theorem 4.1. *Let $\boldsymbol{X}$ be a data matrix of form (4.3), and let Ω be uniformly sampled with size m. Suppose that the incoherence property (4.15) holds and that all measurements are noiseless. Then there exists a universal constant $c_1 > 0$ such that $\boldsymbol{X}$ is the unique solution to EMaC with probability exceeding $1-(n_1 n_2)^{-2}$, provided that*

$$m > c_1 \mu_1 c_{\mathrm{s}} r \log^4(n_1 n_2). \tag{4.16}$$

Theorem 4.1 asserts that under some mild *deterministic* incoherence condition (i.e. μ_1 is bounded away from 0), EMaC admits prefect recovery as soon as the number of measurements exceeds $\mathcal{O}(r \log^4(n_1 n_2))$, which is only a logarithmic factor away from the underlying degrees of freedom (i.e. $\Theta(r)$). This demonstrates the near-optimality of EMaC. We note, however, that the polylog factor might be further refined via finer tuning of concentration inequalities.

It is worth emphasizing that while we assume random observation models, the data model is assumed to be deterministic. This differs drastically from [96] which relies on randomness in both the observation and data models. In particular, our theoretical performance guarantees rely solely on the frequency locations irrespective of the associated amplitudes. In contrast, the results in [96] require the phases of all frequency spikes to be i.i.d. drawn in a uniform manner in addition to a separation condition.

Additionally, our analysis framework can be straightforwardly adapted to the general Hankel (or Toeplitz) matrix completion problems that arise in system identification [114], computer vision [115], MRI [116], etc. We present below the performance guarantee for two-fold Hankel matrix completion.

Theorem 4.2 (**Hankel Matrix Completion**). *Consider a two-fold Hankel matrix* $\boldsymbol{X}_{\mathrm{e}}$ *of rank* r. *The bounds in Theorem 4.1 continue to hold if the incoherence* μ_1 *is defined as the smallest number that satisfies*

$$\max_{(k,l)\in[n_1]\times[n_2]} \left\{ \|\boldsymbol{U}^* \boldsymbol{A}_{(k,l)}\|_{\mathrm{F}}^2, \|\boldsymbol{A}_{(k,l)} \boldsymbol{V}\|_{\mathrm{F}}^2 \right\} \leq \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{4.17}$$

4.3 Numerical Experiments

In this section, we present numerical examples to evaluate the practical applicability of EMaC. Additionally, in order to handle larger scale data, we apply an extension of singular value thresholding (SVT) [117] to exploit the multi-fold Hankel structure.

4.3.1 Phase Transition in the Noiseless Setting

To evaluate the performance of EMaC, we conducted a series of numerical experiments to examine the phase transition for exact recovery. Let $n_1 = n_2$, and take $k_1 = k_2 = \lceil (n_1+1)/2 \rceil$. For each (r, m) pair, 100 Monte Carlo trials were conducted. We generated a spectrally sparse data matrix $\boldsymbol{X}$ by randomly generating r frequency spikes in $[0,1)\times[0,1)$, and sampled a subset Ω of size m entries uniformly at random. The EMaC algorithm was conducted using CVX with the interior-point solver SDPT3 [118]. Each trial is declared successful if the normalized mean squared error (NMSE) satisfies $\|\hat{\boldsymbol{X}} - \boldsymbol{X}\|_{\mathrm{F}} / \|\boldsymbol{X}\|_{\mathrm{F}} \leq 10^{-3}$, where $\hat{\boldsymbol{X}}$ denotes the estimate returned by EMaC. The empirical success rate is calculated by averaging over 100 Monte Carlo trials.

Fig. 4.2 illustrates the results of these Monte Carlo experiments when the dimensions of $\boldsymbol{X}$ are 11×11 and 15×15. The horizontal axis corresponds to the number m of samples revealed to the algorithm, while the vertical axis corresponds to the spectral sparsity level r. The empirical success rate is reflected by the color of each cell. It can be seen from the plot that the number of samples m grows approximately linearly with respect to the spectral sparsity r, and that the slopes of the phase transition lines for two cases are approximately the same. These observations are in line with our theoretical guarantee in Theorem 4.1. These phase transition diagrams justify the practical applicability of our algorithm in the noiseless setting.

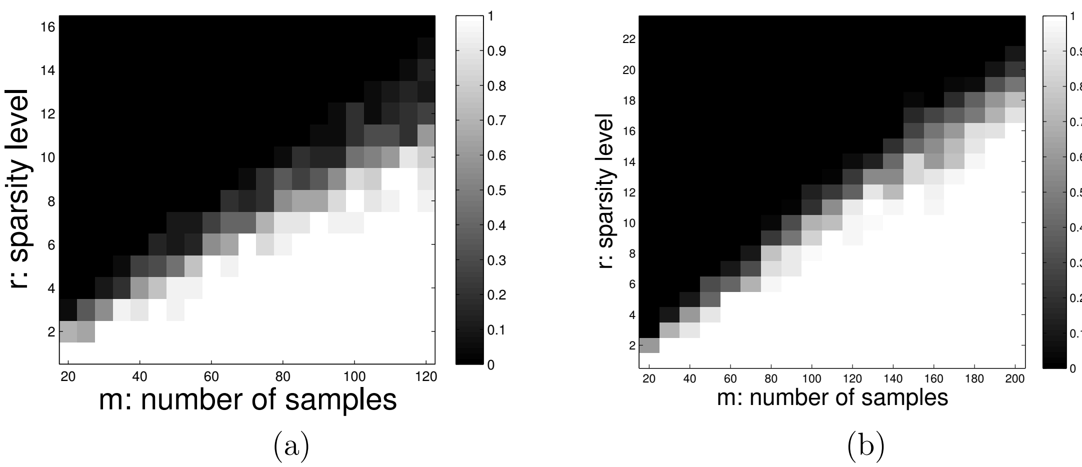


Figure 4.2: Phase transition plots where frequency locations are randomly generated. The plot (a) concerns the case where $n_1 = n_2 = 11$, whereas the plot (b) corresponds to the situation where $n_1 = n_2 = 15$. The empirical success rates are calculated by averaging over 100 Monte Carlo trials.

4.3.2 Singular Value Thresholding for EMaC

The above Monte Carlo experiments were conducted using the advanced SDP solver SDPT3. This solver and many other popular ones (e.g. SeDuMi) are based on second-order methods like interior point methods, which are typically limited to small-scale problems due to expensive computation of Hessian matrices. In fact, SDPT3 fails to handle an $n \times n$ data matrix when n exceeds 19, which corresponds to a 100×100 enhanced matrix.

One alternative for large-scale data is the first-order algorithms tailored to MC problems, e.g. the singular value thresholding (SVT) algorithm [117]. We propose to solve the program via a modified SVT algorithm to exploit the Hankel structure. Details can be found in [73].

Fig. 4.3 illustrates the performance of SVT. We generated a true 101×101 data matrix $\boldsymbol{X}$ through a superposition of 30 random complex sinusoids, and revealed 5.8% of the total entries (i.e. $m = 600$) uniformly at random. The noise was i.i.d. Gaussian giving a signal-to-noise amplitude ratio of 10. As illustrated in Fig. 4.3, the reconstruction error was $\|\hat{\boldsymbol{X}} - \boldsymbol{X}\|_{\mathrm{F}} / \|\boldsymbol{X}\|_{\mathrm{F}} = 0.1098$, validating the stability of our algorithm in the presence of noise.

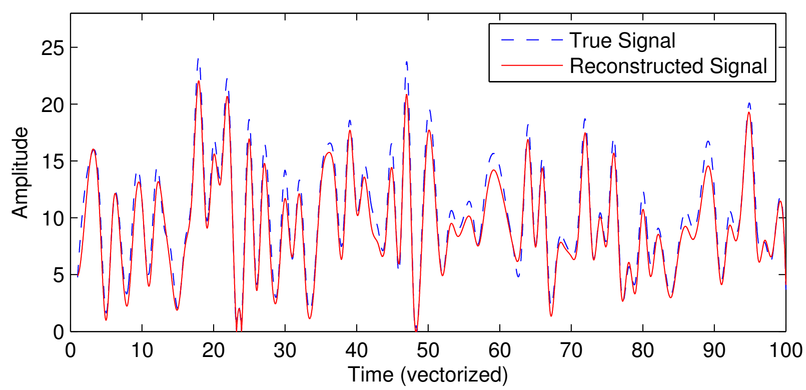


Figure 4.3: The performance of SVT for a 101×101 data matrix that contains 30 random frequency spikes. 5.8% of all entries ($m = 600$) are observed with signal-to-noise amplitude ratio 10. The reconstructed data against the true data for the first 100 time instances are plotted.

4.4 Proof Outline of Theorems 4.1 and 4.2

The proposed EMaC algorithm enjoys a similar spirit as the well-known MC algorithms [88, 61], except that we impose (multi-fold) Hankel structures on the matrices.

While [61, Theorem 3] has presented a general sufficient condition for exact recovery with the MC algorithm, the basis in our case does not exhibit desired coherence properties as required therein, and hence these results cannot deliver informative estimates when applied to our problem. We also note that the analyses adopted in [88, 61] rely on a desired joint incoherence property on $\boldsymbol{U}\boldsymbol{V}^*$, which has been shown to be unnecessary [104].

For concreteness, the analyses herein is concerned with Theorem 4.1, since proving Theorem 4.1 is slightly more involved than proving Theorem 4.2. We note, however, that our analysis already entails all reasoning required for establishing Theorem 4.2.

4.4.1 Dual Certification

Denote by $\mathcal{A}_{(k,l)}(\boldsymbol{M})$ the projection of $\boldsymbol{M}$ onto the subspace spanned by $\boldsymbol{A}_{(k,l)}$, and define the projection operator onto the space spanned by all $\boldsymbol{A}_{(k,l)}$ and its orthogonal complement as

$$\mathcal{A} := \sum_{(k,l)\in[n_1]\times[n_2]} \mathcal{A}_{(k,l)}, \quad \text{and} \quad \mathcal{A}^{\perp} = \mathcal{I} - \mathcal{A}. \tag{4.18}$$

There are two common ways to describe the randomness of Ω: one corresponds to sampling *without* replacement, and another concerns sampling *with* replacement (i.e. Ω contains m i.i.d. generated indices). As discussed in [61, Section II.A], while both situations result in the same order bounds, the latter situation admits simpler analysis due to independence. Therefore, we will assume that Ω is a multi-set (possibly with repeated elements). Let $\boldsymbol{a}_i$'s be independently and uniformly drawn from $[n_1] \times [n_2]$, and we define the associated operators as $\mathcal{A}_{\Omega} := \sum_{i=1}^{m} \mathcal{A}_{\boldsymbol{a}_i}$. We also define another projection operator $\mathcal{A}'_{\Omega}$ similar to $\mathcal{A}_{\Omega}$, but with the sum extending only over *distinct* samples. Its complement operator is defined as $\mathcal{A}'_{\Omega^{\perp}} := \mathcal{A} - \mathcal{A}'_{\Omega}$. With these definitions, EMaC can be rewritten as the following general MC problem

$$\begin{aligned} \underset{\boldsymbol{M}}{\text{minimize}} \quad & \|\boldsymbol{M}\|_* \\ \text{subject to} \quad & \mathcal{A}'_{\Omega}(\boldsymbol{M}) = \mathcal{A}'_{\Omega}(\boldsymbol{X}_{\mathrm{e}}), \quad \mathcal{A}^{\perp}(\boldsymbol{M}) = \mathcal{A}^{\perp}(\boldsymbol{X}_{\mathrm{e}}) = 0. \end{aligned} \tag{4.19}$$

To prove exact recovery of convex optimization, it suffices to produce an appropriate dual certificate, as stated in the following lemma.

Lemma 4.1. *Consider a multi-set Ω that contains m random indices. Suppose that the sampling operator $\mathcal{A}_\Omega$ obeys*

$$\left\| \mathcal{P}_T \mathcal{A} \mathcal{P}_T - \frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{A}_\Omega \mathcal{P}_T \right\| \leq 0.5. \tag{4.20}$$

If there exists a matrix $\boldsymbol{W}$ satisfying

$$\mathcal{A}'_{\Omega^\perp} \left(\boldsymbol{W} \right) = 0, \tag{4.21}$$

$$\left\| \mathcal{P}_T \left(\boldsymbol{W} - \boldsymbol{U}\boldsymbol{V}^* \right) \right\|_{\mathrm{F}} \leq 0.5 n_1^{-2} n_2^{-2}, \tag{4.22}$$

$$\left\| \mathcal{P}_{T^\perp} \left(\boldsymbol{W} \right) \right\| \leq 0.5, \tag{4.23}$$

then $\boldsymbol{X}$ is the unique minimizer of EMaC.

Condition (4.20) will be analyzed in Section 4.4.2, while a dual certificate $\boldsymbol{W}$ will be constructed in Section 4.4.3. The validity of $\boldsymbol{W}$ as a dual certificate will be established in Sections 4.4.3 - 4.4.5. These are the focus of the remainder of this chapter .

4.4.2 Concentration of $\mathcal{P}_T \mathcal{A} \mathcal{P}_T$

Lemma 4.1 requires that $\mathcal{A}_\Omega$ be sufficiently incoherent w.r.t. the tangent space T. The following lemma quantifies the projection of each $\boldsymbol{A}_{(k,l)}$ onto the subspace T.

Lemma 4.2. *Under the hypothesis* (4.15)*, one has*

$$\left\| \boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A}_{(k,l)} \right\|_{\mathrm{F}}^2 \leq \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}, \quad \left\| \boldsymbol{A}_{(k,l)} \boldsymbol{V}\boldsymbol{V}^* \right\|_{\mathrm{F}}^2 \leq \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}, \tag{4.24}$$

for all $(k,l) \in [n_1] \times [n_2]$. For any $\boldsymbol{a}, \boldsymbol{b} \in [n_1] \times [n_2]$, one has

$$\left| \langle \boldsymbol{A}_{\boldsymbol{b}}, \mathcal{P}_T \boldsymbol{A}_{\boldsymbol{a}} \rangle \right| \leq \sqrt{\frac{\omega_{\boldsymbol{b}}}{\omega_{\boldsymbol{a}}}} \frac{3\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{4.25}$$

Recognizing that (4.24) is the same as (4.17), the following proof also establishes Theorem 4.2. Note that Lemma 4.2 immediately leads to

$$\left\|\mathcal{P}_T\left(\boldsymbol{A}_{(k,l)}\right)\right\|_{\mathrm{F}}^2 \leq \left\|\mathcal{P}_U\left(\boldsymbol{A}_{(k,l)}\right)\right\|_{\mathrm{F}}^2 + \left\|\mathcal{P}_V\left(\boldsymbol{A}_{(k,l)}\right)\right\|_{\mathrm{F}}^2 \leq \frac{2\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{4.26}$$

Once (4.26) holds, the fluctuation of $\mathcal{P}_T\mathcal{A}_\Omega\mathcal{P}_T$ can be well controlled, as stated in the following lemma. This justifies Condition (4.20) as required by Lemma 4.1.

Lemma 4.3. *Suppose that* (4.26) *holds. Then for any small constant* $0 < \epsilon \leq \frac{1}{2}$,

$$\left\|\frac{n_1 n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega\mathcal{P}_T - \mathcal{P}_T\mathcal{A}\mathcal{P}_T\right\| \leq \epsilon \tag{4.27}$$

holds with probability exceeding $1 - (n_1 n_2)^{-4}$, *provided that* $m > c_1 \mu_1 c_{\mathrm{s}} r \log(n_1 n_2)$ *for some universal constant* $c_1 > 0$.

4.4.3 Construction of Dual Certificates

Now we are in a position to construct the dual certificate via the golfing scheme, which is the way proposed by Gross [61] for construction of slightly relaxed dual certificates, as detailed below. Suppose that we generate j_0 independent random location multi-sets Ω_i $(1 \leq i \leq j_0)$, each containing $\frac{m}{j_0}$ i.i.d. samples. This way the distribution of Ω is the same as $\Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_{j_0}$. Note that Ω_i's correspond to sampling *with* replacement. Let

$$\rho := m/(n_1 n_2) \quad \text{and} \quad q := \rho/j_0 \tag{4.28}$$

represent the undersampling factors of Ω and Ω_i, respectively.

Consider a small constant $\epsilon < \frac{1}{e}$, and pick $j_0 := 3\log_{\frac{1}{\epsilon}} n_1 n_2$. The construction of the dual matrix $\boldsymbol{W}$ then proceeds as follows:

Construction of a dual certificate $\boldsymbol{W}$ via the golfing scheme.
1. Set $\boldsymbol{F}_0 = \boldsymbol{U}\boldsymbol{V}^*$, and $j_0 := 5\log_{\frac{1}{\epsilon}}(n_1 n_2)$. 2. For all i $(1 \leq i \leq j_0)$, let $\boldsymbol{F}_i = \mathcal{P}_T\left(\mathcal{A} - \frac{1}{q}\mathcal{A}_{\Omega_i}\right)\mathcal{P}_T\left(\boldsymbol{F}_{i-1}\right)$. 3. Set $\boldsymbol{W} := \sum_{j=1}^{j_0}\left(\frac{1}{q}\mathcal{A}_{\Omega_i} + \mathcal{A}^{\perp}\right)(\boldsymbol{F}_{i-1})$.

We will establish that $\boldsymbol{W}$ is a valid dual certificate by showing that $\boldsymbol{W}$ satisfies the conditions stated in Lemma 4.1, which we proceed as follows.

First, by construction, all summands $\left(\frac{1}{q}\mathcal{A}_{\Omega_i}+\mathcal{A}^{\perp}\right)(\boldsymbol{F}_{i-1})$ lie within the subspace supported on Ω or $\mathcal{A}^{\perp}$. This validates that $\mathcal{A}'_{\Omega^{\perp}}(\boldsymbol{W})=0$, as required in (4.21).

Secondly, the recursive construction procedure of $\boldsymbol{F}_i$ allows us to write

$$\begin{aligned}
-\mathcal{P}_T(\boldsymbol{W}-\boldsymbol{F}_0) &= \mathcal{P}_T(\boldsymbol{F}_0)-\sum_{j=1}^{j_0}\mathcal{P}_T\left(\frac{1}{q}\mathcal{A}_{\Omega_i}+\mathcal{A}^{\perp}\right)(\boldsymbol{F}_{i-1})\\
&= \mathcal{P}_T(\boldsymbol{F}_0)-\mathcal{P}_T\left(\frac{1}{q}\mathcal{A}_{\Omega_i}+\mathcal{A}^{\perp}\right)\mathcal{P}_T(\boldsymbol{F}_0)-\sum_{j=2}^{j_0}\mathcal{P}_T\left(\frac{1}{q}\mathcal{A}_{\Omega_i}+\mathcal{A}^{\perp}\right)(\boldsymbol{F}_{i-1})\\
&= \mathcal{P}_T\left(\mathcal{A}-\frac{1}{q}\mathcal{A}_{\Omega_i}\right)\mathcal{P}_T(\boldsymbol{F}_0)-\sum_{j=2}^{j_0}\mathcal{P}_T\left(\frac{1}{q}\mathcal{A}_{\Omega_i}+\mathcal{A}^{\perp}\right)\boldsymbol{F}_{i-1}\\
&= \mathcal{P}_T(\boldsymbol{F}_1)-\sum_{j=2}^{j_0}\mathcal{P}_T\left(\frac{1}{q}\mathcal{A}_{\Omega_i}+\mathcal{A}^{\perp}\right)(\boldsymbol{F}_{i-1})=\cdots=\mathcal{P}_T(\boldsymbol{F}_{j_0}). \qquad (4.29)
\end{aligned}$$

Lemma 4.3 asserts the following: if $qn_1n_2 \geq c_1\mu_1c_{\mathrm{s}}r\log(n_1n_2)$ or, equivalently, $m \geq \tilde{c}_1\mu_1c_{\mathrm{s}}r\log^2(n_1n_2)$ for some constant $\tilde{c}_1>0$, then with overwhelming probability,

$$\left\|\mathcal{P}_T-\mathcal{P}_T\left(\frac{1}{q}\mathcal{A}_{\Omega_i}+\mathcal{A}^{\perp}\right)\mathcal{P}_T\right\|=\left\|\mathcal{P}_T\mathcal{A}\mathcal{P}_T-\frac{1}{q}\mathcal{P}_T\mathcal{A}_{\Omega_i}\mathcal{P}_T\right\|\leq\epsilon<\frac{1}{2}.$$

This allows us to bound $\|\mathcal{P}_T(\boldsymbol{F}_i)\|_{\mathrm{F}}$ as

$$\|\mathcal{P}_T(\boldsymbol{F}_i)\|_{\mathrm{F}}\leq\epsilon^i\|\mathcal{P}_T(\boldsymbol{F}_0)\|_{\mathrm{F}}\leq\epsilon^i\|\boldsymbol{U}\boldsymbol{V}^*\|_{\mathrm{F}}=\epsilon^i\sqrt{r},$$

which together with (4.29) gives

$$\|\mathcal{P}_T(\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^*)\|_{\mathrm{F}}=\|\mathcal{P}_T(\boldsymbol{W}-\boldsymbol{F}_0)\|_{\mathrm{F}}=\|\mathcal{P}_T(\boldsymbol{F}_{j_0})\|_{\mathrm{F}}\leq\epsilon^{j_0}\sqrt{r}<(2n_1^2n_2^2)^{-1}, \qquad (4.30)$$

as required in Condition (4.22).

Finally, it remains to be shown that $\|\mathcal{P}_{T^{\perp}}(\boldsymbol{W})\|\leq\frac{1}{2}$, which we will establish in the next two subsections. In particular, we first introduce two key metrics and characterize their relationships in Section 4.4.4. These metrics are crucial in bounding $\|\mathcal{P}_{T^{\perp}}(\boldsymbol{W})\|$, which will be the focus of Section 4.4.5.

4.4.4 Two Metrics and Associated Bounds

In this subsection, we introduce the following two norms

$$\|\boldsymbol{M}\|_{\mathcal{A},\infty} := \max_{(k,l)\in[n_1]\times[n_2]} \left|\frac{\langle \boldsymbol{A}_{(k,l)}, \boldsymbol{M}\rangle}{\sqrt{\omega_{k,l}}}\right|, \tag{4.31}$$

$$\|\boldsymbol{M}\|_{\mathcal{A},2} := \sqrt{\sum_{(k,l)\in[n_1]\times[n_2]} \frac{\left|\langle \boldsymbol{A}_{(k,l)}, \boldsymbol{M}\rangle\right|^2}{\omega_{k,l}}}. \tag{4.32}$$

Based on these two metrics, we can derive several technical lemmas which, taken collectively, allow us to control $\|\mathcal{P}_{T^{\perp}}(\boldsymbol{W})\|$. Specifically, these lemmas characterize the mutual dependence of three norms $\|\cdot\|$, $\|\cdot\|_{\mathcal{A},2}$ and $\|\cdot\|_{\mathcal{A},\infty}$.

Lemma 4.4. *For any given matrix $\boldsymbol{M}$, there exists some numerical constant $c_2 > 0$ such that with probability at least $1 - (n_1 n_2)^{-10}$, one has*

$$\left\|\left(\frac{n_1 n_2}{m}\mathcal{A}_{\Omega} - \mathcal{A}\right)(\boldsymbol{M})\right\| \le c_2 \left(\sqrt{\frac{n_1 n_2 \log(n_1 n_2)}{m}}\|\boldsymbol{M}\|_{\mathcal{A},2} + \frac{n_1 n_2 \log(n_1 n_2)}{m}\|\boldsymbol{M}\|_{\mathcal{A},\infty}\right). \tag{4.33}$$

Lemma 4.5. *Assume that there exists a quantity μ_5 such that*

$$\omega_{\alpha,\beta}\left\|\mathcal{P}_T\left(\boldsymbol{A}_{(\alpha,\beta)}\right)\right\|_{\mathcal{A},2}^2 \le \frac{\mu_5 r}{n_1 n_2},\ (\alpha,\beta)\in[n_1]\times[n_2]. \tag{4.34}$$

Then for any given matrix $\boldsymbol{M}$, with probability exceeding $1 - (n_1 n_2)^{-10}$, one has

$$\left\|\left(\frac{n_1 n_2}{m}\mathcal{P}_T\mathcal{A}_{\Omega} - \mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},2} \le c_3\sqrt{\frac{\mu_5 r \log(n_1 n_2)}{m}}\cdot\left(\|\boldsymbol{M}\|_{\mathcal{A},2} + \sqrt{\frac{n_1 n_2 \log(n_1 n_2)}{m}}\|\boldsymbol{M}\|_{\mathcal{A},\infty}\right) \tag{4.35}$$

for some absolute constant $c_3 > 0$.

Lemma 4.6. *For any given matrix $\boldsymbol{M} \in T$, there is some absolute constant $c_4 > 0$ such that with probability exceeding* $1-(n_1n_2)^{-10}$,

$$\begin{aligned}&\left\|\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega-\mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},\infty}\\&\quad\leq c_4\left(\sqrt{\frac{\mu_1c_{\mathrm{s}}r\log(n_1n_2)}{m}}\cdot\sqrt{\frac{\mu_1c_{\mathrm{s}}r}{n_1n_2}}\left\|\boldsymbol{M}\right\|_{\mathcal{A},2}+\frac{\mu_1c_{\mathrm{s}}r\log(n_1n_2)}{m}\left\|\boldsymbol{M}\right\|_{\mathcal{A},\infty}\right).\end{aligned}\tag{4.36}$$

Lemma 4.5 combined with Lemma 4.6 gives rise to the following inequality. Consider any given matrix $\boldsymbol{M} \in T$. Applying the bounds (4.35) and (4.36) suggests

$$\begin{aligned}&\left\|\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega-\mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},2}+\sqrt{\frac{n_1n_2\log(n_1n_2)}{m}}\left\|\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega-\mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},\infty}\\&\quad\leq c_5\left(\sqrt{\frac{\mu_5r\log(n_1n_2)}{m}}+\frac{\mu_1c_{\mathrm{s}}r\log(n_1n_2)}{m}\right)\cdot\left\{\left\|\boldsymbol{M}\right\|_{\mathcal{A},2}+\sqrt{\frac{n_1n_2\log(n_1n_2)}{m}}\left\|\boldsymbol{M}\right\|_{\mathcal{A},\infty}\right\},\end{aligned}\tag{4.37}$$

with probability exceeding $1-(n_1n_2)^{-10}$, where $c_5 = \max\{c_3, c_4\}$. This holds under the hypothesis (4.34).

4.4.5 An Upper Bound on $\|\mathcal{P}_{T^\perp}(\boldsymbol{W})\|$

Now we are ready to show how we may combine the above lemmas to develop an upper bound on $\|\mathcal{P}_{T^\perp}(\boldsymbol{W})\|$. By construction,

$$\|\mathcal{P}_{T^\perp}(\boldsymbol{W})\|\leq\sum_{l=1}^{j_0}\left\|\mathcal{P}_{T^\perp}\left(\frac{1}{q}\mathcal{A}_{\Omega_l}+\mathcal{A}^\perp\right)\mathcal{P}_T(\boldsymbol{F}_{l-1})\right\|.$$

Each summand can be bounded above as follows

$$
\begin{aligned}
\left\|\mathcal{P}_{T^{\perp}}\left(\frac{1}{q}\mathcal{A}_{\Omega_l}+\mathcal{A}^{\perp}\right)\mathcal{P}_T\left(\boldsymbol{F}_{l-1}\right)\right\| &= \left\|\mathcal{P}_{T^{\perp}}\left(\frac{1}{q}\mathcal{A}_{\Omega_l}-\mathcal{A}\right)\mathcal{P}_T\left(\boldsymbol{F}_{l-1}\right)\right\| \leq \left\|\left(\frac{1}{q}\mathcal{A}_{\Omega_l}-\mathcal{A}\right)\left(\boldsymbol{F}_{l-1}\right)\right\| \\
&\leq c_2\left(\sqrt{\frac{\log\left(n_1n_2\right)}{q}}\left\|\boldsymbol{F}_{l-1}\right\|_{\mathcal{A},2}+\frac{\log\left(n_1n_2\right)}{q}\left\|\boldsymbol{F}_{l-1}\right\|_{\mathcal{A},\infty}\right) \qquad (4.38)\\
&\leq c_2c_5\left(\sqrt{\frac{\mu_5 r\log\left(n_1n_2\right)}{qn_1n_2}}+\frac{\mu_1c_{\mathrm{s}}r\log\left(n_1n_2\right)}{qn_1n_2}\right)\left\{\sqrt{\frac{\log\left(n_1n_2\right)}{q}}\left\|\boldsymbol{F}_{l-2}\right\|_{\mathcal{A},2}+\frac{\log\left(n_1n_2\right)}{q}\left\|\boldsymbol{F}_{l-2}\right\|_{\mathcal{A},\infty}\right\} \qquad (4.39)\\
&\leq \left(\frac{1}{2}\right)^{l-1}\left(\sqrt{\frac{\log\left(n_1n_2\right)}{q}}\cdot\left\|\boldsymbol{F}_0\right\|_{\mathcal{A},2}+\frac{\log\left(n_1n_2\right)}{q}\left\|\boldsymbol{F}_0\right\|_{\mathcal{A},\infty}\right), \qquad (4.40)
\end{aligned}
$$

where (4.38) follows from Lemma 4.4 together with the fact that $\boldsymbol{F}_i \in T$, and (4.39) is a consequence of (4.37). The last inequality holds under the hypothesis that $qn_1n_2 \gg \max\left\{\mu_1c_{\mathrm{s}},\mu_5\right\} r\log\left(n_1n_2\right)$ or, equivalently, $m \gg \max\left\{\mu_1c_{\mathrm{s}},\mu_5\right\} r\log^2\left(n_1n_2\right)$.

It remains to control $\left\|\boldsymbol{U}\boldsymbol{V}^*\right\|_{\mathcal{A},\infty}$ and $\left\|\boldsymbol{U}\boldsymbol{V}^*\right\|_{\mathcal{A},2}$ as follows.

Lemma 4.7. *With the incoherence measure μ_1, for any $(\alpha,\beta)\in[n_1]\times[n_2]$, one has*

$$\left\|\boldsymbol{U}\boldsymbol{V}^*\right\|_{\mathcal{A},\infty} \leq \frac{\mu_1c_{\mathrm{s}}r}{n_1n_2}, \qquad (4.41)$$

$$\left\|\boldsymbol{U}\boldsymbol{V}^*\right\|_{\mathcal{A},2}^2 \leq \frac{\mu_1c_{\mathrm{s}}r\log^2\left(n_1n_2\right)}{n_1n_2}, \qquad (4.42)$$

$$\left\|\mathcal{P}_T\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\right\|_{\mathcal{A},2}^2 \leq \frac{c_6\mu_1c_{\mathrm{s}}\log^2\left(n_1n_2\right)r}{n_1n_2} \qquad (4.43)$$

for some numerical constant $c_6>0$.

In particular, the bound (4.43) translates into $\mu_5 \leq c_6\mu_1c_{\mathrm{s}}\log^2\left(n_1n_2\right)$. Substituting (4.41) and (4.42) into (4.40) gives

$$\left\|\mathcal{P}_{T^{\perp}}\left(\frac{n_1n_2}{m}\mathcal{A}_{\Omega_l}+\mathcal{A}^{\perp}\right)\mathcal{P}_T\left(\boldsymbol{F}_{l-1}\right)\right\| \leq \left(\frac{1}{2}\right)^{l-1}\left(\sqrt{\frac{\mu_1c_{\mathrm{s}}r\log^2\left(n_1n_2\right)}{qn_1n_2}}+\frac{\mu_1c_{\mathrm{s}}r\log\left(n_1n_2\right)}{qn_1n_2}\right) \ll \left(\frac{1}{2}\right)^{l+1},$$

as soon as $m > c_7 \max\left\{\mu_1 c_{\mathrm{s}} \log^2\left(n_1 n_2\right), \mu_5 \log^2\left(n_1 n_2\right)\right\} r$ or $m > \tilde{c}_7 \mu_1 c_{\mathrm{s}} \log^4\left(n_1 n_2\right)$ for some sufficiently large constants $c_7, \tilde{c}_7 > 0$, indicating that

$$\left\|\mathcal{P}_{T^{\perp}}\left(\boldsymbol{W}\right)\right\| \leq \sum_{l=1}^{j_0}\left\|\mathcal{P}_{T^{\perp}}\left(\frac{1}{q}\mathcal{A}_{\Omega_l} + \mathcal{A}^{\perp}\right)\mathcal{P}_T\left(\boldsymbol{F}_{l-1}\right)\right\| \leq \frac{1}{2} \cdot \sum_{l=1}^{\infty}\left(\frac{1}{2}\right)^l \leq \frac{1}{2}$$

as required. Hence, we have successfully verified that, with high probability, $\boldsymbol{W}$ is a valid dual certificate, and hence by Lemma 4.1 the solution to EMaC is exact.

Chapter 5

Covariance Estimation from Quadratic Measurements

This chapter investigates covariance recovery from random quadratic measurements, an effective data acquisition method that requires minimal storage and admits universally faithful covariance recovery under a variety of structural assumptions.

5.1 Models and Notations

In general, recovering the covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ from $m < n(n+1)/2$ measurements is ill-posed, unless the sampling mechanism can effectively exploit the low-dimensional covariance structure. Random sampling often preserves the information from minimal observations, and allows robust recovery from noisy measurements.

In this chapter, we restrict our attention to the following random sampling model. We assume that the sensing vectors are composed of i.i.d. *sub-Gaussian* entries. In particular, we assume $\boldsymbol{a}_i$'s ($1 \leq i \leq m$) are i.i.d. copies of $\boldsymbol{z} = [z_1, \cdots, z_n]^\top$, where each z_i is i.i.d. drawn from a distribution with the following properties

$$\mathbb{E}[z_i] = 0, \quad \mathbb{E}[z_i^2] = 1, \quad \text{and} \quad \mu_4 := \mathbb{E}\left[z_i^4\right] > 1. \tag{5.1}$$

We assume that the noise $\boldsymbol{\eta} := [\eta_1, \cdots, \eta_m]^\top$ is bounded in either ℓ_1 norm or ℓ_2 norm as specified later. For notational simplicity, let $\boldsymbol{A}_i := \boldsymbol{a}_i \boldsymbol{a}_i^\top$ represent the equivalent sensing matrix. The quadratic measurements $\boldsymbol{y} := [y_1, \cdots, y_m]^\top$ considered herein are given by

$$y_i = \boldsymbol{a}_i^\top \boldsymbol{\Sigma} \boldsymbol{a}_i + \eta_i = \langle \boldsymbol{A}_i, \boldsymbol{\Sigma} \rangle + \eta_i, \quad i = 1, \cdots, m. \tag{5.2}$$

If we define the linear operator $\mathcal{A}(\boldsymbol{M}) : \mathbb{R}^{n\times n} \mapsto \mathbb{R}^m$ that maps a matrix $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ to $\{\langle \boldsymbol{M}, \boldsymbol{A}_i \rangle\}_{i=1}^m$, then the obtained measurements can be succinctly expressed as

$$\boldsymbol{y} = \mathcal{A}(\boldsymbol{\Sigma}) + \boldsymbol{\eta}. \tag{5.3}$$

Finally, we provide a brief summary of useful notations. We denote by $\|\boldsymbol{X}\|$, $\|\boldsymbol{X}\|_{\mathrm{F}}$, $\|\boldsymbol{X}\|_*$, $\|\boldsymbol{X}\|_1$ and $\|\boldsymbol{X}\|_0$ the spectral norm, the Frobenius norm, the nuclear norm, the elementwise ℓ_1 norm, and the support size of of $\boldsymbol{X}$, respectively. The matrix inner product is defined as $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle = \mathrm{tr}(\boldsymbol{X}^\top \boldsymbol{Y})$. We denote by $\mathcal{T}$ the orthogonal projection operator onto Toeplitz matrices, and $\mathcal{T}^\perp$ its orthogonal complement.

5.2 Algorithms and Main Results

5.2.1 Recovery of Low-Rank Covariance Matrices

Suppose that $\boldsymbol{\Sigma}$ is low-rank. A natural heuristic is to perform rank minimization to encourage the low-rank structure

$$\hat{\boldsymbol{\Sigma}} = \arg\min_{\boldsymbol{M}} \mathrm{rank}(\boldsymbol{M}) \quad \text{s.t.} \quad \boldsymbol{M} \succeq 0, \quad \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{M})\|_1 \le \epsilon_1, \tag{5.4}$$

where ϵ_1 is an upper bound on $\|\boldsymbol{\eta}\|_1$ and assumed known *a priori*. However, the rank minimization problem is in general NP-hard. Therefore, we replace it with trace minimization over all matrices compatible with the measurements

$$\hat{\boldsymbol{\Sigma}} = \arg\min_{\boldsymbol{M}} \mathrm{tr}(\boldsymbol{M}) \quad \text{s.t.} \quad \boldsymbol{M} \succeq 0, \quad \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{M})\|_1 \le \epsilon_1. \tag{5.5}$$

Since $\mathbf{\Sigma}$ is PSD, the trace norm forms a convex surrogate for the rank function, which has proved successful in matrix completion and phase retrieval problems [88, 89, 54]. It turns out that this convex relaxation approach (5.5) admits stable and faithful estimates even when $\mathbf{\Sigma}$ is approximately low rank and/or when the measurements are corrupted by bounded noise. This is formally stated in the following theorem.

Theorem 5.1. *Consider the sub-Gaussian sampling model in* (5.1) *and assume that* $\|\boldsymbol{\eta}\|_1 \leq \epsilon_1$*. Then with probability exceeding* $1 - c_2 \exp(-c_1 m)$*, the solution* $\hat{\mathbf{\Sigma}}$ *to* (5.5) *satisfies*

$$\|\hat{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_{\mathrm{F}} \leq C_1 \frac{\|\mathbf{\Sigma} - \mathbf{\Sigma}_r\|_*}{\sqrt{r}} + C_2 \frac{\epsilon_1}{m} \tag{5.6}$$

for all $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$*, provided that* $m > c_0 nr$*. Here,* $\mathbf{\Sigma}_r$ *represents the best rank-r approximation of* $\mathbf{\Sigma}$*, and* c_0*,* c_1*,* c_2*,* C_1 *and* C_2 *are some absolute constants.*

Remark 5.1. We emphasize that the quadratic sampling operator $\mathcal{A}$ fails to satisfy the recovery condition RIP-ℓ_2/ℓ_2 used in [89, 119] for the establishment of matrix recovery using full-rank measurement matrices with i.i.d. sub-Gaussian entries. This fact is formally pointed out by Candès et. al. in [54], which motivates us to propose a new analysis framework.

The main implications of Theorem 5.1 and its associated performance bound (5.6) are listed as follows.

1. **Exact Recovery from Noiseless Measurement**s. Consider the case where $\mathrm{rank}\,(\mathbf{\Sigma}) = r$. In the absence of noise, one can see from (5.6) that the trace minimization program (5.5) (with $\epsilon_1 = 0$) allows perfect covariance recovery with exponentially high probability, provided that the number m of measurements exceeds the order of nr. Notice that each PSD matrix can be uniquely decomposed as $\mathbf{\Sigma} = \boldsymbol{L}\boldsymbol{L}^\top$, where $\boldsymbol{L} \in \mathbb{R}^{n \times r}$ has *orthogonal* columns, which implies that the intrinsic degrees of freedom carried by PSD matrices is about $nr - \frac{r(r-1)}{2}$. That said, the theoretic sampling limit for perfect recovery is $\Theta\,(nr)$, indicating that our algorithm allows order-wise optimal recovery.

2. **Near-Optimal Universal Recovery**. The trace minimization program (5.5) allows universal recovery, in the sense that once the sensing vectors are chosen,

all low-rank covariance matrices can be perfectly recovered in the absence of noise. This highlights the power of convex relaxation, which allows universally accurate estimates as soon as the number of measurements exceeds the theoretic limit. In addition, the universality and optimality results hold for a large class of sub-Gaussian measurements beyond Gaussian sampling.

3. **Robust Recovery for Approximately Low-Rank Matrices**. In the absence of noise ($\epsilon = 0$), if $\boldsymbol{\Sigma}$ is approximately low-rank, then by (5.6) the reconstruction inaccuracy is at most

$$\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{F}} \leq \mathcal{O}\left(\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r\|_* / \sqrt{r}\right)$$

with probability at least $1 - \exp(-c_1 m)$, as soon as m is about the same order of nr. One can obtain a more intuitive understanding through the following *power-law* covariance model. Let λ_ℓ represent the ℓth largest singular value of $\boldsymbol{\Sigma}$, and suppose the decay of λ_ℓ obeys a power law, i.e. $\lambda_\ell \leq \frac{\alpha}{\ell^\beta}$ for some constant $\alpha > 0$ and decay rate exponent $\beta > 1$. Then

$$\frac{\|\boldsymbol{\Sigma} - \boldsymbol{\Sigma}_r\|_*}{\sqrt{r}} \leq \frac{1}{\sqrt{r}} \sum_{\ell=r+1}^{n} \frac{\alpha}{\ell^\beta} \leq \frac{\alpha}{(\beta - 1) r^{\beta - \frac{1}{2}}},$$

which in turn implies

$$\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{F}} = \mathcal{O}\left(r^{-\left(\beta - \frac{1}{2}\right)}\right). \tag{5.7}$$

This asserts that (5.5) reconstructs an almost accurate estimate of $\boldsymbol{\Sigma}$ in a manner which requires no prior knowledge on the signal (other than the power law decay that is natural for a broad class of data).

4. **Stable Recovery from Noisy Measurements**. When $\boldsymbol{\Sigma}$ is exactly of rank r and the noise is bounded $\|\boldsymbol{\eta}\|_1 \leq \epsilon_1$, the reconstruction inaccuracy satisfies

$$\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{F}} \leq C_2 \epsilon_1 / m \tag{5.8}$$

with exponentially high probability, provided that m exceeds $\Theta(nr)$. This reveals that (5.5) recovers an object with an error at most proportional to the average *per-entry* noise level, which makes it practically appealing.

5. **Phase Retrieval with Sub-Gaussian Measurements**. The proposed algorithm (5.5) appears in the same form as the convex algorithm called *PhaseLift*, which was proposed in [54] for phase retrieval. It is equivalent to treating $\boldsymbol{\Sigma}$ as the rank-one lifted matrix $\boldsymbol{x}\boldsymbol{x}^\top$. It has been established in [58] that with high probability it is possible to recover $\boldsymbol{x}$ exactly from $\Theta(n)$ quadratic measurements, assuming that the sensing vectors are i.i.d. Gaussian-distributed. Our result immediately recovers all results of [54, 58] including exact and stable recovery. In fact, our analysis framework yields a much simpler proof of all these results, and immediately extends to a broader class of sub-Gaussian sampling mechanisms. We will further discuss our improvement of sparse recovery from magnitude measurements [59, 120] in Section 5.2.3.

Remark 5.2. A lower bound on the minimax risk has recently been established by Cai and Zhang [121, Theorem 2.4]. Specifically, if the noise $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$ with $\sigma = \Theta\left(\frac{\epsilon_1}{m}\right)$, then for any estimator $\tilde{\boldsymbol{\Sigma}}(\boldsymbol{y})$,

$$\inf_{\tilde{\boldsymbol{\Sigma}}(\cdot)} \sup_{\boldsymbol{\Sigma}:\, \mathrm{rank}(\boldsymbol{\Sigma})=r} \sqrt{\mathbb{E}_{\boldsymbol{\eta}}\left[\left\|\tilde{\boldsymbol{\Sigma}}(\boldsymbol{y}) - \boldsymbol{\Sigma}\right\|_{\mathrm{F}}^2\right]} \geq c_8 \sigma = \Theta\left(\frac{\epsilon_1}{m}\right),$$

provided that $m = \Theta(nr)$. While our results are established for bounded (possibly adversarial) noise, it is straightforward to see that the above argument reveals the orderwise optimality of our stability bound when applying to stochastic Gaussian noise model.

5.2.2 Recovery of Low-Rank Covariance Matrices for Stationary Instances

Suppose that $\boldsymbol{\Sigma}$ is low-rank and represents the covariance matrix of n-dimensional stationary data instances. Similar to recovery in the general low-rank case, we seek a nuclear norm minimizer over all matrices compatible with the measurements. Since it is known *a priori* that $\boldsymbol{x}_i$ is stationary, we further impose a Toeplitz constraint to enforce stationarity conditions, which results in the following estimate

$$\hat{\boldsymbol{\Sigma}} = \arg\min_{\boldsymbol{M}} \operatorname{Tr}(\boldsymbol{M}) \quad \text{s.t.} \quad \boldsymbol{M} \succeq 0,\ \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{M})\|_2 \leq \epsilon_2,\ \boldsymbol{M} \text{ is Toeplitz}, \tag{5.9}$$

where ϵ_2 is an upper bound of $\|\boldsymbol{\eta}\|_2$.

Encouragingly, the semidefinite relaxation (5.9) is exact under noise-free measurements and provides stable recovery from noisy measurements, as asserted below.

Theorem 5.2. *Consider the sub-Gaussian sampling model in* (5.1)*, and assume that* $\mu_4 \leq 3$ *and* $\|\boldsymbol{\eta}\|_2 \leq \epsilon_2$*. Then with probability exceeding* $1 - 1/n^2$*,*

$$\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{F}} \leq C_2 \epsilon_2 / \sqrt{m} \tag{5.10}$$

holds for all Toeplitz covariance matrices $\boldsymbol{\Sigma}$ *of rank at most* r*, provided that* $m > c_0 r \log^{10} n$*. Here,* c_0 *and* C_2 *are some universal constants.*

Once we obtain accurate recovery of $\boldsymbol{\Sigma}$, the underlying spectrum can be identified by conventional harmonic retrieval methods, e.g. ESPRIT [34]. We highlight some implications of Theorem 5.2 as follows.

1. **Exact Recovery without Noise**. As any rank-r PSD Toeplitz matrix admits a unique rank-r Vandemonde decomposition that can be specified by $2r$ parameters, by Theorem 5.2, exact recovery of Toeplitz low-rank covariance matrices occurs as soon as m is slightly larger than the theoretic sampling limit $\Omega(r)$ (with some poly-logarithmic factor). Note that this sampling requirement is much smaller than that for general low-rank matrices, and also much smaller than the degrees of freedom for general Toeplitz matrices (which is n).

2. **Stable and Universal Recovery from Noisy Measurements**. The proposed convex relaxation (5.9) returns faithful estimates in the presence of noise, as revealed by Theorem 5.2. This feature is universal: if $\mathcal{A}$ is randomly sampled and then fixed thereafter, then, with high probability, the error bounds (5.10)

hold for all Toeplitz low-rank matrices. Note that the error bound (5.10) is stated in terms of the ℓ_2 norm of $\boldsymbol{\eta}$. This is out of mathematical convenience for this special setup, which will be discussed later.

Remark 5.3. Two aspects of Theorem 5.2 are worth noting. (a) Theorem 5.2 does not guarantee recovery with exponentially high probability as ensured in Theorem 5.1. This arises from our use of stochastic RIP, as will be seen in the analysis. (b) We are only able to provide theoretical guarantee when $\mu_4 \leq 3$; roughly speaking, the tails of these distributions are typically no heavier than Gaussian measure (e.g. $\mu_4 = 3$ for Gaussian distribution and $\mu_4 = 1$ for Bernoulli distribution). We conjecture that these two aspects can be improved via more delicate proof techniques.

5.2.3 Recovery of Jointly Sparse and Rank-One Matrices

If we set the covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{x}\boldsymbol{x}^\top$ to be a rank-one matrix, then covariance estimation from quadratic measurements is equivalent to phase retrieval as studied in [54]. In addition to the general rank-one model, our approach allows simple analysis for recovering jointly sparse and rank-one covariance matrices or, equivalently, sparse signal recovery from magnitude measurements.

Specifically, suppose that the dominant component $\boldsymbol{x}$ of the matrix $\boldsymbol{\Sigma}$ is (approximately) sparse, and the goal is to recover $\boldsymbol{x}\boldsymbol{x}^\top$ from a small number of phaseless measurements. The measurements we obtain can be expressed as

$$\boldsymbol{y} := \left\{ \left| \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \right|^2 + \eta_i \right\}_{1 \leq i \leq m}.$$

When $\boldsymbol{x}$ is sparse, the lifted matrix $\boldsymbol{x}\boldsymbol{x}^\top$ is *simultaneously* low rank and sparse [122], which motivates us to adapt the convex program proposed in [59] as follows

$$\underset{\boldsymbol{M} \in \mathcal{S}^{n \times n}}{\text{minimize}} \quad \operatorname{tr}(\boldsymbol{M}) + \lambda \left\| \boldsymbol{M} \right\|_1 \quad \text{s.t.} \quad \boldsymbol{M} \succeq 0, \;\; \left\| \boldsymbol{y} - \mathcal{A}(\boldsymbol{M}) \right\|_1 \leq \epsilon_1. \tag{5.11}$$

Here, λ is a regularization parameter that balances the two convex surrogates (i.e. trace norm and ℓ_1 norm) associated with the low-rank and sparse structural assumptions, respectively, and ϵ_1 is an upper bound of $\|\boldsymbol{\eta}\|_1$. Our analysis framework ensures stable recovery of an approximately sparse signal, as stated in the following theorem.

Theorem 5.3. *Set λ to be any number within the interval $\left[\frac{1}{n}, \frac{1}{\sqrt{k}}\rho\right]$ for some quantity ρ. Consider the sub-Gaussian sampling model in* (5.1) *and assume that* $\|\boldsymbol{\eta}\|_1 \leq \epsilon_1$. *Then with probability at least $1 - c_1 \exp(-c_0 m)$, the solution $\hat{\boldsymbol{X}}$ to (5.11) satisfies*

$$\left\|\hat{\boldsymbol{X}} - \boldsymbol{x}\boldsymbol{x}^\top\right\|_{\mathrm{F}} \leq C_1 \left\{ \left\|\boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{x}_\Omega \boldsymbol{x}_\Omega^\top\right\|_* + \lambda \left\|\boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{x}_\Omega \boldsymbol{x}_\Omega^\top\right\|_1 + \frac{\epsilon_1}{m} \right\} \tag{5.12}$$

for all signals $\boldsymbol{x}$ satisfying $\frac{\|\boldsymbol{x}_\Omega\|_2}{\|\boldsymbol{x}_\Omega\|_1} \geq \rho$, provided that $m > \frac{C_2 \log n}{\lambda^2}$. Here, $\boldsymbol{x}_\Omega$ denotes the best k-sparse approximation of $\boldsymbol{x}$, and C_1, C_2 c_0, and c_1 are absolute constants.

Theorem 5.3 recovers all the theoretical performance guarantees established in [59] with a simpler proof, and improves upon them in two aspects: (i) Theorem 5.3 establishes the performance guarantees of the algorithm (5.11) when the structural assumption is *imperfect* and when the samples are *noisy*; (ii) [59] considers only Gaussian sensing vectors, whereas we extend the results to a large class of sub-Gaussian sensing vectors. Some implications of Theorem 5.3 are as follows.

1. **Stable and Universal Recovery for Imperfect Models and Noisy Samples**. The recovered signal is a highly accurate estimate even when the sparsity assumption is inexact, provided that the true signal exhibits sufficiently fast decay outside the support Ω. The estimation inaccuracy due to noise corruption is also small, in the sense that it is at most proportional to the per-entry noise level. Besides, the recovery guarantee depends on the choice of λ, and is universal over a large class of signals with $\frac{\|\boldsymbol{x}_\Omega\|_2}{\|\boldsymbol{x}_\Omega\|_1} \geq \rho$.

2. **Near-Optimal Recovery for Power-Law Sparse Signals**. In general, by setting $\lambda = \frac{1}{k}$, one can obtain *universal* recovery for all k-sparse signals from $O(k^2 \log n)$ samples with exponentially high probability. Somewhat surprisingly, if the *nonzero entries* of $\boldsymbol{x}$ are known to be decaying, then the algorithm (5.11) allows near-optimal recovery. For instance, suppose that the *non-zero*

entries of $\boldsymbol{x}$ satisfies the power-law decay such that the magnitude of the lth largest entry of $\boldsymbol{x}_\Omega / \|\boldsymbol{x}_\Omega\|_2$ is bounded above by c_{pl}/l^α for some constants c_{pl} and exponent $\alpha > 1$. By setting $\lambda = \Theta\left(\frac{1}{\sqrt{k}\log n}\right)$, one can obtain accurate recovery from $O\left(k\log^2 n\right)$ noiseless samples, which is only a logarithmic factor from the theoretic sampling limit (which is $\Theta(k)$).

5.3 Approximate ℓ_2/ℓ_1 Isometry for Low-rank and Sparse Matrices

In this section, we present a novel concept called the mixed-norm restricted isometry property (RIP-ℓ_2/ℓ_1) that allows us to establish Theorems 5.1 and 5.3 concerning recovery of low-rank, and jointly sparse and rank-one covariance matrices.

Prevailing wisdom in CS asserts that perfect recovery from minimal samples is possible if the dimensionality reduction projection preserves the signal strength when acting on the class of matrices of interest [24, 89]. While there are various ways to define the restricted isometry properties (RIP), an appropriately chosen approximate isometry leads to a very simple yet powerful theoretical framework.

5.3.1 Mixed-Norm Restricted Isometry (RIP-ℓ_2/ℓ_1)

Recall that the RIP occurs if the sampling output preserves the input strength under certain metrics. The most commonly used one is RIP-ℓ_2/ℓ_2, for which the signal strength before and after the projection are both measured in terms of the Frobenius norm [123, 89]. This, however, fails to hold under rank-one measurements – see detailed argument by Candès et al [54]. Another isometry concept called RIP-ℓ_1/ℓ_1 has also been investigated, for which the signal strength before and after the operation $\mathcal{A}$ are measured both in terms of the ℓ_1 norms[1]. This is initially developed to account for measurements from expander graphs [124], and has become a powerful metric

[1] Note that the nuclear norm is the ℓ_1-norm counterpart for matrices.

when analyzing phase retrieval [54, 59]. Nevertheless, when considering general low-rank matrices, RIP-ℓ_1/ℓ_1 no longer holds [2]. Moreover, the proof based on RIP-ℓ_1/ℓ_1 typically relies on delicate construction of dual certificates [54, 59], which is often mathematically complicated.

One of the key and novel ingredients in our analysis is a mixed-norm approximate isometry, which measures the signal strength before and after sampling with different metrics. Specifically, we introduce RIP-ℓ_2/ℓ_1, where the input and output are measured in terms of the Frobenius norm and the ℓ_1 norm, respectively. It turns out that as long as the input is measured with the Frobenius norm, the trick pioneered in [123] can be carried over to our problem, which saves the need for dual construction. We make formal definitions of RIP-ℓ_2/ℓ_1 for low-rank/sparse matrices as follows.

Definition 5.1 (**RIP-ℓ_2/ℓ_1 for low-rank matrices**). *For the set of rank-r matrices, we define the RIP-ℓ_2/ℓ_1 constants δ_r^{lb} and δ_r^{ub} with respect to an operator $\mathcal{B}$ as the smallest numbers such that for all $\boldsymbol{X}$ of rank at most r:*

$$\left(1-\delta_r^{\mathrm{lb}}\right)\|\boldsymbol{X}\|_{\mathrm{F}} \leq \frac{1}{m}\|\mathcal{B}(\boldsymbol{X})\|_1 \leq \left(1+\delta_r^{\mathrm{ub}}\right)\|\boldsymbol{X}\|_{\mathrm{F}}.$$

Definition 5.2 (**RIP-ℓ_2/ℓ_1 for low-rank plus sparse matrices**). *Consider the class of index sets $\mathcal{S}_k := \{\omega\times\omega \mid \omega\in[n] \text{ and } \mathrm{card}(\omega)=k\}$. For the set of matrices*

$$\mathcal{M}_{r,l}^k = \left\{\boldsymbol{X}_1+\boldsymbol{X}_2 \mid \mathrm{supp}\left(\boldsymbol{X}_1\right)\in\mathcal{S}_k,\ \mathrm{rank}\left(\boldsymbol{X}_1\right)\leq r,\ \ \left\|\boldsymbol{X}_2\right\|_0\leq l\right\}, \tag{5.13}$$

we define the RIP-ℓ_2/ℓ_1 constants $\delta_{r,l}^{\mathrm{lb},k}$ and $\delta_{r,l}^{\mathrm{lb},k}$ with respect to an operator $\mathcal{B}$ as the smallest numbers such that $\forall\boldsymbol{X}\in\mathcal{M}_{r,l}^k$,

$$\left(1-\delta_{r,l}^{\mathrm{lb},k}\right)\|\boldsymbol{X}\|_{\mathrm{F}} \leq \frac{1}{m}\|\mathcal{B}(\boldsymbol{X})\|_1 \leq \left(1+\delta_{r,l}^{\mathrm{ub},k}\right)\|\boldsymbol{X}\|_{\mathrm{F}}.$$

Remark 5.4. In short, any matrix within $\mathcal{M}_{r,l}^k$ can be decomposed into two components $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, where $\boldsymbol{X}_1$ is simultaneously low-rank and sparse, and $\boldsymbol{X}_2$ is sparse.

[2] For instance, consider two matrices $\boldsymbol{X}_1=\mathrm{diag}\{\boldsymbol{I}_{r/2},\boldsymbol{I}_{r/2},\boldsymbol{0}\}$ and $\boldsymbol{X}_2=\mathrm{diag}\{\boldsymbol{I}_{r/2},-\boldsymbol{I}_{r/2},\boldsymbol{0}\}$ with the same nuclear norm. When $m=\Omega(nr)$, one can show from the Bernstein inequality for sub-exponential random variables that $\frac{1}{m}\|\mathcal{A}(\boldsymbol{X}_1)\|_1=\Theta(r)$ and $\frac{1}{m}\|\mathcal{A}(\boldsymbol{X}_2)\|_1=\Theta(\sqrt{r})$, precluding the existence of a small RIP-ℓ_1/ℓ_1 constant.

This allows us to treat each matrix perturbation as a superposition of a collection of jointly low-rank and sparse matrices and a collection of general sparse matrices, where the rank-one measurements of each term can be well controlled under minimal sample complexity.

5.3.2 RIP-ℓ_2/ℓ_1 of Quadratic Measurements for Low-rank and Sparse Matrices

Unfortunately, the original sampling operator $\mathcal{A}$ does not satisfy RIP-ℓ_2/ℓ_1. This occurs primarily since each measurement matrix $\boldsymbol{A}_i$ has non-zero mean, which biases the output measurements. In order to get rid of this undesired bias effect, we introduce a set of "debiased" auxiliary measurement matrices as follows

$$\boldsymbol{B}_i := \boldsymbol{A}_{2i-1} - \boldsymbol{A}_{2i}. \tag{5.14}$$

Without loss of generality, denote $\mathcal{B}_i(\boldsymbol{X}) := \langle \boldsymbol{B}_i, \boldsymbol{X} \rangle$ for all $1 \le i \le m$, and let $\mathcal{B}(\boldsymbol{X})$ represent the linear transformation that maps $\boldsymbol{X}$ to $\{\mathcal{B}_i(\boldsymbol{X})\}_{i=1}^m$. Note that by representing the sensing process using m rank-2 measurements $\mathcal{B}_i$, we have implicitly doubled the number of measurements for notational simplicity. This, however, will not change our order-wise results.

It turns out that the auxiliary operator $\mathcal{B}$ exhibits RIP-ℓ_2/ℓ_1 in the presence of minimal measurements, which can be shown by combining the following proposition with a standard covering argument as applied in [119].

Proposition 5.1. *Let $\mathcal{A}$ be sampled from the sub-Gaussian model in* (5.1)*. For any matrix $\boldsymbol{X}$, there exist universal constants $c_1, c_2, c_3, C > 0$ such that with probability exceeding $1 - C\exp(-c_3 m)$, one has*

$$c_1 \|\boldsymbol{X}\|_{\mathrm{F}} \le \frac{1}{m} \|\mathcal{B}(\boldsymbol{X})\|_1 \le c_2 \|\boldsymbol{X}\|_{\mathrm{F}}. \tag{5.15}$$

An immediate consequence of Proposition 5.1 is the establishment of the RIP-ℓ_2/ℓ_1 of the sampling operator $\mathcal{B}$ for either general low-rank or sparse matrices. The proof of the corollaries below follows immediately from a standard covering argument detailed in [119, Section III.B] and [125, Section 5]. We thus omit the details but refer interested readers to the above references for details.

Corollary 5.1 (RIP-ℓ_2/ℓ_1 for low-rank matrices). *Consider the sub-Gaussian sampling model in* (5.1) *and the constants* $c_1, c_2 > 0$ *given in (5.15). There exist universal constants* $c_3, c_4, C > 0$ *such that with probability exceeding* $1 - C\exp(-c_3 m)$*,* $\mathcal{B}$ *satisfies the RIP-*ℓ_2/ℓ_1 *for all* $\boldsymbol{X}$ *of rank at most* r*, and obeys*

$$1 - \delta_r^{\mathrm{lb}} \geq c_1/2, \quad 1 + \delta_r^{\mathrm{ub}} \leq 2c_2, \tag{5.16}$$

provided that $m > c_4 nr$*.*

Corollary 5.2 (RIP-ℓ_2/ℓ_1 for low-rank plus sparse matrices). *Consider the sub-Gaussian sampling model in* (5.1) *and the constants* $c_1, c_2 > 0$ *given in (5.15). There exist universal constants* $c_3, c_4, C > 0$ *such that with probability at least* $1 - C\exp(-c_3 m)$*,* $\mathcal{B}$ *satisfies the RIP-*ℓ_2/ℓ_1 *w.r.t.* $\mathcal{M}_{r,l}^k$ *(defined in (5.13)), which obeys*

$$1 - \delta_{r,l}^{\mathrm{lb},k} \geq \frac{c_1}{2}, \quad 1 + \delta_{r,l}^{\mathrm{ub},k} \leq 2c_2, \tag{5.17}$$

provided that $m > c_4 \max\{kr\log n, l\log(n/l)\}$*.*

5.3.3 Proof Outline of Theorems 5.1 and 5.3 via RIP-ℓ_2/ℓ_1

Theorems 5.1 and 5.3 can thus be proved given that reasonably small RIP-ℓ_2/ℓ_1 constants with respect to the auxiliary operator $\mathcal{B}$ are guaranteed via Corollaries 5.1. We first present Lemma 5.1 which establishes Theorem 5.1.

Lemma 5.1. *Consider any matrix* $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_r + \boldsymbol{\Sigma}_{\mathrm{c}}$*, where* $\boldsymbol{\Sigma}_r$ *is the best rank-*r *approximation of* $\boldsymbol{\Sigma}$*. If there exists a number* $K_1 > 2r$ *such that*

$$\frac{1 - \delta_{2r+K_1}^{\mathrm{lb}}}{\sqrt{2}} - \left(1 + \delta_{K_1}^{\mathrm{ub}}\right)\sqrt{\frac{2r}{K_1}} \geq \beta_1 > 0 \tag{5.18}$$

holds for some absolute constant β, then the minimizer $\hat{\mathbf{\Sigma}}$ to (5.5) *obeys*

$$\|\hat{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_{\mathrm{F}} \le C_1 \frac{\|\mathbf{\Sigma}_{\mathrm{c}}\|_*}{\sqrt{K_1}} + C_2 \frac{\epsilon_1}{m} \tag{5.19}$$

for some constants C_1 and C_2 depending only on the restricted isometry constants and β_1.

By choosing $K_1 = 8\left(\frac{4c_2}{c_1}\right)^2 r \ge 8\left(\frac{1+\delta_{K_1}^{\mathrm{ub}}}{1-\delta_{2r+K_1}^{\mathrm{ub}}}\right)^2 r$ for the universal constants c_1, c_2 given in Corollary 5.1, we obtain (5.18) when $m > c_4 (K_1 + 2r) n$ for some constant c_4. This establishes Theorem 5.1.

Furthermore, the specialized RIP-ℓ_2/ℓ_1 concept allows us to prove Theorem 5.3 through the following lemma.

Lemma 5.2. *Set λ to be any number within the interval $\left[\frac{1}{n}, \frac{1}{\sqrt{k}} \frac{\|\boldsymbol{x}_\Omega\|_2}{\|\boldsymbol{x}_\Omega\|_1}\right]$. Suppose that $\boldsymbol{x}_\Omega$ is the best k-sparse approximation of $\boldsymbol{x}_0$. If there exists a number K_1 such that*

$$\frac{\frac{1-\delta^{lb}_{2K_1, 2\frac{K_1}{\lambda^2}}}{\sqrt{3}} - \frac{3\left(1+\delta^{ub,k}_{K_1, \frac{K_1}{\lambda^2}}\right)}{\sqrt{K_1}}}{2\max\left\{\frac{1+\delta^{ub,k}_{K_1, \frac{K_1}{\lambda^2}}}{\sqrt{K_1}}, 1\right\}} \ge \beta_3 > 0, \quad \textit{and} \quad \frac{1+\delta^{ub,k}_{K_1, \frac{K_1}{\lambda^2}}}{\left(1-\delta^{lb,k}_{K_1, \frac{K_1}{\lambda^2}}\right)\sqrt{K_1}} \le \beta_4 \tag{5.20}$$

for some absolute constants β_3 and β_4, then the solution $\hat{\boldsymbol{X}}$ to (5.11) satisfies

$$\left\|\hat{\boldsymbol{X}} - \boldsymbol{x}_\Omega \boldsymbol{x}_\Omega^\top\right\|_{\mathrm{F}} \le C\left(\left\|\boldsymbol{X} - \boldsymbol{x}_\Omega \boldsymbol{x}_\Omega^\top\right\|_* + \lambda\left\|\boldsymbol{X} - \boldsymbol{x}_\Omega \boldsymbol{x}_\Omega^\top\right\|_1 + \frac{\epsilon_1}{m}\right) \tag{5.21}$$

for some constant C that depends only on β_3 and β_4.

By Corollary 5.2, one can ensure small RIP-ℓ_2/ℓ_1 constants as soon as

$$m > c_4 \max\left\{kK_1 \log n, \frac{K_1}{\lambda^2}\log n\right\} = c_4 \frac{K_1}{\lambda^2}\log n.$$

This in turn establishes Theorem 5.3.

Finally, note that we have not discussed general Toeplitz low-rank matrices using RIP-ℓ_2/ℓ_2. We are unaware of a rigorous approach to prove exact recovery using RIP-ℓ_2/ℓ_1 for the Toeplitz case. Fortunately, the analysis for Toeplitz low-rank matrices can be performed via a different method, which we detail in the next section.

5.4 Approximate ℓ_2/ℓ_2 Isometry for Toeplitz Low-Rank Matrices

While quadratic measurements in general do not exhibit RIP-ℓ_2/ℓ_2 (as introduced in [89]) with respect to the set of general low-rank matrices (as pointed out in [54]), a slight variant of them can indeed satisfy RIP-ℓ_2/ℓ_2 when restricted to *Toeplitz* low-rank matrices. In this subsection, we first provide a characterization of RIP-ℓ_2/ℓ_2 for the set of general low-rank matrices under bounded and near-isotropic measurements, and then transform quadratic measurements into equivalent isotropic measurements.

5.4.1 RIP-ℓ_2/ℓ_2 for Near-Isotropic Measurements

Before proceeding to the Toeplitz low-rank matrices, we investigate near-isotropic and bounded operators for the set of general low-rank matrices as follows. For convenience of presentation, we repeat the definition of RIP-ℓ_2/ℓ_2 as follows, followed by a theorem characterizing RIP-ℓ_2/ℓ_2 for near-isotropic and bounded operators.

Definition 5.3 (**RIP-ℓ_2/ℓ_2 for low-rank matrices**). *For the set of rank-r matrices, we define the RIP-ℓ_2/ℓ_2 constants δ_r w.r.t. an operator $\mathcal{B}$ as the smallest numbers such that for all $\boldsymbol{X}$ of rank at most r,*

$$(1-\delta_r)\left\|\boldsymbol{X}\right\|_{\mathrm{F}} \leq \frac{1}{m}\left\|\mathcal{B}\left(\boldsymbol{X}\right)\right\|_2 \leq (1+\delta_r)\left\|\boldsymbol{X}\right\|_{\mathrm{F}}.$$

Theorem 5.4. *Suppose that for all $1 \leq i \leq m$,*

$$\left\|\boldsymbol{B}_i\right\| \leq K \quad \textit{and} \quad \left\|\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]-\mathcal{I}\right\| \leq \frac{c_5}{n} \tag{5.22}$$

hold for some quantity $K \leq n^2$. For any small constant $\delta > 0$, if[3] $m > c_0 r K^2 \log^7 n$, *then with probability at least $1 - 1/n^2$, one has*

i) $\mathcal{B}$ satisfies RIP-ℓ_2/ℓ_2 w.r.t. all matrices of rank at most r and obeys $\delta_r \leq \delta$;

ii) Suppose that $\mathcal{K}$ is some convex set. Then for all $\boldsymbol{\Sigma}$ of rank at most r and $\boldsymbol{\Sigma} \in \mathcal{K}$, if $\|\boldsymbol{y} - \mathcal{B}(\boldsymbol{\Sigma})\|_2 \leq \epsilon_2$, the solution

$$\hat{\boldsymbol{\Sigma}} = \operatorname{argmin}_{\boldsymbol{M}} \|\boldsymbol{M}\|_* \quad \text{subject to} \quad \|\boldsymbol{y} - \mathcal{B}(\boldsymbol{M})\|_2 \leq \epsilon_2, \quad \boldsymbol{M} \in \mathcal{K} \tag{5.23}$$

satisfies

$$\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{F}} \leq C_2 \frac{\epsilon_2}{\sqrt{m}}. \tag{5.24}$$

Here, $c_0, C_2, c_5 > 0$ are some universal constants.

In fact, the bound on $\|\boldsymbol{B}_i\|$ can be as small as $\Theta\left(\sqrt{n}\right)$, and we say a measurement matrix $\boldsymbol{B}_i$ is *well-bounded* if $K = O\left(\sqrt{n}\text{poly}\log\left(n\right)\right)$. Simultaneously well-bounded and near-isotropic operators (i.e. those satisfying (5.22)) subsume the Fourier-type basis as discussed in [61], which admits a small RIP-ℓ_2/ℓ_2 constant as soon as $m = \Omega\left(nr\text{poly}\log\left(n\right)\right)$. Theorem 5.4 strengthens the result in [61] by justifying RIP-ℓ_2/ℓ_2, universal and stable recovery, which are not revealed by the approach of [61].

Unfortunately, Theorem 5.4 cannot be directly applied to the class of Toeplitz low-rank matrices for the following reasons: i) The sampling operator $\mathcal{A}$ is neither isotropic nor well-bounded; ii) Theorem 5.4 requires $m > c_0 r K^2 \text{poly}\log(n) = \Omega\left(nr\text{poly}\log(n)\right)$ measurements, which far exceeds the measurement complexity stated in Theorem 5.2. This motivates us to construct another set of equivalent sampling operators that satisfy the assumptions of Theorem 5.4, which is the focus of the following subsection.

[3] The proof of Theorem 5.4 follows the entropy method introduced in [60]. The $\log^7 n$ factor is a consequence of the entropy method, which might be refined a bit by generic chaining due to Talagrand [126] as employed in [91].

5.4.2 Construction of RIP-ℓ_2/ℓ_2 Operators for Toeplitz Low-rank Matrices

Note that the quadratic measurement matrices $\boldsymbol{A}_i = \boldsymbol{a}_i \cdot \boldsymbol{a}_i^\top$ are neither non-isotropic nor well bounded. For instance, when $\boldsymbol{a}_i \sim \mathcal{N}(0, \boldsymbol{I}_n)$, simple calculation reveals that

$$\|\boldsymbol{A}_i\| = \Theta\left(\sqrt{n}\right), \quad \text{and} \quad \mathbb{E}\left[\boldsymbol{A}_i \langle \boldsymbol{A}_i, \boldsymbol{X}\rangle\right] = 2\boldsymbol{X} + \operatorname{tr}(\boldsymbol{X}) \cdot \boldsymbol{I}, \tag{5.25}$$

precluding $\boldsymbol{A}_i$'s from being isotropic and well-bounded. In order to apply Theorem 5.4, we generate a new set of measurement matrices $\tilde{\boldsymbol{B}}_i$ via the following procedure.

1. Define a set of matrices $\boldsymbol{B}_i$ of rank at most 3

$$\boldsymbol{B}_i := \begin{cases} \frac{1}{2}\left(\boldsymbol{A}_{2i-1} - \boldsymbol{A}_{2i}\right), & \text{if } \mu_4 = 3, \\ \alpha \boldsymbol{A}_{3i} + \beta \boldsymbol{A}_{3i-1} + \gamma \boldsymbol{A}_{3i-2}, & \text{if } \mu_4 < 3, \end{cases} \tag{5.26}$$

   where α, β, γ are specified in Lemma 5.3.

2. Generate M matrices independently such that[4]

$$\hat{\boldsymbol{B}}_i = \begin{cases} \sqrt{n}\mathcal{T}\left(\boldsymbol{B}_i\right), & \text{with probability } \frac{1}{n}, \\ \sqrt{\frac{n}{n-1}}\mathcal{T}^\perp\left(\boldsymbol{G}_i\right), & \text{with probability } \frac{n-1}{n}, \end{cases} \tag{5.27}$$

   where $\boldsymbol{G}_i$ is a random matrix with i.i.d. standard Gaussian entries.

3. Define a *truncated* version $\tilde{\boldsymbol{B}}_i$ of $\hat{\boldsymbol{B}}_i$ as follows

$$\tilde{\boldsymbol{B}}_i := \hat{\boldsymbol{B}}_i \mathbf{1}_{\left\{\left\|\hat{\boldsymbol{B}}_i\right\| \leq c_{10} \log^{3/2} n\right\}}, \quad 1 \leq i \leq M. \tag{5.28}$$

We will demonstrate that the $\tilde{\boldsymbol{B}}_i$'s are nearly-isotropic and well-bounded, and hence by Theorem 5.4 the associated operator $\tilde{\mathcal{B}}$ enables exact and stable recovery for all rank-r matrices when M exceeds $nr\operatorname{poly}\log(n)$. This in turn establishes Theorem 5.2 through an equivalence argument, detailed below.

[4] We choose M to be about $\Theta(nm)$, which will be made clear later.

5.4.2.1 Isotropy Trick

While $\boldsymbol{A}_i$'s are in general non-isotropic, a linear combination of them can be made isotropic when restricted to Toeplitz matrices. This is stated in the following lemma.

Lemma 5.3. *Consider the sub-Gaussian sampling model in* (5.1).

1) When $\mu_4 = 3$, *then for any* $\boldsymbol{X}$, *the matrix*

$$\boldsymbol{B}_i = \frac{1}{2}\left(\boldsymbol{A}_{2i-1} - \boldsymbol{A}_{2i}\right) \tag{5.29}$$

satisfies $\mathbb{E}\left[\boldsymbol{B}_i \left\langle \boldsymbol{B}_i, \boldsymbol{X}\right\rangle\right] = \boldsymbol{X}$.

2) When $\mu_4 < 3$, *take any constant* $\xi > 0$ *obeying* $\xi^2 > 1.5 \cdot (3 - \mu_4)$ *and set*

$$\boldsymbol{B}_i = \alpha \boldsymbol{A}_{3i} + \beta \boldsymbol{A}_{3i-1} + \gamma \boldsymbol{A}_{3i-2}, \tag{5.30}$$

with the choice of $\Delta := -\left(1 - \frac{\xi}{n}\right)^2 - 2 + \frac{2\xi^2}{3-\mu_4}$,

$$\alpha = \sqrt{\frac{3-\mu_4}{2\xi^2}}, \quad \beta := \frac{-\left(1 - \frac{\xi}{\sqrt{n}}\right) + \sqrt{\Delta}}{2}\alpha, \quad \textit{and} \quad \gamma := \frac{-\left(1 - \frac{\xi}{\sqrt{n}}\right) - \sqrt{\Delta}}{2}\alpha. \tag{5.31}$$

Then, for any norm $\|\cdot\|_{\mathrm{n}}$ *and any* $\boldsymbol{X}$ *that satisfies* $\boldsymbol{X}_{11} = \boldsymbol{X}_{22} = \cdots = \boldsymbol{X}_{nn}$, *one has*

$$\begin{cases} \mathbb{E}\left[\boldsymbol{B}_i\right] = \sqrt{\frac{3-\mu_4}{2n}}; \\ \mathbb{E}\left[\boldsymbol{B}_i \left\langle \boldsymbol{B}_i, \boldsymbol{X}\right\rangle\right] = \boldsymbol{X}; \\ \left\|\boldsymbol{B}_i\right\|_{\mathrm{n}} \leq \sqrt{3}\max_{i:1\leq i\leq m}\left\|\boldsymbol{A}_i\right\|_{\mathrm{n}}. \end{cases} \tag{5.32}$$

Lemma 5.3 asserts that a large class of measurement matrices can made isotropic when restricted to the class of matrices with equal diagonal entries (e.g. Toeplitz matrices). This immediately implies that the operator $\hat{\mathcal{B}}$ associated with $\hat{\boldsymbol{B}}_i$'s (defined in (5.27)) are isotropic. Specifically, for any symmetric $\boldsymbol{X}$,

$$\begin{aligned} \mathbb{E}\left[\hat{\boldsymbol{B}}_i \left\langle \hat{\boldsymbol{B}}_i, \boldsymbol{X}\right\rangle\right] &= \mathbb{E}\left[\mathcal{T}\left(\boldsymbol{B}_i\right)\left\langle \boldsymbol{B}_i, \mathcal{T}\left(\boldsymbol{X}\right)\right\rangle\right] + \mathbb{E}\left[\mathcal{T}^{\perp}\left(\boldsymbol{G}_i\right)\left\langle \boldsymbol{G}_i, \mathcal{T}^{\perp}\left(\boldsymbol{X}\right)\right\rangle\right] \\ &= \mathcal{T}\left(\boldsymbol{X}\right) + \mathcal{T}^{\perp}\left(\boldsymbol{X}\right) = \boldsymbol{X}, \end{aligned} \tag{5.33}$$

which is a consequence of Lemma 5.3.

5.4.2.2 Truncation of $\hat{\mathcal{B}}$ is near-isotropic

The operators associated with $\hat{\boldsymbol{B}}_i$'s are in general not well-bounded. Fortunately, $\hat{\boldsymbol{B}}_i$'s are well-bounded with high probability, which follows from the following lemma.

Lemma 5.4. *Consider $\boldsymbol{z}$ follows the sub-Gaussian sampling model in* (5.1)*. There exists an absolute constant $c_{10} > 0$ such that*

$$\left\|\mathcal{T}\left(\boldsymbol{z}\boldsymbol{z}^{\top}\right)\right\| \leq c_{12} \log^{\frac{3}{2}} n \tag{5.34}$$

holds with probability exceeding $1 - n^{-10}$.

As $\|\boldsymbol{B}_i\|$ can be bounded above by $\max_{1\leq i\leq m}\|\boldsymbol{A}_i\|$ up to some constant factor, Lemma 5.4 provides a tight estimate (within some logarithmic factor) of $\|\mathcal{T}(\boldsymbol{B}_i)\|$ for sub-Gaussian vectors, i.e.

$$\|\mathcal{T}(\boldsymbol{B}_i)\| \leq c_{10} \log^{\frac{3}{2}} n, \quad 1 \leq i \leq m \tag{5.35}$$

with probability exceeding $1 - 3n^{-8}$. Similarly, classical results in random matrices (e.g. [127]) assert that $\|\boldsymbol{G}_i\|$ can also be bounded above by $\mathcal{O}(\sqrt{n}\log n)$ with overwhelming probability. These bounds taken collectively suggest that

$$\|\hat{\boldsymbol{B}}_i\| \leq K := c_{10}\sqrt{n} \log^{\frac{3}{2}} n, \quad 1 \leq i \leq m \tag{5.36}$$

for some constant $c_{10} > 0$ with probability exceeding $1 - n^{-7}$.

The above stochastically boundedness property motivates us to study the truncated version $\tilde{\boldsymbol{B}}_i$ of $\hat{\boldsymbol{B}}_i$ as defined in (5.28). Interestingly, $\tilde{\boldsymbol{B}}_i$ is near-isotropic, a consequence of the following lemma whose proof can be found in Appendix D.7.

Lemma 5.5. *Suppose that the restriction of $\mathcal{B}_i$ to Toeplitz matrices is isotropic. Consider any event E obeying $\mathbb{P}(E) \geq 1 - \frac{1}{n^5}$. Then there is some constant $c_5 > 0$ such that*

$$\|\mathbb{E}[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_E] - \mathcal{T}\| \leq \frac{c_5}{n^2}. \tag{5.37}$$

The truncated version of $\boldsymbol{G}_i$ can be easily bounded as in [91], which we omit for simplicity of presentation. This combined with (5.37) indicates that

$$\left\|\mathbb{E}\left[\tilde{\mathcal{B}}_i^*\tilde{\mathcal{B}}_i\right]-\mathcal{I}\right\| \leq \left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\right]-\mathcal{T}\right\| + \left\|\mathbb{E}\left[\mathcal{T}^{\perp}\mathcal{G}_i^*\mathcal{G}_i\mathcal{T}^{\perp}\right]-\mathcal{T}^{\perp}\right\| \leq \frac{c_5}{n}. \tag{5.38}$$

5.4.3 Proof of Theorem 5.2

So far we have demonstrated that $\tilde{\boldsymbol{B}}_i$'s are near-isotropic and satisfy $\left\|\tilde{\boldsymbol{B}}_i\right\| = O\left(\sqrt{n}\log^{\frac{3}{2}} n\right)$. Suppose that $\left\|\boldsymbol{y}-\tilde{\mathcal{B}}\left(\boldsymbol{\Sigma}\right)\right\|_2 \leq \tilde{\epsilon}_2$. Theorem 5.4 implies that if M exceeds $\Theta\left(nr\log^{10}(n)\right)$, then the solution to

$$\tilde{\boldsymbol{\Sigma}} := \operatorname{argmin}_{\boldsymbol{M}} \|\boldsymbol{M}\|_* \quad \text{subject to } \left\|\boldsymbol{y}-\tilde{\mathcal{B}}\left(\boldsymbol{M}\right)\right\|_2 \leq \tilde{\epsilon}_2, \quad \boldsymbol{M} \text{ is Toeplitz} \tag{5.39}$$

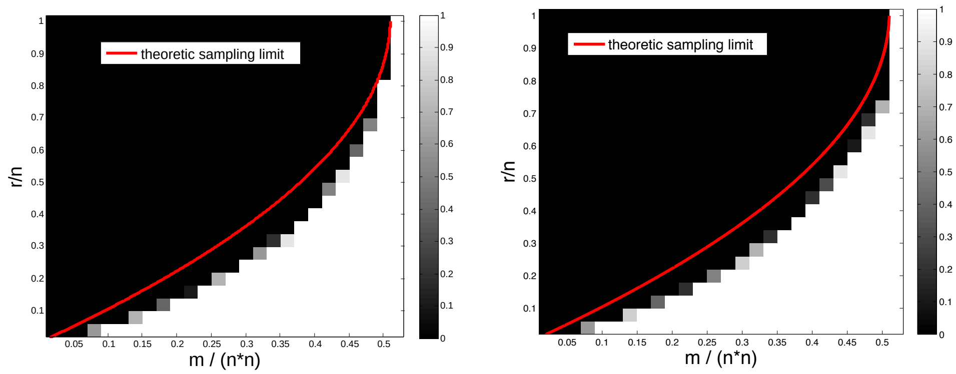
satisfies

$$\left\|\boldsymbol{\Sigma}-\tilde{\boldsymbol{\Sigma}}\right\|_{\mathrm{F}} \leq C_2\frac{\tilde{\epsilon}_2}{\sqrt{M}} \tag{5.40}$$

for the entire set of rank-r matrices $\boldsymbol{\Sigma}$. Apparently, such low-rank manifold subsumes all rank-r Toeplitz matrices as special cases. This claim in turn establishes Theorem 5.2 through the following argument:

1. From (5.27) and the Chernoff bound, $\tilde{\mathcal{B}}$ entails $\Theta\left(\frac{M}{n}\right)=\Theta\left(r\log^{10} n\right)$ independent copies of $\sqrt{n}\mathcal{T}\left(\boldsymbol{B}_i\right)$, and all other measurements are on the orthogonal complement of the Toeplitz space.

2. For any rank-r Toeplitz matrix $\boldsymbol{\Sigma}$, the original $\mathcal{A}$ entails $\frac{m}{3} > \Theta\left(r\log^{10} n\right)$ measurement matrices of the form $\mathcal{T}\left(\boldsymbol{B}_i\right)$, and any non-Toeplitz component of $\boldsymbol{X}$ is perfectly known (i.e. equal to 0). This indicates that the convex program (5.9) is tighter than (5.39) when $\tilde{\epsilon}_2=\Theta\left(\sqrt{n}\epsilon_2\right)$, i.e. one can construct (via coupling) a new probability space over which if the solution $\tilde{\boldsymbol{\Sigma}}$ to (5.39) is exact and unique, then it will be the unique solution to (5.9) as well. This combined with the universal bound (5.40) establishes Theorem 5.2.



(a) (b)

Figure 5.1: Recovery of covariance matrices from quadratic measurements when $n = 50$. The colormap for each cell indicates the empirical probability of success, and the red line reflects the fundamental degrees of freedom. The results are shown for (a) Gaussian sensing vectors and (b) symmetric Bernoulli sensing vectors.

5.5 Numerical Examples and Discussions

To demonstrate the practical applicability of the proposed convex relaxation under quadratic sensing, in this section we present a variety of numerical examples for low-rank or sparse covariance matrix estimation.

5.5.1 Recovery of Low-Rank Covariance Matrices

We conduct a series of Monte Carlo trials for various parameters. Specifically, we choose $n = 50$, and for each (m, r) pair, we repeat the following experiments 20 times. We generate $\boldsymbol{\Sigma}$, an $n \times n$ PSD matrix via $\boldsymbol{\Sigma} = \boldsymbol{L}\boldsymbol{L}^\top$, where $\boldsymbol{L}$ is a randomly generated $n \times r$ matrix with independent Gaussian components. The sensing vectors are generated as i.i.d. Gaussian vectors and Bernoulli vectors, and we obtain noiseless quadratic measurements $\boldsymbol{y}$. We use the off-the-shelf SDP solver SDPT3 with the modeling software CVX [118], and declare a matrix $\boldsymbol{\Sigma}$ to be recovered if the solution $\hat{\boldsymbol{\Sigma}}$ returned by the solver satisfies $\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\|_{\mathrm{F}} / \|\boldsymbol{\Sigma}\|_{\mathrm{F}} < 10^{-3}$. Figure 5.1 illustrates the empirical probability of successful recovery in these Monte Carlo trials, which is reflected through the color of each cell. In order to compare the optimality of the practical performance, we also plot the theoretic limit in red lines, i.e. the fundamental lower limit on m required to recover all rank-r matrices, which is $nr - \frac{r(r-1)}{2}$ in our case. It turns out that the practical phase transition curve is very close to the theoretic sampling limit, which demonstrates the optimality of our algorithm.

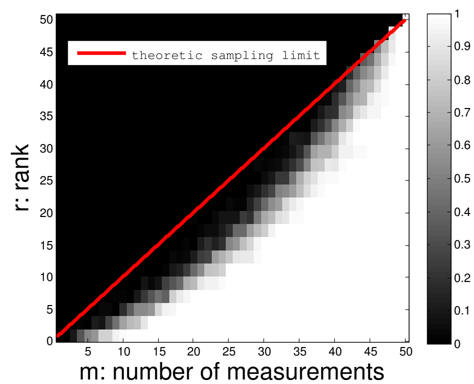


Figure 5.2: Phase transition plots where frequency locations are randomly generated. The plot corresponds to the situation where $n = 50$. The empirical success rate is calculated by averaging over 50 Monte Carlo trials.

5.5.2 Recovery of Toeplitz Low-rank Matrices

To justify the convex heuristic for Toeplitz low-rank matrices, we perform a series of numerical experiments for matrices of dimension $n = 50$. By Caratheodory's theorem, each PSD Toeplitz matrix can be uniquely decomposed into a linear combination of line spectrums [128]. Thus, we generate the PSD Toeplitz matrix by randomly generating the frequencies and amplitudes of each line spectra. In the real-valued situation, the underlying spectral spikes occur in conjugate pairs (i.e. $(f_1, -f_1), (f_2, -f_2), \cdots$). We independently generate $r/2$ frequency pairs within the unit disk uniformly at random, and the amplitudes are generated as the absolute values of i.i.d. Gaussian variables. Figure 5.2 illustrates the phase transition diagram for varying choices of (m, r). Each trial is declared successful if the estimate $\hat{\mathbf{\Sigma}}$ satisfies $\|\hat{\mathbf{\Sigma}} - \mathbf{\Sigma}\|_{\mathrm{F}} / \|\mathbf{\Sigma}\|_{\mathrm{F}} < 10^{-3}$. The empirical success rate is calculated by averaging over 50 Monte Carlo trials, and is reflected by the color of each cell. While there are in total r degrees of freedom, our algorithm exhibits approximately linear phase transition curve, which confirms our theoretical prediction in the absence of noise.

Chapter 6

Summary and Future Directions

The central theme of this thesis has been to explore the interplay between the signal structures and reduced-rate data acquisition, particularly in the high-dimensional realm. We end this thesis with a summary of our findings as well as several future directions.

6.1 Summary of Contributions

Capacity under Channel-Optimized Sampling

We have characterized sampled channel capacity as a function of sampling rate for the class of periodic sampling methods, thereby forming a new connection between sampling theory and information theory. We show how the capacity of a sampled analog channel is affected by reduced sampling rate and identify optimal sampling structures for several canonical sampling methods, which exploit structure in the sampling design. These results also indicate that capacity is not always monotonic in sampling rate, and illuminate an intriguing connection between MIMO channel capacity and capacity of undersampled analog channels. The capacity optimizing sampling structures are shown to extract the frequency components with the highest SNRs from each aliased set, and hence suppress aliasing and out-of-band noise.

When we go beyond periodic sampling structures, we characterize the maximum achievable information rate for a general class of time-preserving nonuniform sampling methods under a sampling rate constraint. It is shown that nonuniformly spaced sampling sets, while requiring fairly complicated reconstruction algorithms, do not provide any capacity gain in the presence of perfect CSI. Encouragingly, filterbank sampling with varied sampling rates on different branches, or a single branch of sampling with modulation and filtering, are sufficient to achieve the sampled channel capacity. In addition, both strategies suppress aliasing effects. In terms of maximizing capacity, there is no need to employ irregular sampling sets that are more complicated to implement in practical hardware systems. The resulting sampled capacity is shown to be monotonically increasing in sampling rate.

In summary, our work establishes a framework for using the information-theoretic metric of capacity to optimize sampling structures, offering a different angle from traditional design of sampling methods based on other performance metrics.

Capacity under Channel-Independent Sampling

We have investigated optimal universal sampling design from a capacity perspective. In order to evaluate the loss due to universal sub-Nyquist sampling design, we introduce the notion of sampled capacity loss relative to Nyquist-rate capacity, and characterize overall robustness of the sampling design in terms of the minimax capacity loss metric. Specifically, we have determined the fundamental minimax limit on the sampled capacity loss achievable by a class of channel-blind periodic sampling systems. This minimax limit turns out to be a constant that only depends on the band sparsity ratio and undersampling factor, modulo a residual term that vanishes in the SNR and a large number of subbands. Our results demonstrate that with exponentially high probability, random sampling is minimax in terms of a universal sampler design. This highlights the power of random sampling methods in the channel-blind design. In addition, our results extend to discrete-time counterparts without difficulty.

It would be interesting to extend this framework to situations beyond compound multiband channels, and our notion of sampled capacity loss will be useful in evaluating the robustness for these scenarios. Our framework and results may also be appropriate for other channels with state where sparsity exists in other transform domains. In addition, when it comes to multiple access channels or random access channels [20], it remains to be seen how to find a channel-blind sampler that is robust for the entire capacity region.

Spectral Compressed Sensing

We present an efficient algorithm to estimate a spectrally sparse signal from its partial time-domain samples that does not require prior knowledge on the model order, which poses spectral compressed sensing as a low-rank Hankel structured matrix completion problem. Under mild incoherence conditions, our algorithm enables recovery of the multi-dimensional unknown frequencies with infinite precision, which addresses the basis mismatch issue that arises in conventional CS paradigms. To the best of our knowledge, our result on Hankel matrix completion is also the first theoretical guarantee that is close to the information-theoretical limit (up to some logarithmic factor).

Our results are based on uniform random observation models. In particular, this thesis considers directly taking a random subset of the time domain samples, it is also possible to take a random set of linear mixtures of the time domain samples, as in the renowned CS setting [23]. This again can be translated into taking linear measurements of the low-rank K-fold Hankel matrix, given as $\boldsymbol{y} = \mathcal{B}(\boldsymbol{X}_{\mathrm{e}})$. Unfortunately, due to the Hankel structures, it is not clear whether $\mathcal{B}$ exhibits the approximate isometry property. Nonetheless, the technique developed in this thesis can be extended without difficulty to analyze linear measurements.

Covariance Estimation

We investigate covariance estimation under a quadratic sampling model. This sampling model acts as an effective signal processing method for real-time data with limited processing power and memory at the sensor side, and subsumes many sampling strategies where we can only obtain magnitude or energy samples. Three of the most popular covariance structures (i.e. low rank, Toeplitz low-rank structure, and jointly sparse and rank-one structure) have been explored. Encouragingly, the same quadratic sampling schemes prove simultaneously effective for all these structures.

Our results indicate that covariance matrices under the above structural assumptions can be faithfully recovered from a small set of quadratic measurements and minimal storage, as long as the sensing vectors are i.i.d. drawn from sub-Gaussian distributions. The recovery can be achieved via efficient convex programming as soon as the number of stored measurements exceeds the information theoretic limit. We also observe universal recovery phenomena, in the sense that once the sensing vectors are chosen, all covariance matrices possessing the presumed structure can be recovered. Our results highlight the stability and robustness of the convex program in the presence of noise and imperfect structural assumptions. The performance guarantees for low-rank, sparse and jointly rank-one and sparse models are established via a novel notion of a mixed-norm restricted isometry property (RIP-ℓ_2/ℓ_1). In addition, our innovation includes a systematic approach to analyze Toeplitz low-rank structure, which relies on RIP-ℓ_2/ℓ_2 under near-isotropic and bounded operators.

6.2 Future Directions

Duality between Channel Capacity and Distortion-Rate Function under Sub-Nyquist Sampling

Our work uncovers additional questions at the intersection of sampling theory and information theory from both channel coding and source coding perspectives. In particular, the fundamental rate-distortion function of Gaussian sources has recently been characterized under sub-Nyquist sampling with filtering [129], revealing that the alias suppressing sampler design also achieves the optimal rate distortion function in the source coding problem. This suggests an intriguing duality between the channel coding and source coding problems, namely, the optimal sub-Nyquist sampler design for one problem is identical to the optimal sampler for another problem, which might potentially extend to a very general class of scenarios including both channel-optimized and channel-blind sampling. Such duality relations, which might be established via the interplay among MMSE, channel capacity, and rate distortion functions, would be useful in deriving new results concerning sampled Gaussian channels or sources.

Information Loss Metrics for Robust or Universal Information Theory

Practical communication, storage, and data acquisition systems are typically comprised of two parts: 1) hardware components, which are not easily configurable and hence cannot be easily changed during operations; 2) software (reconfigurable) components, which have much higher flexibility. This dichotomy raises interesting design tradeoffs in time-varying environments, since the hardware components are more constrained and cannot be adapted to all channel / system realizations. The information loss metrics proposed in this thesis (i.e. the minimax information loss metrics) become a natural criterion in measuring the robustness for data transmission and can be adpated to the storage problem as well.

In fact, even in the software components like encoders and decoders, universal designs that support near-optimal information rate would be desirable for practical systems, especially when accurate channel or source state information is not available. That said, one would wish to adopt a single codebook design that is capable of accommodating composite time-varying channels or sources. In some hardware-constrained scenarios where there is no coding strategy that achieves optimal information rate uniformly across all possible states, the minimax (or Bayesian) information loss metrics would be of great importance in comparing the overall performances of different system designs.

Super-Resolution via Structured Matrix Completion

Our EMaC algorithm succeeds by exploiting the underlying connection between spectral sparsity and low-rank Hankel structure. It remains to be seen whether the success of EMaC is a result of the random sampling pattern or if it is applicable to a broader class of deterministic sampling patterns. For instance, if the observation samples cover only the low-end time domain, one could explore if it is provably effective to apply EMaC to extrapolate high-frequency components in a faithful manner. In addition, if the underlying frequency spikes are not sufficiently separable, one could investigate how the performance of EMaC will be affected. This might uncover a deeper connection between spectral separability and the incoherence properties of the associated structured matrix.

Estimation of Inverse Covariance Matrices

Another covariance structure of great practical interest is sparsity in the graphical model underlying the *inverse* covariance matrix. Specifically, many complex systems or large data streams can be captured through succinct cross-variable dependencies using graphical models. It is well known that when the data are jointly Gaussian, the inverse covariance matrix encodes the conditional independence, which is typically sparse. Given limited memory, one would wish to exploit the sparsity in such conditional independence in order to enable reliable covariance estimation from highly incomplete measurements. It remains to be seen whether the quadratic measurement scheme in (1.1) is effective in preserving the sparse graphical model representation and, more importantly, how to design efficient and provably stable recovery paradigms in inferring the underlying graphical structure.

Appendix A

Proof of Theorems and Lemmas in Chapter 2

A.1 Proof of Corollary 2.2

Corollary 2.2 immediately follows from the following proposition.

Proposition A.1. *The kth largest eigenvalue λ_k of the positive semidefinite matrix $\tilde{\boldsymbol{F}}_s \boldsymbol{F}_h \boldsymbol{F}_h^* \tilde{\boldsymbol{F}}_s^*$ is bounded by*

$$0 \le \lambda_k \le \lambda_k \left(\boldsymbol{F}_h \boldsymbol{F}_h^* \right), \quad 1 \le k \le M. \tag{A.1}$$

These upper bounds can be attained simultaneously by the filter (2.13).

Proof. Recall that at a given f, $\boldsymbol{F}_h$ is an infinite diagonal matrix satisfying $\left(\boldsymbol{F}_h\right)_{l,l} = H\left(f - \frac{l f_s}{M}\right)$ for all $l \in \mathbb{Z}$, and that $\tilde{\boldsymbol{F}}_s = \left(\boldsymbol{F}_s \boldsymbol{F}_s^*\right)^{-\frac{1}{2}} \boldsymbol{F}_s$. Hence, $\tilde{\boldsymbol{F}}_s \boldsymbol{F}_h \boldsymbol{F}_h^* \tilde{\boldsymbol{F}}_s^*$ is an $M \times M$ dimensional matrix. We observe that

$$\tilde{\boldsymbol{F}}_s \tilde{\boldsymbol{F}}_s^* = \left(\boldsymbol{F}_s \boldsymbol{F}_s^*\right)^{-\frac{1}{2}} \boldsymbol{F}_s \boldsymbol{F}_s^* \left(\boldsymbol{F}_s \boldsymbol{F}_s^*\right)^{-\frac{1}{2}} = \boldsymbol{I}, \tag{A.2}$$

which indicates that the rows of $\tilde{\boldsymbol{F}}_s$ are orthonormal. Hence, the operator norm of $\tilde{\boldsymbol{F}}_s$ is no larger than 1, which leads to

$$\lambda_1\left(\tilde{\boldsymbol{F}}_s\boldsymbol{F}_h\boldsymbol{F}_h^*\tilde{\boldsymbol{F}}_s^*\right)=\left\|\tilde{\boldsymbol{F}}_s\boldsymbol{F}_h\right\|_2^2\leq\left\|\boldsymbol{F}_h\right\|_2^2=\lambda_1\left(\boldsymbol{F}_h\boldsymbol{F}_h^*\right).$$

Denote by $\boldsymbol{e}_k$ the kth standard basis vector. We introduce the index set $\{i_1,i_2,\cdots,i_M\}$ such that $\boldsymbol{e}_{i_k}$ $(1\leq k\leq M)$ is the eigenvector associated with the kth largest eigenvalues of the diagonal matrix $\boldsymbol{F}_h\boldsymbol{F}_h^*$.

Suppose that $\boldsymbol{v}_k$ is the eigenvector associated with the kth largest eigenvalue λ_k of $\tilde{\boldsymbol{F}}_s\boldsymbol{F}_h\boldsymbol{F}_h^*\tilde{\boldsymbol{F}}_s^*$, and denote by $\left(\tilde{\boldsymbol{F}}_s\right)_k$ the kth *column* of $\tilde{\boldsymbol{F}}_s$. Since $\tilde{\boldsymbol{F}}_s\boldsymbol{F}_h\boldsymbol{F}_h^*\tilde{\boldsymbol{F}}_s^*$ is Hermitian positive semidefinite, its eigendecomposition yields an orthogonal basis of eigenvectors. Observe that $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ spans a k-dimensional space and that $\left\{\left(\tilde{\boldsymbol{F}}_s\right)_j,1\leq j\leq k-1\right\}$ spans a subspace of dimension no more than $k-1$. For any $k\geq 2$, there exists k scalars $a_1,\cdots,a_k$ such that

$$\sum_{i=1}^{k}a_i\boldsymbol{v}_i\perp\left\{\left(\tilde{\boldsymbol{F}}_s\right)_{i_j},1\leq j\leq k-1\right\}\text{ and }\sum_{i=1}^{k}a_i\boldsymbol{v}_i\neq 0. \tag{A.3}$$

This allows us to define the following unit vector

$$\tilde{\boldsymbol{v}}_k\triangleq\sum_{i=1}^{k}\frac{a_i}{\sqrt{\sum_{j=1}^{k}\left|a_j\right|^2}}\boldsymbol{v}_i, \tag{A.4}$$

which is orthogonal to $\left\{\left(\tilde{\boldsymbol{F}}_s\right)_j,1\leq j\leq k-1\right\}$. We observe that

$$\begin{aligned}\left\|\tilde{\boldsymbol{F}}_s\boldsymbol{F}_h\boldsymbol{F}_h^*\tilde{\boldsymbol{F}}_s^*\tilde{\boldsymbol{v}}_k\right\|_2^2&=\left\|\sum_{i=1}^{k}\frac{a_i}{\sqrt{\sum_{j=1}^{k}\left|a_j\right|^2}}\tilde{\boldsymbol{F}}_s\boldsymbol{F}_h\boldsymbol{F}_h^*\tilde{\boldsymbol{F}}_s^*\boldsymbol{v}_i\right\|_2^2=\left\|\sum_{i=1}^{k}\frac{a_i\lambda_i}{\sqrt{\sum_{j=1}^{k}\left|a_j\right|^2}}\boldsymbol{v}_i\right\|_2^2\\&=\sum_{i=1}^{k}\frac{\lambda_i^2\left|a_i\right|^2}{\sum_{j=1}^{k}\left|a_j\right|^2}\geq\lambda_k^2.\end{aligned} \tag{A.5}$$

Define $\boldsymbol{u}_k := \tilde{\boldsymbol{F}}_s^* \tilde{\boldsymbol{v}}_k$. From (A.3) we can see that $(\boldsymbol{u}_k)_i = \left\langle \left(\tilde{\boldsymbol{F}}_s\right)_i, \tilde{\boldsymbol{v}}_i \right\rangle = 0$ holds for all $i \in \{i_1, i_2, \cdots, i_{k-1}\}$. In other words, $\boldsymbol{u}_k \perp \left\{\boldsymbol{e}_{i_1}, \cdots, \boldsymbol{e}_{i_{k-1}}\right\}$. This further implies that

$$\begin{aligned}\lambda_k^2 &\leq \left\|\tilde{\boldsymbol{F}}_s \boldsymbol{F}_h \boldsymbol{F}_h^* \tilde{\boldsymbol{F}}_s^* \tilde{\boldsymbol{v}}_k\right\|_2^2 \leq \left\|\tilde{\mathbf{F}}_s\right\|_2^2 \left\|\boldsymbol{F}_h \boldsymbol{F}_h^* \tilde{\boldsymbol{F}}_s^* \tilde{\boldsymbol{v}}_k\right\|_2^2 \leq \left\|\mathbf{F}_h \mathbf{F}_h^* \boldsymbol{u}_k\right\|_2^2 \\ &\leq \sup_{\boldsymbol{x} \perp \mathrm{span}\left\{\boldsymbol{e}_{i_1}, \cdots, \boldsymbol{e}_{i_{k-1}}\right\}} \left\|\boldsymbol{F}_h \boldsymbol{F}_h^* \boldsymbol{x}\right\|_2^2 = \lambda_k^2 \left(\boldsymbol{F}_h \boldsymbol{F}_h^*\right)\end{aligned}$$

by observing that $\boldsymbol{F}_h \boldsymbol{F}_h^*$ is a diagonal matrix.

Setting

$$S_k\left(f - \frac{l f_s}{M}\right) = \begin{cases} 1, & \text{if } \left|H\left(f - \frac{l f_s}{M}\right)\right|^2 = \lambda_k\left(\boldsymbol{F}_h(f) \boldsymbol{F}_h^*(f)\right), \\ 0, & \text{otherwise,} \end{cases}$$

yields $\tilde{\boldsymbol{F}}_s = \boldsymbol{F}_s$ and hence $\tilde{\boldsymbol{F}}_s \boldsymbol{F}_h \boldsymbol{F}_h^* \tilde{\boldsymbol{F}}_s$ is a diagonal matrix such that

$$\left(\tilde{\boldsymbol{F}}_s \boldsymbol{F}_h \boldsymbol{F}_h^* \tilde{\boldsymbol{F}}_s^*\right)_{k,k} = \lambda_k\left(\boldsymbol{F}_h \boldsymbol{F}_h^*\right). \tag{A.6}$$

Apparently, this choice of $S_k(f)$ allows the upper bounds

$$\lambda_k\left(\tilde{\boldsymbol{F}}_s \boldsymbol{F}_h \boldsymbol{F}_h^* \tilde{\boldsymbol{F}}_s^*\right) = \lambda_k\left(\boldsymbol{F}_h \boldsymbol{F}_h^*\right), \quad \forall 1 \leq k \leq M \tag{A.7}$$

to be attained simultaneously. □

By picking the M frequencies with the highest SNR from each aliased set $\{f - l f_s/M \mid l \in \mathbb{Z}\}$, we achieve $\lambda_k = \lambda_k\left(\boldsymbol{F}_h \boldsymbol{F}_h^*\right)$ and hence the maximum capacity.

A.2 Proof of Proposition 2.1

Denote by $y^k(t)$ the analog signal after passing through the k^{th} prefilter prior to ideal sampling. When both the input signal $x(t)$ and the noise $\eta(t)$ are Gaussian, the MMSE estimator of $x(t)$ from samples $\left\{y^k[n] \mid 1 \leq k \leq M\right\}$ is linear. Recall that

$\tilde{T}_s = MT_s$ and $\tilde{f}_s = f_s/M$. A linear estimator of $x(t)$ from $\boldsymbol{y}[n]$ can be given as

$$\hat{x}(t) = \sum_{k\in\mathbb{Z}} \boldsymbol{g}^{\top}(t - k\tilde{T}_s) \cdot \boldsymbol{y}(k\tilde{T}_s), \tag{A.8}$$

where we use the vector form $\boldsymbol{g}(t) = [g^1(t), \cdots, g^M(t)]^{\top}$ and $\boldsymbol{y}(t) = [y^1(t), \cdots, y^M(t)]^{\top}$ for notational simplicity. Here, $g^l(t)$ denotes the interpolation function operating upon the samples in the l^{th} branch. We propose to find the optimal estimator $\boldsymbol{g}(t)$ that minimizes the mean square estimation error $\mathbb{E}\left[|x(t) - \hat{x}(t)|^2\right]$ for some t.

From the orthogonality principle, the MMSE estimate $\hat{x}(t)$ obeys

$$\mathbb{E}\left[x(t)\boldsymbol{y}^*(l\tilde{T}_s)\right] = \mathbb{E}\left[\hat{x}(t)\boldsymbol{y}^*(l\tilde{T}_s)\right], \quad \forall l \in \mathbb{Z}. \tag{A.9}$$

Since $x(t)$ and $\eta(t)$ are both stationary Gaussian processes, we can define $\boldsymbol{R}_{XY}(\tau) := \mathbb{E}\left[x(t)\boldsymbol{y}^*(t-\tau)\right]$ to be the cross correlation function between $x(t)$ and $\boldsymbol{y}(t)$, and $\boldsymbol{R}_Y(\tau) := \mathbb{E}\left[\boldsymbol{y}(t)\boldsymbol{y}^*(t-\tau)\right]$ the autocorrelation function of $\boldsymbol{y}(t)$. Plugging (A.8) into (A.9) leads to the following relation

$$\boldsymbol{R}_{XY}\left(t - l\tilde{T}_s\right) = \sum_{k\in\mathbb{Z}} \boldsymbol{g}^{\top}\left(t - k\tilde{T}_s\right)\boldsymbol{R}_Y\left(k\tilde{T}_s - l\tilde{T}_s\right).$$

Replacing t by $t + l\tilde{T}_s$, we can equivalently express it as

$$\boldsymbol{R}_{XY}(t) = \sum_{k\in\mathbb{Z}} \boldsymbol{g}^{\top}\left(t + l\tilde{T}_s - k\tilde{T}_s\right)\boldsymbol{R}_Y\left(k\tilde{T}_s - l\tilde{T}_s\right) = \sum_{l\in\mathbb{Z}} \boldsymbol{g}^{\top}\left(t - l\tilde{T}_s\right)\boldsymbol{R}_Y\left(l\tilde{T}_s\right), \tag{A.10}$$

which is equivalent to the convolution of $\boldsymbol{g}(t)$ and $\boldsymbol{R}_Y(t) \cdot \sum_{l\in\mathbb{Z}} \delta\left(t - l\tilde{T}_s\right)$.

Let $\mathcal{F}(\cdot)$ denote Fourier transform operator. Define the cross spectral density $\boldsymbol{S}_{XY}(f) := \mathcal{F}\left(\boldsymbol{R}_{XY}(t)\right)$ and $\mathbf{S}_Y(f) = \mathcal{F}\left(\boldsymbol{R}_Y\left(t\right)\right)$. By taking the Fourier transform on both sides of (A.10) , we have

$$\boldsymbol{S}_{XY}(f) = \boldsymbol{G}(f)\mathcal{F}\left(\boldsymbol{R}_Y(\tau)\sum_{l\in\mathbb{Z}}\delta(\tau - l\tilde{T}_s)\right),$$

which immediately yields that $\forall f \in \left[-\tilde{f}_s/2, \tilde{f}_s/2\right]$,

$$\boldsymbol{G}(f) = \boldsymbol{S}_{XY}(f)\left[\mathcal{F}\left(\boldsymbol{R}_Y(\tau)\sum_{l\in\mathbb{Z}}\delta(\tau - l\tilde{T}_s)\right)\right]^{-1} = \boldsymbol{S}_{XY}(f)\left(\sum_{l\in\mathbb{Z}}\boldsymbol{S}_Y\left(f - l\tilde{f}_s\right)\right)^{-1}.$$

Since the noise $\eta(t)$ is independent of $x(t)$, the cross correlation function $\boldsymbol{R}_{XY}(t)$ is

$$\boldsymbol{R}_{XY}(\tau) = \mathbb{E}\left[x(t+\tau)\cdot\left\{(s_1 * h * x)^*(t), \cdots, (s_M * h * x)^*(t)\right\}\right],$$

which allows the cross spectral density to be derived as

$$\boldsymbol{S}_{XY}(f) = H^*(f)\mathcal{S}_X(f)\left[S_1^*(f), \cdots, S_M^*(f)\right]. \tag{A.11}$$

Additionally, the power spectral density of $\boldsymbol{y}(t)$ can be represented by the following $M \times M$ matrix

$$\boldsymbol{S}_Y(f) = \left(|H(f)|^2\mathcal{S}_X(f) + \mathcal{S}_\eta(f)\right)\boldsymbol{S}(f)\boldsymbol{S}^*(f), \tag{A.12}$$

where $\boldsymbol{S}(f) = [S_1(f), \cdots, S_m(f)]^\top$.

Define

$$\boldsymbol{K}(f) := \sum_{l\in\mathbb{Z}}\left(|H(f - lf_s)|^2\mathcal{S}_X(f - lf_s) + \mathcal{N}(f - lf_s)\right)\boldsymbol{S}(f - lf_s)\boldsymbol{S}^*(f - lf_s).$$

The Wiener-Hopf linear reconstruction filter can now be written as

$$\mathbf{G}(f) = H^*(f)\mathcal{S}_X(f)\boldsymbol{S}^*(f)\boldsymbol{K}^{-1}(f)$$

Define $R_X(\tau) = \mathbb{E}\left[x(t)x^*(t-\tau)\right]$. Since $\int_{-\infty}^{\infty} \mathcal{S}_X(f)\mathrm{d}f = R_X(0)$, the resulting MSE is

$$\begin{aligned}
\xi(t) &= \mathbb{E}\left[|x(t)|^2\right] - \mathbb{E}\left[|\hat{x}(t)|^2\right] = \mathbb{E}\left[|x(t)|^2\right] - \mathbb{E}\left[x(t)\hat{x}^*(t)\right] \\
&= R_X(0) - \mathbb{E}\left[x(t)\left(\sum_{l\in\mathbb{Z}} \boldsymbol{g}^\top(t-lT_s)\boldsymbol{y}(lT_s)\right)^*\right] \\
&= R_X\left(0\right) - \sum_{l\in\mathbb{Z}} \boldsymbol{R}_{XY}(t-lT_s)\boldsymbol{g}(t-lT_s).
\end{aligned}$$

Since $\mathcal{F}\left(\boldsymbol{g}(-t)\right) = \left(\boldsymbol{G}^*(f)\right)^\top$ and $\boldsymbol{S}_{XY} = H^*(f)\mathcal{S}_X(f)\boldsymbol{S}^*(f)$, Parseval's identity implies that

$$\begin{aligned}
\xi(t) &= \int_{-\infty}^{\infty} \left[\mathcal{S}_X(f) - \boldsymbol{G}^*(f)\boldsymbol{S}_{XY}^\top\right]\mathrm{d}f \\
&= \int_{-\infty}^{\infty} \left[\mathcal{S}_X(f) - \left|H(f)\mathcal{S}_X(f)\right|^2 \boldsymbol{S}^*(f)\boldsymbol{K}^{-1}(f)\boldsymbol{S}(f)\right]\mathrm{d}f \\
&= \int_{-\tilde{f}_s/2}^{\tilde{f}_s/2} \left[\sum_{l=-\infty}^{\infty} \mathcal{S}_X(f-l\tilde{f}_s) - \tilde{T}_s\boldsymbol{V}_\zeta^\top(f,\tilde{f}_s)\cdot\mathbf{1}\right]\mathrm{d}f.
\end{aligned}$$

Suppose that we impose power constraints $\sum_{l\in\mathbb{Z}} \mathcal{S}_X(f-l\tilde{f}_s) = P(f)$, and define $\zeta(f) := \left|H(f)\mathcal{S}_X(f)\right|^2 \boldsymbol{S}^*(f)\boldsymbol{K}^{-1}(f)\boldsymbol{S}(f)$. For a given input process $x(t)$, the problem of finding the optimal prefilter $\boldsymbol{S}(f)$ that minimizes MSE then becomes

$$\underset{\{S(f-lf_s),l\in\mathbb{Z}\}}{\text{maximize}} \quad \boldsymbol{V}_\zeta^\top(f,\tilde{f}_s)\cdot\mathbf{1},$$

where the objective function can be alternatively rewritten in matrix form

$$\mathrm{tr}\left\{\boldsymbol{F}_X^{\frac{1}{2}}\boldsymbol{F}_h^*\boldsymbol{F}_s^*\left(\boldsymbol{F}_s\left(\boldsymbol{F}_h\boldsymbol{F}_h^* + \boldsymbol{F}_\eta\right)\boldsymbol{F}_s^*\right)^{-1}\boldsymbol{F}_s\boldsymbol{F}_h\boldsymbol{F}_X^{\frac{1}{2}}\right\} \tag{A.13}$$

Here $\boldsymbol{F}_X$ and $\boldsymbol{F}_\eta$ are diagonal matrices such that $(\boldsymbol{F}_X)_{l,l} = \mathcal{S}_X(f - lf_s)$ and $(\boldsymbol{F}_\eta)_{l,l} = \mathcal{S}_\eta(f + kf_s)$. We observe that

$$
\begin{aligned}
&\operatorname{tr}\left\{\boldsymbol{F}_X^{\frac{1}{2}}\boldsymbol{F}_h^*\boldsymbol{F}_s^*\left(\boldsymbol{F}_s\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)\boldsymbol{F}_s^*\right)^{-1}\boldsymbol{F}_s\boldsymbol{F}_h\boldsymbol{F}_X^{\frac{1}{2}}\right\}\\
&\quad=\operatorname{tr}\left\{\left(\boldsymbol{F}_s\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)\boldsymbol{F}_s^*\right)^{-1}\boldsymbol{F}_s\boldsymbol{F}_h\boldsymbol{F}_X\boldsymbol{F}_h^*\boldsymbol{F}_s^*\right\}\\
&\quad\overset{(a)}{=}\operatorname{tr}\left\{\left(\boldsymbol{Y}\boldsymbol{Y}^*\right)^{-1}\boldsymbol{Y}\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)^{-\frac{1}{2}}\boldsymbol{F}_h\boldsymbol{F}_X\boldsymbol{F}_h^*\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)^{-\frac{1}{2}}\boldsymbol{Y}^*\right\}\\
&\quad\overset{(b)}{=}\operatorname{tr}\left\{\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)^{-1}\boldsymbol{F}_h\boldsymbol{F}_X\boldsymbol{F}_h^*\boldsymbol{Y}^*\left(\boldsymbol{Y}\boldsymbol{Y}^*\right)^{-1}\boldsymbol{Y}\right\}\\
&\quad\overset{(c)}{=}\operatorname{tr}\left\{\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)^{-1}\boldsymbol{F}_h\boldsymbol{F}_X\boldsymbol{F}_h^*\tilde{\boldsymbol{Y}}^*\tilde{\boldsymbol{Y}}\right\}\\
&\quad\overset{(d)}{\leq}\sup_{\boldsymbol{Z}\cdot\boldsymbol{Z}^*=\boldsymbol{I}_M}\operatorname{tr}\left\{\boldsymbol{Z}\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)^{-1}\boldsymbol{F}_h\boldsymbol{F}_X\boldsymbol{F}_h^*\boldsymbol{Z}^*\right\}=\sum_{i=1}^{M}\lambda_i(\boldsymbol{D}),
\end{aligned}
$$

where (a) follows by introducing $\boldsymbol{Y} := \boldsymbol{F}_s\left(\boldsymbol{F}_h\boldsymbol{F}_h^*+\boldsymbol{F}_\eta\right)^{\frac{1}{2}}$, (b) follows from the fact that $\boldsymbol{F}_h$, $\boldsymbol{F}_X$, $\boldsymbol{F}_\eta$ are all diagonal matrices, (c) follows by introducing $\tilde{\boldsymbol{Y}} = \left(\boldsymbol{Y}\boldsymbol{Y}^*\right)^{-\frac{1}{2}}\boldsymbol{Y}$, and (d) follows by observing that $\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^* = \left(\boldsymbol{Y}\boldsymbol{Y}^*\right)^{-\frac{1}{2}}\boldsymbol{Y}\boldsymbol{Y}^*\left(\boldsymbol{Y}\boldsymbol{Y}^*\right)^{-\frac{1}{2}} = \boldsymbol{I}$. Here, $\boldsymbol{D}$ is an infinite diagonal matrix such that $\boldsymbol{D}_{l,l} = \frac{|H(f-lf_s)|^2\mathcal{S}_X(f-lf_s)}{|H(f-lf_s)|^2\mathcal{S}_X(f-lf_s)+\mathcal{S}_\eta(f-lf_s)}$. In other words, the upper bound is the sum of the M largest $\boldsymbol{D}_{i,i}$ which are associated with M frequency points of highest SNR $\frac{|H(f+lf_s)|^2\mathcal{S}_X(f+lf_s)}{\mathcal{S}_\eta(f+lf_s)}$.

Therefore, when restricted to the set of all permutations of $\{\mathcal{S}_X(f), \mathcal{S}_X(f \pm f_s), \cdots\}$, the minimum MSE is achieved when assigning the M largest $\mathcal{S}_X(f + lf_s)$ to the M branches with the largest SNR. In this case, the corresponding optimal filter can be chosen such that

$$
S_k(f - lf_s) = \begin{cases} 1, & \text{if } l = \hat{k} \\ 0, & \text{otherwise.} \end{cases} \tag{A.14}
$$

where $\hat{k}$ is the index of the k^{th} largest element in $\left\{|H(f - lf_s)|^2 / \mathcal{S}_\eta(f - lf_s) : l \in \mathbb{Z}\right\}$.

A.3 Proof of Theorem 2.5

For simplicity of presentation, we assume throughout that the noise is white, i.e. $S_\eta(f) \equiv 1$. In fact, under the assumption (2.22), we can always split the channel filter $H(f)$ into two parts with respective frequency response $H(f)/\sqrt{\mathcal{S}_\eta(f)}$ and $\sqrt{\mathcal{S}_\eta(f)}$. Since the colored noise is equivalent to a white Gaussian noise passed through a filter with transfer function $\sqrt{\mathcal{S}_\eta(f)}$, the original system can be redrawn as in Fig. A.1. The filter with frequency response $\sqrt{\mathcal{S}_\eta(f)}$ can then be incorporated into the preprocessing system to generate a new time-preserving preprocessor.

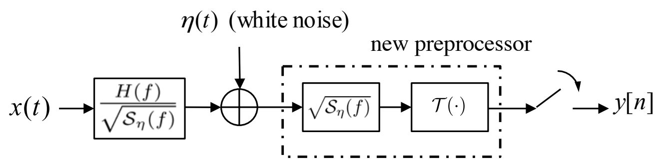


Figure A.1: Equivalent representation of sampling systems in the presence of colored noise.

Since any multibranch sampling system can be easily converted into a single branch sampling system without loss of information, we restrict our proof to the class of single branch sampling systems.

A.3.1 General Upper Bound: Finite-duration $h(t)$

In this subsection, we focus on the channel whose impulse response is of finite duration $2L_0$, i.e.

$$h(t) = 0, \quad \forall t \ (|t| > L_0).$$

Our goal is to prove that the capacity upper bound (2.23) holds for this type of channel.

For any transmission block of duration $2T$, we call the transmit signal $x(t)$ over this block a codeword (or symbol) of code length $2T$. The information conveyed through such finite-duration codewords can be bounded via certain analog channel capacity, as long as we can preclude inter-symbol interference. The key idea here is to separate consecutive codewords with a guard zone with sufficient length and then use capacity-achieving strategies separately for each codeword. When the code length $2T$ is sufficiently large, the transmission time wasted on the guard zones becomes negligible, which in turn allows us to approach the true capacity arbitrarily well. The detailed analysis proceeds as follows.

Step 1. Consider an input $x(t)$ that is constrained to the interval $[-T, T]$. Since $h(t)$ is of finite duration $2L_0$, the channel output $r(t) = h(t) * x(t) + \eta(t)$ will be affected by the input only when $t \in [-T - L_0, T + L_0]$. Define a *window operator* and its complement operator such that

$$w_T(f(t)) = \begin{cases} f(t), & \text{if } |t| \leq T + L_0, \\ 0, & \text{else}; \end{cases} \tag{A.15}$$

$$w_T^{\perp}(f(t)) = \begin{cases} 0, & \text{if } |t| \leq T + L_0, \\ f(t), & \text{else}. \end{cases} \tag{A.16}$$

Then for any linear sampling operator $\mathcal{P}$ with impulse response $q(t, \tau)$, the sampled output is $\mathcal{P}(r(t)) = \mathcal{P}(w_T(r(t))) + \mathcal{P}\left(w_T^{\perp}(r(t))\right)$. One can easily observe that the component $\mathcal{P}\left(w_T^{\perp}(r(t))\right)$ contains no information about $x(t)$, and is statistically *independent* of $\mathcal{P}(w_T(r(t)))$ due to the whiteness assumption of the noise. In other words, the sampling input outside the interval $[-T - L_0, T + L_0]$ does not improve capacity at all. Consequently, it suffices to restrict attention to the class of sampling systems whose system input is constrained to the interval $[-T - L_0, T + L_0]$.

Step 2. Construct a periodization of the above sampled channel model with finite input duration. Set the impulse response $q_{T+L_0}^{\mathrm{p}}(t, \tau)$ of the preprocessor of the periodized sampling system to be a periodic extension of $q(t, \tau)$ in the block $[-T - L_0, T + L_0] \times [-T, T]$. Specifically, if $\tau = k \cdot 2(T + L_0) + \tau_{\mathrm{r}}$ for some $k \in \mathbb{Z}$ and $\tau_{\mathrm{r}} \in [-T - L_0, T + L_0]$, then

$$q_{T+L_0}^{\mathrm{p}}(t, \tau) = \begin{cases} q\left(t - 2k(T + L_0), \tau_{\mathrm{r}}\right), & \text{if } |t - 2k(T + L_0)| \leq T + L_0, \\ 0, & \text{else}. \end{cases} \tag{A.17}$$

Apparently, $q_{T+L_0}^{\mathrm{p}}(t, \tau)$ corresponds to a periodic preprocessing system with period $2(T + L_0)$.

Suppose without loss of generality that the indices of the sample times that fall in $[-T-L_0, T+L_0]$ are $0, 1, \cdots, K-1$, i.e. $\{k \mid t_k \in [-T-L_0, T+L_0]\} = \{0, 1, \cdots, K-1\}$. We can then set the sampling set $\Lambda^{\mathrm{p}}_{T+L_0}$ of the periodized system such that for any sampling time $t_k \in \Lambda^{\mathrm{p}}_{T+L_0}$, we have

$$t_k = t_{k \bmod K} + 2(T+L_0) \cdot \left\lfloor \frac{k}{K} \right\rfloor, \tag{A.18}$$

where $\lfloor x \rfloor \overset{\Delta}{=} \max\{n \mid n \in \mathbb{Z}, n \leq x\}$. Clearly, this forms a periodic sampling set with period $2(T+L_0)$. The definition of Beurling density ensures that for any $\epsilon > 0$, there exists a T_D such that for every $T > T_D$,

$$f_s - \epsilon \leq D\left(\Lambda^{\mathrm{p}}_{T+L_0}\right) \leq f_s + \epsilon.$$

Due to the finite-duration assumption of $h(t)$, our construction (A.17) guarantees that the input $x(t)$ within time interval $[2k(T+L_0) - T, 2k(T+L_0) + T]$ will only affect the sampled output at the *kth time block* $[(2k-1)(T+L_0), (2k+1)(T+L_0)]$, as illustrated in Fig. A.2. Since the noise $\eta(t)$ is assumed to be white, the noise components across different time blocks are independent. In fact, the intervals $[2k(T+L_0)+T, (2k+1)(T+L_0)-T]$ $(k \in \mathbb{Z})$ act effectively as *guard zones* in order to avoid leakage of signals across different time blocks.

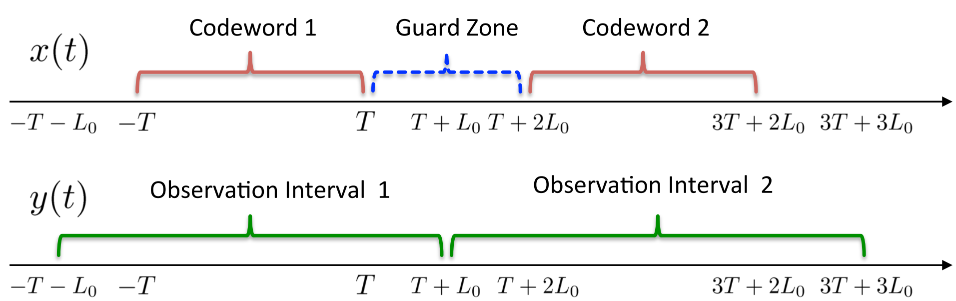


Figure A.2: The codewords of duration $2T$ are separated by guard zones of duration $2L_0$. There is no inter-symbol interference among different observation intervals.

Based on the above argument, we can separate codewords of duration $2T$ in $[2k(T+L_0)+T,(2k+1)(T+L_0)-T]$ ($k\in\mathbb{Z}$) on the analog channel by a guard zone $2L_0$ (as illustrated in Fig. A.2). The ratio of guard space to the length of the time block vanishes as $T\rightarrow\infty$, and there is no intersymbol interference under the new system we construct. By our capacity definition, for any $\delta>0$, there exists a T_0 such that $\forall T>T_0$, we have

$$\frac{T+L_0}{T}<1+\delta,\quad \text{and}\quad \frac{T}{T+L_0}>1-\delta.$$

Consequently,

$$C_T^{\mathcal{P}}\left(P\right)\overset{\text{(i)}}{\leq}\frac{T+L_0}{T}C_{\text{p}}^{\mathcal{P}}\left(\frac{T}{T+L_0}P\right)\leq\left(1+\delta\right)C_{\text{p}}^{\mathcal{P}}\left(\left(1-\delta\right)P\right),\tag{A.19}$$

where $C_{\text{p}}^{\mathcal{P}}$ denotes the capacity under our periodized sampling system. The inequality (i) follows from the following three arguments:

- $C_T^{\mathcal{P}}$ is the information rate when we observe the samples within the interval $[-T,T]$, which is smaller than the information rate, termed $\hat{C}_T^{\mathcal{P}}$, when we observe all samples within $[-T-L_0,T+L_0]$;
- $\hat{C}_T^{\mathcal{P}}$ is equivalent to the maximum information rate achievable by the periodized system, under the constraint that there is no input signal transmitted over the guard zones. Clearly, this rate will be smaller than the capacity without this transmission constraint, which is $\frac{T+L_0}{T}C_{\text{p}}^{\mathcal{P}}$. Here, the multiplication factor $\frac{T+L_0}{T}$ arises from the fact that we only use a portion $\frac{T}{T+L_0}$ of time for transmission;
- Since the total energy over each transmission block is PT and each guard zone has zero power, the average power allocated to the transmitted signal is $\frac{T}{T+L_0}P$.

We know from Corollary 2.3(b) that

$$C_{\text{p}}^{\mathcal{P}}\left(\left(1-\delta\right)P\right)\leq C_{\text{u}}\left(D\left(\Lambda_{T+L_0}^{\text{p}}\right),\left(1-\delta\right)P\right)\leq C_{\text{u}}\left(f_s+\epsilon,\left(1-\delta\right)P\right),\tag{A.20}$$

where the last inequality arises from observing that $C_{\mathrm{u}}(f_s, P)$ is monotonically non-decreasing in f_s and P. Putting (A.19) and (A.20) together yields

$$C_T^{\mathcal{P}}(P) \leq (1+\delta)\, C_{\mathrm{u}}(f_s+\epsilon, P) \tag{A.21}$$

as soon as $T > \max\{T_0, T_D\}$. Since ϵ and δ can be arbitrarily small, we have that

$$\limsup_{T\to\infty} C_T^{\mathcal{P}}(P) \leq C_{\mathrm{u}}(f_s, P)$$

when $h(t)$ is of finite duration and $\eta(t)$ is white.

A.3.2 General Upper Bound: Infinite-Duration $h(t)$

We now investigate the capacity bound when $h(t)$ is not time-limited. We would like to prove that for any given sampling system $\mathcal{P}$ and any $\epsilon > 0$, there exists T_1 such that for any $T > T_1$, one has

$$C_T^{\mathcal{P}} \leq C_{\mathrm{u}}(f_s, P) + \epsilon.$$

Our proof proceeds by comparing the original channel with a truncated channel whose channel response $\tilde{h}(f)$ satisfies

$$\tilde{h}(t) \triangleq \begin{cases} h(t), & \text{if } |t| \leq L_1, \\ 0, & \text{otherwise.} \end{cases}$$

Let $\xi > 0$ be an arbitrary small constant, and L_1 chosen such that

$$\int_{-\infty}^{-L_1} |h(t)|^2 \,\mathrm{d}t + \int_{L_1}^{\infty} |h(t)|^2 \,\mathrm{d}t \leq \xi. \tag{A.22}$$

We further constrain the input and the observed sampled output to the time interval $[-T, T]$. For both the original and truncated channel, the sampled noise is not white, which motivates us to first perform prewhitening.

Suppose without loss of generality that the sampled times within $[-T,T]$ are $\{t_i \mid 1 \leq i \leq K_T\}$. For convenience of notation, we introduce a linear operator $\hat{\mathcal{P}}_T$ associated with the sampling system such that

$$\hat{\mathcal{P}}_T(\hat{r}(t)) = [y[1], y[2], \cdots, y[K_T]],$$

where $\hat{r}(t) = g(x) * x(t) + \hat{\eta}(t)$ is the sampling system input, $\hat{\eta}(t)$ is white, and $\{y[n]\}$ are the corresponding sampled output. Thus, for the original channel, one can write

$$\begin{bmatrix} y[1] \\ \vdots \\ y[K_T] \end{bmatrix} = \hat{\mathcal{P}}_T\left(g(t) * x(t)\right) + \hat{\mathcal{P}}_T\left(\hat{\eta}(t)\right).$$

Denote by $\hat{q}(t_i, \tau)$ the impulse response associated with this sampling system. Then, the noise component $\hat{\mathcal{P}}_T\left(\hat{\eta}(t)\right)$ can be whitened by left-multiplying it with a K_T-dimensional square matrix $\boldsymbol{W}_{\hat{\mathcal{P}}}^{-1/2}$ defined by

$$\boldsymbol{W}_{\hat{\mathcal{P}}}(i,j) = \int_{-\infty}^{\infty} \hat{q}\left(t_i, \tau\right) \hat{q}^*\left(t_j, \tau\right) \mathrm{d}\tau.$$

The invertibility is guaranteed by our assumptions. To see this, if we denote by $\tilde{\eta} \stackrel{\Delta}{=} \boldsymbol{W}_{\hat{\mathcal{P}}}^{-1/2} \hat{\mathcal{P}}_T\left(\hat{\eta}(t)\right)$ the K_T-dimensional "prewhitened" noise, then one can verify that for every i and j,

$$\begin{aligned}\left[\mathbb{E}\left(\hat{\mathcal{P}}_T\left(\hat{\eta}(t)\right)\left(\hat{\mathcal{P}}_T\left(\hat{\eta}(t)\right)\right)^{\top}\right)\right]_{ij} &= \mathbb{E}\left[\left(\int_{-\infty}^{\infty} \hat{q}\left(t_i, \tau\right) \hat{\eta}\left(\tau\right) \mathrm{d}\tau\right)\left(\int_{-\infty}^{\infty} \hat{q}^*\left(t_j, \tau\right) \hat{\eta}(\tau) \mathrm{d}\tau\right)\right] \\ &= \int_{-\infty}^{\infty} \hat{q}\left(t_i, \tau\right) \hat{q}^*\left(t_j, \tau\right) \mathrm{d}\tau = \boldsymbol{W}_{\hat{\mathcal{P}}}(i,j)\end{aligned}$$

or, equivalently,

$$\mathbb{E}\left[\hat{\mathcal{P}}_T\left(\hat{\eta}(t)\right)\left(\hat{\mathcal{P}}_T\left(\hat{\eta}(t)\right)\right)^{\top}\right] = \boldsymbol{W}_{\hat{\mathcal{P}}}.$$

As a result, the covariance of $\tilde{\eta}$ obeys

$$\mathbb{E}\left[\tilde{\eta}\tilde{\eta}^{\top}\right] = \boldsymbol{W}_{\hat{\mathcal{P}}}^{-1/2}\mathbb{E}\left(\hat{\mathcal{P}}_{T}\left(\hat{\eta}(t)\right)\left(\hat{\mathcal{P}}_{T}\left(\hat{\eta}(t)\right)\right)^{\top}\right)\boldsymbol{W}_{\hat{\mathcal{P}}}^{-1/2} = \boldsymbol{I}.$$

If we denote by $\hat{\mathcal{P}}_{\mathrm{w}} \triangleq \boldsymbol{W}_{\hat{\mathcal{P}}}^{-\frac{1}{2}} \cdot \hat{\mathcal{P}}_{T}$ and let $\hat{q}_{\mathrm{w}}\left(t_i,\tau\right)$ represent its associated impulse response, then the above calculation reveals that

$$\int_{-\infty}^{\infty} \hat{q}_{\mathrm{w}}\left(t_i,\tau\right)\hat{q}_{\mathrm{w}}^{*}\left(t_j,\tau\right)\mathrm{d}\tau = \begin{cases} 1, & \text{if } i=j; \\ 0, & \text{else,} \end{cases} \tag{A.23}$$

indicating that $\left\{\hat{q}_{\mathrm{w}}\left(t_i,\cdot\right), 1 \leq i \leq K_T\right\}$ forms a set of orthonormal sequences in the corresponding Hilbert space.

For an operator $\mathcal{A}$ with an impulse response $a(t,\tau)$ ($-T \leq \tau \leq T$, $t \in \{t_i \mid 1 \leq i \leq K_T\}$) and input domain $\mathcal{D}\left(\mathcal{A}\right)$, we denote by $\left\Vert \mathcal{A}\right\Vert_{\mathrm{F}}$ the generalized Frobenius norm of the operator $\mathcal{A}$ with respect to its associated domain, namely,

$$\left\Vert \mathcal{A}\right\Vert_{\mathrm{F}} := \begin{cases} \sqrt{\sum_{i=1}^{K_T}\int_{-T}^{T}\left|a(t_i,\tau)\right|^2\mathrm{d}\tau}, & \text{if } \mathcal{D}\left(\mathcal{A}\right) = \{t_i \mid 1 \leq i \leq K_T\}\times[-T,T], \\ \sqrt{\sum_{i=1}^{K_T}\int_{-\infty}^{\infty}\left|a(t_i,\tau)\right|^2\mathrm{d}\tau}, & \text{if } \mathcal{D}\left(\mathcal{A}\right) = \{t_i \mid 1 \leq i \leq K_T\}\times[-\infty,\infty], \\ \sqrt{\int_{-\infty}^{\infty}\int_{-T}^{T}\left|a(t,\tau)\right|^2\mathrm{d}\tau\mathrm{d}t}, & \text{if } \mathcal{D}\left(\mathcal{A}\right) = [-\infty,\infty]\times[-T,T]. \end{cases}$$

Recall that $\hat{q}_{\mathrm{w}}\left(t_i,\cdot\right)$ ($1 \leq i \leq K_T$) forms orthonormal sequences. By Bessel's inequality [130], an operator $\mathcal{A}$ with $\mathcal{D}\left(\mathcal{A}\right) = [-\infty,\infty] \times [-T,T]$ satisfies

$$\sum_{i=1}^{K_T}\left|\left\langle \hat{q}_{\mathrm{w}}\left(t_i,\cdot\right), a\left(\cdot,\tau\right)\right\rangle\right|^2 \leq \int_{-\infty}^{\infty}\left|a\left(\tau_1,\tau\right)\right|^2\mathrm{d}\tau_1$$

for every $\tau \in [-T,T]$, which immediately gives

$$\begin{aligned} \left\Vert \hat{\mathcal{P}}_{\mathrm{w}}\mathcal{A}\right\Vert_{\mathrm{F}}^2 &= \sum_{i=1}^{K_T}\int_{-T}^{T}\left|\int_{-\infty}^{\infty}\hat{q}_{\mathrm{w}}(t_i,\tau_1)a(\tau_1,\tau)\mathrm{d}\tau_1\right|^2\mathrm{d}\tau = \int_{-T}^{T}\sum_{i=1}^{K_T}\left|\left\langle \hat{q}_{\mathrm{w}}\left(t_i,\cdot\right), a\left(\cdot,\tau\right)\right\rangle\right|^2\mathrm{d}\tau \\ &\leq \int_{-T}^{T}\int_{-\infty}^{\infty}\left|a\left(\tau_1,\tau\right)\right|^2\mathrm{d}\tau_1\mathrm{d}\tau \leq \left\Vert \mathcal{A}\right\Vert_{\mathrm{F}}^2. \end{aligned}$$

Denote by $\{\lambda_i\}$ and $\{\tilde{\lambda}_i\}$ the set of *squared* singular values associated with the original sampled channel operator $\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}$ and the operator $\hat{\mathcal{P}}_{\mathrm{w}}\tilde{\mathcal{G}}$ of the truncated sampled channel, respectively. Here, $\mathcal{G}$ and $\tilde{\mathcal{G}}$ represent respectively the operator associated with the original channel response and the truncated channel response. We can obtain some properties of $\{\lambda_i\}$ and $\{\tilde{\lambda}_i\}$ as stated in the following lemma.

Lemma A.1. *Suppose that $\int_{-\infty}^{\infty} |g(t)|^2 \,\mathrm{d}t < C_g < \infty$ for some constant C_g. For any $\xi > 0$, there exists T_0 such that for every $T > T_0$, one has*

(1) $\left| \frac{1}{2T}\sum_i \lambda_i - \frac{1}{2T}\sum_i \tilde{\lambda}_i \right| \leq \xi + 2\sqrt{\xi C_g}$.

(2) $\frac{1}{2T}\sum_i \lambda_i \leq \int_{-\infty}^{\infty} |g(t)|^2 \,\mathrm{d}t < \infty$.

(3) Suppose that $h(t) = O\left(\frac{1}{t^{1.5+\varepsilon}}\right)$ for some small $\varepsilon > 0$. Then there exists $T_{0,\epsilon}$ such that for every $T > T_{0,\epsilon}$, one has $\left|\lambda_i - \tilde{\lambda}_i\right| \leq \xi$.

For notational simplicity, define two functions as follows

$$C_T^{\mathcal{P}}\left(\nu, \{\lambda_i\}\right) := \frac{1}{2T}\sum_{i=1}^{K_T} \frac{1}{2}\left[\log\left(\nu\lambda_i\right)\right]^+, \quad F_T\left(\nu, \{\lambda_i\}\right) := \frac{1}{2T}\sum_{i=1}^{K_T}\left[\nu - \frac{1}{\lambda_i}\right]^+ \tag{A.24}$$

for some water level ν. Note that if ν is chosen such that $F_T\left(\nu, \{\lambda_i\}\right) = P$, then

$$C_T^{\mathcal{P}}\left(\nu, \{\lambda_i\}\right) = C_T^{\mathcal{P}}\left(P\right).$$

Apparently, both $C_T^{\mathcal{P}}\left(P\right)$ and $C_T^{\mathcal{P}}\left(\nu, \{\lambda_i\}\right)$ are non-decreasing functions of $\{\lambda_i\}$, which implies that

$$C_T^{\mathcal{P}}\left(\nu, \{\lambda_i\}\right) \leq C_T^{\mathcal{P}}\left(\nu, \left\{\max\left\{\lambda_i, \xi^{\frac{1}{3}}\right\}\right\}\right), \quad C_T^{\mathcal{P}}\left(P\right) \leq C_T^{\mathcal{P}}\left(\nu, \left\{\max\left\{\lambda_i, \xi^{\frac{1}{3}}\right\}\right\}\right),$$

where ν is determined by

$$F_T\left(\nu, \left\{\max\left(\lambda_i, \xi^{\frac{1}{3}}\right)\right\}\right) = P. \tag{A.25}$$

Here, $\xi > 0$ is some arbitrarily small constant. In fact, one can easily verify that $C_T^{\mathcal{P}}\left(\nu, \left\{\max\left\{\lambda_i, \xi^{\frac{1}{3}}\right\}\right\}\right)$ with ν determined by (A.25) is no larger than the sum capacity of two separate channels with respective eigenvalues $\{\lambda_i\}$ and $\left\{\breve{\lambda}_i := \xi^{\frac{1}{3}}\right\}$

each with power allocation P. In other words,

$$\begin{aligned}
C_T^{\mathcal{P}}\left(\nu,\left\{\max\left\{\lambda_i,\xi^{\frac{1}{3}}\right\}\right\}\right) &\leq C_T^{\mathcal{P}}(P)+C_T^{\mathcal{P}}\left(\nu,\left\{\xi^{\frac{1}{3}}\right\}_{1\leq i\leq K_T}\right)\\
&\leq C_T^{\mathcal{P}}(P)+\frac{K_T}{2T}\log\left(1+\frac{PT}{K_T}\xi^{\frac{1}{3}}\right)\leq C_T^{\mathcal{P}}(P)+\frac{K_T}{2T}\cdot\frac{PT}{K_T}\xi^{\frac{1}{3}}\\
&= C_T^{\mathcal{P}}(P)+\frac{P}{2}\xi^{\frac{1}{3}}. \qquad \text{(A.26)}
\end{aligned}$$

For any positive water level ν and some small constant $\xi > 0$, the Lipschitz constants of the functions

$$f_1(x):=\frac{1}{2}\left[\log\left(\nu\max\left\{x,\xi^{\frac{1}{3}}\right\}\right)\right]^+,\quad f_2(x):=\left[\nu-\max\left\{x,\xi^{\frac{1}{3}}\right\}^{-1}\right]^+.$$

are bounded above in magnitude by $\frac{1}{2}\xi^{-1/3}$ and $\xi^{-2/3}$, respectively. Using the same water level ν, the corresponding power for both channels can be computed as

$$P=\frac{1}{2T}\sum_{i=1}^{K_T}\left[\nu-\frac{1}{\max\left\{\lambda_i,\xi^{\frac{1}{3}}\right\}}\right]^+,\quad \tilde{P}=\frac{1}{2T}\sum_{i=1}^{K_T}\left[\nu-\frac{1}{\max\left\{\tilde{\lambda}_i,\xi^{\frac{1}{3}}\right\}}\right]^+.$$

Combining Lemma A.1 and the Lipschitz constants of $f_2(x)$ immediately suggests that there exists $T_{0,\epsilon}$ such that for any $T > T_{0,\epsilon}$, one has

$$\left|\tilde{P}-P\right|=\frac{1}{2T}\sum_{i=1}^{K_T}\frac{1}{\xi^{\frac{2}{3}}}\left|\lambda_i-\tilde{\lambda}_i\right|\leq\frac{K_T}{2T\xi^{\frac{2}{3}}}\xi\leq(f_s+\epsilon)\,\xi^{\frac{1}{3}}. \qquad \text{(A.27)}$$

Similarly, we can bound

$$\begin{aligned}
&\left|C_T^{\mathcal{P}}\left(\nu,\left\{\max\left\{\lambda_i,\xi^{\frac{1}{3}}\right\}\right\}\right)-C_T^{\mathcal{P}}\left(\nu,\left\{\max\left\{\lambda_i,\xi^{\frac{1}{3}}\right\}\right\}\right)\right|\\
&\leq\frac{1}{2T}\sum_{i=1}^{K_T}\frac{1}{2\xi^{\frac{1}{3}}}\left|\lambda_i-\tilde{\lambda}_i\right|\leq\frac{1}{4}(f_s+\epsilon)\,\xi^{\frac{2}{3}}. \qquad \text{(A.28)}
\end{aligned}$$

Combining (A.27), (A.28) and (A.21) suggests that

$$
\begin{aligned}
C_T^{\mathcal{P}}(P) &\leq C_T^{\mathcal{P}}\left(\nu, \left\{\max\left\{\lambda_i, \xi^{\frac{1}{3}}\right\}\right\}\right) \leq \ C_T^{\tilde{\mathcal{P}}}\left(\nu, \left\{\max\left\{\tilde{\lambda}_i, \xi^{\frac{1}{3}}\right\}\right\}\right) + \frac{1}{4}\left(f_s+\epsilon\right)\xi^{\frac{2}{3}} \\
&\leq C_T^{\tilde{\mathcal{P}}}\left(\tilde{P}\right) + \frac{\tilde{P}}{2}\xi^{\frac{1}{3}} + \frac{1}{4}\left(f_s+\epsilon\right)\xi^{\frac{2}{3}} \qquad\qquad \text{(A.29)} \\
&\leq C_T^{\tilde{\mathcal{P}}}\left(P + \left(f_s+\epsilon\right)\xi^{\frac{1}{3}}\right) + \frac{\tilde{P}}{2}\xi^{\frac{1}{3}} + \frac{1}{4}\left(f_s+\epsilon\right)\xi^{\frac{2}{3}} \\
&\leq \frac{P + \left(f_s+\epsilon\right)\xi^{\frac{1}{3}}}{2}\xi^{\frac{1}{3}} + \ \frac{1}{4}\left(f_s+\epsilon\right)\xi^{\frac{2}{3}} + (1+\delta)\, C_{\mathrm{u}}\left(f_s+\epsilon, P + \left(f_s+\epsilon\right)\xi^{\frac{1}{3}}\right),
\end{aligned}
$$

where (A.29) is a consequence of (A.26). Since δ, ϵ, and ξ can all be made arbitrarily small, it follows that

$$\limsup_{T\to\infty} C_T^{\mathcal{P}}(P) \leq C_{\mathrm{u}}\left(f_s, P\right),$$

completing the proof.

A.4 Proof of Lemma 2.4

The proof is restricted to the channel with white noise, i.e. $\mathcal{S}_\eta(f) \equiv 1$. It is straightforward to extend the analysis to colored noise through the argument presented in the first paragraph of Appendix A.3.

Our proof proceeds in the following three steps.

1. We first introduce several correlation functions and compute the Fourier series associated with them. These quantities are crucial in deriving the capacity expression. In particular, when the sampling system is periodic, the infinite correlation matrices are block Toeplitz.

2. When constrained to a finite time interval $[-nT_q, nT_q]$, the sampled output is a finite vector. The sampled noise is in general not white, which motivates us to whiten it first. In fact, the covariance matrix of the sampled noise can be easily derived in terms of the proposed correlation functions.

3. For any time interval $[-nT_q, nT_q]$, the capacity is obtained through the Karhunen–Loève expansion. Specifically, the capacity depends on the eigenvalues of the associated system operator, which is related to the correlation functions. The asymptotic properties of block Toeplitz matrices guarantee the convergence when $n \to \infty$, which allow us to derive in closed form the sampled channel capacity.

A.4.1 Correlation functions and Fourier series

For a concatenated linear system consisting of the channel filter followed by the sampling system, we denote by

$$s\left(t_{\mathrm{o}}, t_{\mathrm{i}}\right) := \int_{-\infty}^{\infty} h\left(\tau - t_{\mathrm{i}}\right) q\left(t_{\mathrm{o}}, \tau\right) \mathrm{d}\tau \tag{A.30}$$

its system output seen at time t_{o} due to an impulse input at time t_{i}. For notational convenience, we define $q_k\left(\tau\right) := q\left(t_k, \tau\right)$ as the sampling output response at time t_k due to an impulse input to the sampling system at time τ. Two *output autocorrelation functions* are defined as follows

$$\mathcal{R}_{hq}\left(t_k, t_l\right) \overset{\Delta}{=} \int_{-\infty}^{\infty} s\left(t_k, \tau\right) s^*\left(t_l, \tau\right) \mathrm{d}\tau, \quad \text{and} \quad \mathcal{R}_{q}\left(t_k, t_l\right) \overset{\Delta}{=} \int_{-\infty}^{\infty} q\left(t_k, \tau\right) q^*\left(t_l, \tau\right) \mathrm{d}\tau.$$

For notational simplicity, we use $\mathcal{R}_{hq}(k,l)$ (resp. $\mathcal{R}_{q}\left(k,l\right)$) and $\mathcal{R}_{hq}\left(t_k, t_l\right)$ (resp. $\mathcal{R}_{q}\left(t_k, t_l\right)$) interchangeably. When the sampling system is periodic with period T_q, one can easily see that both $[\mathcal{R}_{hq}(k,l)]_{k,l=-\infty}^{\infty}$ and $[\mathcal{R}_{q}\left(k,l\right)]_{k,l=-\infty}^{\infty}$ are infinite block Toeplitz matrices.

The spectral properties associated with the system operators are captured by *Fourier series matrices* $\boldsymbol{F}_{hq}$, $\boldsymbol{F}_{qq}$, $\boldsymbol{F}_{h}$ and $\boldsymbol{F}_{q}$. Specifically, $\boldsymbol{F}_{hq}$ is an f_sT_q-dimensional square matrix such that for any frequency f and all $1 \leq k, i \leq f_sT_q$,

$$\left(\boldsymbol{F}_{hq}\right)_{k,i}(f) := \sum_{l=-\infty}^{\infty} \mathcal{R}_{hq}\left(t_k, t_{i+lf_sT_q}\right) \exp\left(j2\pi l f\right); \tag{A.31}$$

$$\left(\boldsymbol{F}_{qq}\right)_{k,i}(f) := \sum_{l=-\infty}^{\infty} \mathcal{R}_q\left(t_k, t_{i+lf_sT_q}\right)\exp\left(j2\pi lf\right). \tag{A.32}$$

In addition, for every frequency f, we define an $f_qT_s \times \infty$ dimensional matrix $\boldsymbol{F}_q(f)$ and an infinite square diagonal matrix $\boldsymbol{F}_h(f)$ such that for all $l \in \mathbb{Z}$ and $1 \leq k \leq f_qT_s$:

$$\left(\boldsymbol{F}_q\right)_{k,l}(f) := Q_k\left(f + lf_q\right), \quad \text{and} \quad \left(\boldsymbol{F}_h\right)_{l,l}(f) := H\left(f + lf_q\right), \tag{A.33}$$

where $Q_k(f) \overset{\Delta}{=} \mathcal{F}\left(q_k(\cdot)\right) = \mathcal{F}\left(q(t_k, \cdot)\right)$.

The key properties of the above autocorrelation functions and Fourier series are summarized in the following lemma.

Lemma A.2. *The Fourier series matrices satisfy:*

$$\boldsymbol{F}_{hq} = \boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*, \quad \textit{and} \quad \boldsymbol{F}_{qq} = \boldsymbol{F}_q\boldsymbol{F}_q^*. \tag{A.34}$$

A.4.2 Noise whitening

Denote by $\mathcal{Q}_k(\cdot)$ the sampling operator associated with the sample time t_k such that $\mathcal{Q}_k(x) \overset{\Delta}{=} \int_{-\infty}^{\infty} q\left(t_k, \tau\right) x\left(\tau\right) \mathrm{d}\tau$. The correlation of noise components $\mathcal{Q}_k(\eta)$ at different times can be calculated as

$$\begin{aligned}
\mathbb{E}\left[\mathcal{Q}_k(\eta)\,\mathcal{Q}_l^*(\eta)\right] &= \mathbb{E}\left[\int_{-\infty}^{\infty} q\left(t_k, \tau_k\right)\eta\left(\tau_k\right)\mathrm{d}\tau_k\left(\int_{-\infty}^{\infty} q\left(t_l, \tau_l\right)\eta\left(\tau_l\right)\mathrm{d}\tau_l\right)^*\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} q\left(t_k, \tau_k\right) q^*\left(t_l, \tau_l\right)\mathbb{E}\left(\eta\left(\tau_k\right)\eta^*\left(\tau_l\right)\right)\mathrm{d}\tau_k\mathrm{d}\tau_l \\
&= \int_{-\infty}^{\infty} q\left(t_k, \tau\right) q^*\left(t_l, \tau\right)\mathrm{d}\tau,
\end{aligned}$$

which immediately implies that $\mathcal{Q}(\eta) = \left[\cdots, \mathcal{Q}_1(\eta), \mathcal{Q}_2(\eta), \cdots\right]^\top$ is a zero-mean Gaussian vector with covariance matrix $\mathcal{R}_q$.

We now constrain both the transmit interval and the observation interval to $[-nT_q, nT_q]$. Let $\boldsymbol{y}_n = \left[y\left[-nf_sT_q+1\right],,\cdots,y\left[nf_sT_q-1\right]y\left[nf_sT_q\right]\right]^\top$, where the sampled output sequence satisfies

$$y[k] = \mathcal{Q}_k\left(h(t) * x(t)\right) + \mathcal{Q}_k\left(\eta(t)\right). \tag{A.35}$$

Introduce two $2nf_sT_q$-dimensional *truncated* autocorrelation matrices $\mathcal{R}_{hq}^n$ and $\mathcal{R}_q^n$ such that for all $-nf_sT_q < k,l \leq nf_sT_q$,

$$\left(\mathcal{R}_{hq}^n\right)_{k,l} = \mathcal{R}_{hq}\left(t_k,t_l\right), \quad \text{and} \quad \left(\mathcal{R}_q^n\right)_{k,l} = \mathcal{R}_q\left(t_k,t_l\right).$$

Clearly, the noise components of $\boldsymbol{y}_n$ exhibit a covariance matrix $\mathcal{R}_q^n$, which motivates to whiten it first.

By left multiplying $\boldsymbol{y}_n$ with $\left(\mathcal{R}_q^n\right)^{-\frac{1}{2}}$, we obtain a new input-output relation as

$$\tilde{y}_n[k] = \tilde{\mathcal{Q}}_k\left(h(t) * x\left(t\right)\right) + \tilde{\eta}\left[k\right], \quad \forall k\left(|k| \leq nf_sT_q\right),$$

where $\{\tilde{\eta}\left[k\right]\}$ are i.i.d. Gaussian random variables each of unit variance. Denote by $\tilde{q}\left(t_k,\tau\right)$ the equivalent impulse response of this new system. The truncated output autocorrelation function $\mathcal{R}_{\tilde{q}}^n$ is given as $\left(\mathcal{R}_{\tilde{q}}^n\right)_{k,l} = \mathcal{R}_{\tilde{q}}\left(t_k,t_l\right) = \int_{-\infty}^{\infty}\tilde{q}\left(t_k,\tau\right)\tilde{q}^*\left(t_l,\tau\right)\mathrm{d}\tau$, satisfying

$$\mathcal{R}_{\tilde{q}}^n = \left(\mathcal{R}_q^n\right)^{-\frac{1}{2}}\mathcal{R}_{hq}^n\left(\mathcal{R}_q^n\right)^{-\frac{1}{2}} \tag{A.36}$$

by construction.

A.4.3 Capacity via asymptotic properties of block Toeplitz matrices

While both $\mathcal{R}_q^n$ and $\mathcal{R}_{hq}^n$ are block Toeplitz matrices, $\mathcal{R}_{\tilde{q}}^n$ is in general not a block Toeplitz matrix. By exploiting the asymptotic equivalence in Toeplitz matrix theory [131], one can see that $\mathcal{R}_{\tilde{q}}^n$ is asymptotically equivalent to a block-Toeplitz matrix generated by the Fourier series

$$\mathcal{F}\left(\mathcal{R}_q^{-\frac{1}{2}}\right)\mathcal{F}\left(\mathcal{R}_{hq}\right)\mathcal{F}\left(\mathcal{R}_q^{-\frac{1}{2}}\right)=\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}.$$

Therefore, the asymptotic spectral properties of a block-Toeplitz matrix (e.g. [80]) state that for any nondecreasing continuous function $g(t)$ with a bounded slope,

$$\lim_{n\rightarrow\infty}\frac{1}{2nT_q}\sum_{i=1}^{2nf_sT_q}g\left(\lambda_i\left(\mathcal{R}_{\tilde{q}}^n\right)\right)=\frac{1}{2\pi T_q}\int_{-\pi}^{\pi}\sum_{i=1}^{f_sT_q}g\left(\hat{\lambda}_i\right)\mathrm{d}\omega, \tag{A.37}$$

where $\hat{\lambda}_i$ represents the ith eigenvalue of $\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}$.

(1) The capacity of the sampled channel with an optimal water level ν_{p} can now be calculated as

$$C^{\mathcal{P}}(P)=\lim_{n\rightarrow\infty}\frac{1}{2nT_q}\sum_{i=1}^{2nf_sT_q}\frac{1}{2}\left[\log\left(\nu_{\mathrm{p}}\lambda_i\left(\mathcal{R}_{\tilde{q}}^n\right)\right)\right]^+ \tag{A.38}$$

$$=\frac{1}{2\pi T_q}\int_{-\pi}^{\pi}\sum_{i=1}^{f_sT_q}\frac{1}{2}\left[\log\left(\nu_{\mathrm{p}}\hat{\lambda}_i\right)\right]^+\mathrm{d}\omega \tag{A.39}$$

$$=\int_{-f_q/2}^{f_q/2}\sum_{i=1}^{f_sT_q}\frac{1}{2}\left[\log\left(\nu_{\mathrm{p}}\hat{\lambda}_i\right)\right]^+\mathrm{d}f,$$

where (A.39) is a consequence of (A.37).

The water level ν is computed through the following parametric equation

$$\lim_{n\rightarrow\infty}\frac{1}{2nT_q}\sum_{i=1}^{2nf_sT_q}\left[\nu_{\mathrm{p}}-\frac{1}{\lambda_i\left(\mathcal{R}_{\tilde{q}}^n\right)}\right]^+\mathrm{d}f=P,$$

which by (A.37) is asymptotically equivalent to

$$\frac{1}{2\pi T_q}\int_{-\pi}^{\pi}\sum_{i=1}^{f_sT_q}\left[\nu_{\mathrm{p}}-\frac{1}{\hat{\lambda}_i}\right]^+\mathrm{d}\omega=P,\quad\text{or}\quad\int_{-f_q/2}^{f_q/2}\sum_{i=1}^{f_sT_q}\left[\nu_{\mathrm{p}}-\frac{1}{\hat{\lambda}_i}\right]^+\mathrm{d}f=P$$

by change of variables. This establishes the claim.

(2) We consider now the scenario where equal power allocation is employed. Classical MIMO channel capacity results [79] indicate that the optimal power allocation for the transmitter is to allocate equal power in all transmit branches. It remains to see how much power is allocated to the branch associated with $\lambda_i\left(\mathcal{R}_{\tilde{q}}^n\right)$.

In fact, if the transmitter knows the channel bandwidth, almost all power (except for negligible leakage due to the finite-time approximation) will be allocated inside the channel bandwidth $[0,W]$. Therefore, by the Shannon-Nyquist sampling theorem, all transmit signals can be equivalently transformed to a delta train $\sum_{i=-\infty}^{\infty} x_i\delta(t-i/W)$, where x_i's are randomly generated transmit signals. Consider the input time block $[-nT_q, nT_q]$, then there are equivalently $2nT_qW$ transmit branches inside this time block. Since the total power is $P_{\text{tot}} = 2nT_qP$, the power allocated to each transmit branch is given by

$$P_0 = \lim_{n\to\infty} \frac{P_{\text{tot}}}{2nT_q/\left(\frac{1}{W}\right)} = \lim_{n\to\infty} \frac{2nT_qP}{2nWT_q} = \frac{P}{W}.$$

As a result, the sampled capacity under equal power allocation is computed as

$$\begin{aligned} C_{\text{eq}}^{\mathcal{P}}(P) &= \lim_{n\to\infty} \frac{1}{2nT_q} \sum_{i=1}^{2nf_sT_q} \frac{1}{2}\log\left(1+\frac{P}{W}\lambda_i\left(\mathcal{R}_{\tilde{q}}^n\right)\right) = \frac{1}{2\pi T_q}\int_{-\pi}^{\pi}\sum_{i=1}^{f_sT_q}\frac{1}{2}\log\left(1+\frac{P}{W}\hat{\lambda}_i\right)\mathrm{d}\omega \\ &= \int_0^{f_q}\frac{1}{2}\log\det\left(\boldsymbol{I}+\frac{P}{W}\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*\left(\boldsymbol{F}_q\boldsymbol{F}_q^*\right)^{-\frac{1}{2}}\right)\mathrm{d}f. \end{aligned}$$

A.5 Proof of Lemma A.1

(1) Let $\mathcal{G}$ and $\tilde{\mathcal{G}}$ denote respectively the operators associated with $g(t)$ and $\tilde{g}(t)$. Then, the triangle inequality yields

$$\left\|\hat{\mathcal{P}}_{\text{w}}\mathcal{G}\right\|_{\text{F}} \leq \left\|\hat{\mathcal{P}}_{\text{w}}\tilde{\mathcal{G}}\right\|_{\text{F}} + \left\|\hat{\mathcal{P}}_{\text{w}}\left(\mathcal{G}-\tilde{\mathcal{G}}\right)\right\|_{\text{F}} \leq \left\|\hat{\mathcal{P}}_{\text{w}}\tilde{\mathcal{G}}\right\|_{\text{F}} + \left\|\mathcal{G}-\tilde{\mathcal{G}}\right\|_{\text{F}},$$

and hence

$$\left\|\hat{\mathcal{P}}_{\text{w}}\mathcal{G}\right\|_{\text{F}}^2 \leq \left\|\hat{\mathcal{P}}_{\text{w}}\tilde{\mathcal{G}}\right\|_{\text{F}}^2 + \left\|\mathcal{G}-\tilde{\mathcal{G}}\right\|_{\text{F}}^2 + 2\left\|\tilde{\mathcal{G}}\right\|_{\text{F}}\left\|\mathcal{G}-\tilde{\mathcal{G}}\right\|_{\text{F}}. \tag{A.40}$$

From (A.22) one can easily show that for any $\xi > 0$, there exists a T_0 such that for every $T > T_0$, one has

$$\left\|\mathcal{G} - \tilde{\mathcal{G}}\right\|_{\mathrm{F}} \leq \sqrt{2T\left(\int_{-\infty}^{-T} + \int_{T}^{\infty}\right)|h(t)|^2\,\mathrm{d}t} \leq \sqrt{2T\xi}.$$

Additionally, suppose that $\int_{-\infty}^{\infty}|h(t)|^2\,\mathrm{d}t \leq C_g < \infty$. Then, we have

$$\left\|\tilde{\mathcal{G}}\right\|_{\mathrm{F}} \leq \sqrt{2T\int_{-\infty}^{\infty}|h(t)|^2\,\mathrm{d}t} \leq \sqrt{2TC_g}.$$

This together with (A.40) immediately gives us

$$\left\|\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}\right\|_{\mathrm{F}}^2 \leq \left\|\hat{\mathcal{P}}_{\mathrm{w}}\tilde{\mathcal{G}}\right\|_{\mathrm{F}}^2 + 2T\xi + 4T\sqrt{\xi C_g}.$$

Similar to [3, Theorem 8.4.1], we can obtain that

$$\sum_i \lambda_i = \left\|\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}\right\|_{\mathrm{F}}^2 \quad \text{and} \quad \sum_i \tilde{\lambda}_i = \left\|\hat{\mathcal{P}}_{\mathrm{w}}\tilde{\mathcal{G}}\right\|_{\mathrm{F}}^2,$$

which further leads to

$$\frac{1}{2T}\sum_i \lambda_i - \frac{1}{2T}\sum_i \tilde{\lambda}_i = \frac{1}{2T}\left\|\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}\right\|_{\mathrm{F}}^2 - \frac{1}{2T}\left\|\hat{\mathcal{P}}_{\mathrm{w}}\tilde{\mathcal{G}}\right\|_{\mathrm{F}}^2 \leq \xi + 2\sqrt{\xi C_g}.$$

Similarly, one has $\frac{1}{2T}\sum_i \lambda_i - \frac{1}{2T}\sum_i \tilde{\lambda}_i \geq -\xi - 2\sqrt{\xi C_g}$.

(2) We can also bound the sum of eigenvalues as follows

$$\frac{1}{2T}\sum_i \lambda_i = \frac{1}{2T}\left\|\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}\right\|_{\mathrm{F}}^2 \leq \frac{1}{2T}\left\|\mathcal{G}\right\|_{\mathrm{F}}^2 = \int_{-\infty}^{\infty}|h(t)|^2\,\mathrm{d}t < \infty.$$

(3) If $g(t) = O\left(\frac{1}{t^{1+\epsilon}}\right)$, then one can further control

$$\left\|\mathcal{G} - \tilde{\mathcal{G}}\right\|_{\mathrm{F}}^2 \leq 2T\left(\int_{-\infty}^{-T} + \int_{T}^{\infty}\right)|h(t)|^2\,\mathrm{d}t \leq 2T\mathcal{O}\left(\frac{1}{T^{2+2\epsilon}}\right) = \mathcal{O}\left(\frac{1}{T^{1+2\epsilon}}\right).$$

Therefore, applying Weyl's Theorem yields that

$$\begin{aligned}\left|\lambda_i - \tilde{\lambda}_i\right| &\leq \left\|\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}\left(\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}\right)^* - \hat{\mathcal{P}}_{\mathrm{w}}\tilde{\mathcal{G}}\left(\hat{\mathcal{P}}_{\mathrm{w}}\tilde{\mathcal{G}}\right)^*\right\|_{\mathrm{F}} \leq \left\|\hat{\mathcal{P}}_{\mathrm{w}}\left(\mathcal{G}-\tilde{\mathcal{G}}\right)\right\|_{\mathrm{F}}\left(\left\|\hat{\mathcal{P}}_{\mathrm{w}}\mathcal{G}\right\|_{\mathrm{F}} + \left\|\hat{\mathcal{P}}_{\mathrm{w}}\tilde{\mathcal{G}}\right\|_{\mathrm{F}}\right)\\ &\leq \left\|\mathcal{G}-\tilde{\mathcal{G}}\right\|_{\mathrm{F}}\left(\left\|\mathcal{G}\right\|_{\mathrm{F}} + \left\|\tilde{\mathcal{G}}\right\|_{\mathrm{F}}\right) \leq \mathcal{O}\left(\frac{1}{T^{\epsilon}}\right).\end{aligned}$$

Therefore, for any small $\xi > 0$, there exists a constant $T_{0,\epsilon}$ such that for every $T > T_{0,\epsilon}$, one has $\left|\lambda_i - \tilde{\lambda}_i\right| < \xi$.

A.6 Proof of Lemma A.2

Simple manipulation yields

$$\begin{aligned}\mathcal{R}_{hq}\left(t_k, t_l\right) &= \int s\left(t_k,\tau\right)s^*\left(t_l,\tau\right)\mathrm{d}\tau = \iiint q\left(t_k,\tau_k\right)h\left(\tau_k-\tau\right)h^*\left(\tau_l-\tau\right)q^*\left(t_l,\tau_l\right)\mathrm{d}\tau_k\mathrm{d}\tau_l\mathrm{d}\tau\\ &= \iint q_k\left(\tau_k\right)\mathcal{R}_h\left(\tau_l-\tau_k\right)q_l^*\left(\tau_l\right)\mathrm{d}\tau_k\mathrm{d}\tau_l,\end{aligned}$$

where

$$\begin{aligned}\mathcal{R}_h\left(\tau_l-\tau_k\right) &:= \int_\tau h\left(\tau_k-\tau\right)h^*\left(\tau_l-\tau\right)\mathrm{d}\tau = \int_\tau h\left(\tau_k-\tau_l+\tau\right)h^*\left(\tau\right)\mathrm{d}\tau\\ &= \left(h * h^{-*}\right)\left(\tau_k-\tau_l\right).\end{aligned}$$

Here, for any function $f(t)$, we use $f^-(t)$ to denote $f(-t)$.

By the periodicity assumption of the sampling system, one can derive

$$\begin{aligned}\mathcal{R}_{hq}\left(t_{k+af_sT_q}, t_{l+bf_sT_q}\right) &= \iint q\left(t_k+aT_q,\tau_k\right)\mathcal{R}_h\left(\tau_l-\tau_k\right)q^*\left(t_l+bT_q,\tau_l\right)\mathrm{d}\tau_k\mathrm{d}\tau_l\\ &= \iint q\left(t_k,\tau_k-aT_q\right)\mathcal{R}_h\left(\tau_l-\tau_k\right)\cdot q^*\left(t_l+\left(b-a\right)T_q,\tau_l-aT_q\right)\mathrm{d}\tau_k\mathrm{d}\tau_l\\ &= \mathcal{R}_{hq}\left(t_k, t_{l+(b-a)f_sT_q}\right).\end{aligned}$$

Observing that

$$\begin{aligned}\mathcal{R}_{hq}\left(t_k, t_{i+lf_sT_q}\right) &= \iint q\left(t_k,\tau_k\right)\mathcal{R}_h\left(\tau_i-\tau_k\right)q^*\left(t_i+lT_q,\tau_i\right)\mathrm{d}\tau_k\mathrm{d}\tau_i \\ &= \iint q_k\left(\tau_k\right)\mathcal{R}_h\left(\tau_i+lT_q-\tau_k\right)q_i^*\left(\tau_i\right)\mathrm{d}\tau_k\mathrm{d}\tau_i \\ &= \left(\mathcal{R}_h * q_k * q_i^{-*}\right)\left(lT_q\right),\end{aligned}$$

we can see that $\left(\boldsymbol{F}_{hq}\right)_{k,i}$ is simply the Fourier transform of the sampled sequence of $\mathcal{R}_h * q_k * q_i^{-*}$. The properties of the Fourier transform suggest that

$$\mathcal{F}\left(\mathcal{R}_h * q_k * q_i^{-*}\right)(f) = \mathcal{F}\left(\mathcal{R}_h\right)(f)\cdot Q_k(f)\cdot Q_i^*(f) = \left|H(f)\right|^2 Q_k(f)\cdot Q_i^*(f),$$

where $Q_k(f) := \mathcal{F}\left(q_k\right)$. By definition in (A.31), one can write

$$\left(\boldsymbol{F}_{hq}\right)_{k,i}(f) := \sum_{l=-\infty}^{\infty}\mathcal{R}_{hq}\left(t_k, t_{i+lf_sT_q}\right)\exp\left(j2\pi lf\right) = \sum_{l=-\infty}^{\infty}\left(\mathcal{R}_h * q_k * q_i^{-*}\right)\left(lT_q\right)\exp\left(j2\pi lf\right),$$

which immediately leads to

$$\left(\boldsymbol{F}_{hq}\right)_{k,i} = \sum_{l=-\infty}^{\infty} Q_k\left(f+lf_q\right)\left|H\left(f+lf_q\right)\right|^2 Q_i^*\left(f+lf_q\right).$$

This allows us to express $\boldsymbol{F}_{hq}$ as

$$\boldsymbol{F}_{hq} = \boldsymbol{F}_q\boldsymbol{F}_h\boldsymbol{F}_h^*\boldsymbol{F}_q^*. \tag{A.41}$$

Similarly, the equality $\boldsymbol{F}_{qq} = \boldsymbol{F}_q\boldsymbol{F}_q^*$ is then an immediate consequence of (A.41) by setting $\boldsymbol{F}_h = \boldsymbol{I}$.

Appendix B

Proofs of Theorems and Lemmas in Chapter 3

B.1 Proof of Theorem 3.2

We would like to bound the gap between $C_{\boldsymbol{s}}^{\mathrm{wf}}$ and $C_{\boldsymbol{s}}^{\mathrm{eq}}$. In fact, the equation that determines the water level implies that

$$P = \int_0^{W/n} \sum_{i=1}^{k} \left(\nu - (\boldsymbol{H}_{\boldsymbol{s}}(f))_{ii}^{-2} \right)^+ \mathrm{d}f \geq \int_0^{W/n} \sum_{i=1}^{k} \left(\nu - (\boldsymbol{H}_{\boldsymbol{s}}(f))_{ii}^{-2} \right) \mathrm{d}f \tag{B.1}$$

which in turns yields $\nu \leq \frac{P}{\beta W} + \frac{\int_0^{W/n} \sum_{i=1}^{k} \frac{1}{(\boldsymbol{H}_{\boldsymbol{s}}(f))_{ii}^2} \mathrm{d}f}{\beta W}$. With the above bound on the water level, the capacity can be bounded above as

$$\begin{aligned}
C_{\boldsymbol{s}}^{\mathrm{opt}} &= \int_0^{W/n} \frac{1}{2} \sum_{i=1}^{k} \log^+ \left(\nu \left(\boldsymbol{H}_{\boldsymbol{s}}(f) \right)_{ii}^2 \right) \mathrm{d}f \\
&\leq \int_0^{W/n} \frac{1}{2} \sum_{i=1}^{k} \log^+ \left(\frac{\int_0^{W/n} \sum_{j=1}^{k} \frac{(\boldsymbol{H}_{\boldsymbol{s}}(f))_{ii}^2}{(\boldsymbol{H}_{\boldsymbol{s}}(f))_{jj}^2} \mathrm{d}f}{\beta W} + \frac{P}{\beta W} \left(\boldsymbol{H}_{\boldsymbol{s}}(f) \right)_{ii}^2 \right) \mathrm{d}f \\
&\leq \int_0^{W/n} \frac{1}{2} \sum_{i=1}^{k} \log \left(A + \frac{P}{\beta W} \left(\boldsymbol{H}_{\boldsymbol{s}}(f) \right)_{ii}^2 \right) \mathrm{d}f = \int_0^{W/n} \frac{1}{2} \log\det \left(A\boldsymbol{I} + \frac{P}{\beta W} \boldsymbol{H}_{\boldsymbol{s}}^2(f) \right) \mathrm{d}f,
\end{aligned}$$

where $A := \max_{\boldsymbol{s},i} \frac{\int_0^{W/n} \sum_{j=1}^{k} \frac{(\boldsymbol{H_s}(f))_{ii}^2}{(\boldsymbol{H_s}(f))_{jj}^2} \mathrm{d}f}{\beta W}$. One can easily verify that $A \geq 1$. Therefore,

$$\begin{aligned}
C_{\boldsymbol{s}}^{\text{opt}} - C_{\boldsymbol{s}}^{P_{\text{eq}}} \leq\ & \frac{1}{2}\int_0^{W/n} \log\det\left(A\boldsymbol{I} + \frac{P}{\beta W}\boldsymbol{H}_{\boldsymbol{s}}^2(f)\right)\mathrm{d}f - \frac{1}{2}\int_0^{W/n} \log\det\left(\boldsymbol{I} + \frac{P}{\beta W}\boldsymbol{H}_{\boldsymbol{s}}^2(f)\right)\mathrm{d}f \\
\leq\ & \frac{1}{2}\int_0^{W/n} \sum_{i=1}^{k} \log \frac{A + \frac{P}{\beta W}\left(\boldsymbol{H}_{\boldsymbol{s}}^2(f)\right)_{ii}}{1 + \frac{P}{\beta W}\left(\boldsymbol{H}_{\boldsymbol{s}}^2(f)\right)_{ii}}\mathrm{d}f \\
\leq\ & \frac{1}{2}\int_0^{W/n} \sum_{i=1}^{k} \log \frac{A + \inf_{0\leq f\leq W} \frac{P}{\beta W}\frac{|H(f,\boldsymbol{s})|}{\sqrt{\mathcal{S}_\eta(f)}}}{1 + \inf_{0\leq f\leq W} \frac{P}{\beta W}\frac{|H(f,\boldsymbol{s})|}{\sqrt{\mathcal{S}_\eta(f)}}}\mathrm{d}f \\
\leq\ & \frac{1}{2}\beta W \log\left(1 + \frac{A-1}{1+\mathsf{SNR}_{\min}}\right) \leq \frac{1}{2} W \frac{\beta\,(A-1)}{1+\mathsf{SNR}_{\min}}.
\end{aligned}$$

We also observe that

$$\begin{aligned}
A = \max_{\boldsymbol{s},i} \frac{\int_0^{W/n} \sum_{j=1}^{k} \frac{(\boldsymbol{H_s}(f))_{ii}^2}{(\boldsymbol{H_s}(f))_{jj}^2}\mathrm{d}f}{\beta W} &\leq \frac{\int_0^{W/n} \sum_{j=1}^{k} \frac{(\boldsymbol{H_s}(f))_{ii}^2}{\inf_{f,\boldsymbol{s}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}}\mathrm{d}f}{\beta W} \\
&\leq \min\left\{ \frac{\max_{\boldsymbol{s}\in\binom{[n]}{k}} \int_0^W \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}\mathrm{d}f}{\beta W \inf_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}}, \frac{\sup_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}}{\inf_{0\leq f\leq W, \boldsymbol{s}\in\binom{[n]}{k}} \frac{|H(f,\boldsymbol{s})|^2}{\mathcal{S}_\eta(f)}} \right\}.
\end{aligned}$$

Combining the above bounds and Theorem 3.1 completes the proof.

B.2 Proof of Theorem 3.6

Before proving the results, we first state two facts. Consider any $m \times m$ matrix $\boldsymbol{A}$ with eigenvalues $\lambda_1, \cdots, \lambda_m$. Define the characteristic polynomial of $\boldsymbol{A}$ as

$$p_{\boldsymbol{A}}(t) = \det(t\boldsymbol{I} - \boldsymbol{A}) = t^m - S_1(\lambda_1, \cdots, \lambda_m)t^{m-1} + \cdots + (-1)^m S_m(\lambda_1, \cdots, \lambda_m),$$

where $S_l(\lambda_1, \cdots, \lambda_m)$ is the lth elementary symmetric function of $\lambda_1, \cdots, \lambda_m$ defined as follows:

$$S_l(\lambda_1, \cdots, \lambda_m) := \sum_{1 \leq i_1 < \cdots < i_l \leq m} \prod_{j=1}^{l} \lambda_{i_j}.$$

We also define $E_l(\boldsymbol{A})$ as the sum of determinants of all l-by-l principal minors of $\boldsymbol{A}$. According to [132, Theorem 1.2.12], $S_l(\lambda_1, \cdots, \lambda_m) = E_l(\boldsymbol{A})$ holds for all $1 \leq l \leq m$. After a little manipulation we obtain

$$\det(t\boldsymbol{I} + \boldsymbol{A}) = t^m + E_1(\boldsymbol{A})t^{m-1} + \cdots + E_m(\boldsymbol{A}). \tag{B.2}$$

Another fact we would like to stress is the entropy formula of binomial coefficients. Specifically, for any $0 < k < n$, one has [133, Page 43]

$$\frac{e^{n\mathcal{H}(\beta)}}{n+1} \leq \binom{n}{k} \leq e^{n\mathcal{H}(\beta)},$$

and hence

$$\mathcal{H}(\beta) - \frac{\log(n+1)}{n} \leq \frac{1}{n} \log \binom{n}{k} \leq \mathcal{H}(\beta), \tag{B.3}$$

where $\mathcal{H}(x) := x \log \frac{1}{x} + (1-x) \log \frac{1}{1-x}$ denotes the entropy function.

Now we are in a position to derive the proof for our main results.

(1) Consider an $m \times n$ matrix $\boldsymbol{B}$ with orthonormal rows ($m \geq k$), i.e. $\boldsymbol{B}\boldsymbol{B}^* = \boldsymbol{I}_m$. Using this identity (B.2), we can derive

$$\begin{aligned}\sum_{\boldsymbol{s} \in \binom{[n]}{k}} \det(\epsilon \boldsymbol{I}_m + \boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^*) &= \sum_{\boldsymbol{s} \in \binom{[n]}{k}} \left\{ \epsilon^m + \sum_{l=1}^{m} \epsilon^{m-l} E_l(\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^*) \right\} \\ &= \epsilon^m \binom{n}{k} + \sum_{l=1}^{k} \epsilon^{m-l} \sum_{\boldsymbol{s} \in \binom{[n]}{k}} E_l(\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^*),\end{aligned} \tag{B.4}$$

where the last equality follows by observing that any lth order ($l > k$) minor of $\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^*$ is rank deficient, and hence $E_l(\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^*) = 0$.

Consider an index set $\boldsymbol{r} \in \binom{[m]}{l}$ with $l \leq k$, and denote by $(\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^{*})_{\boldsymbol{r}}$ the submatrix of $\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^{*}$ with rows and columns coming from the index set $\boldsymbol{r}$. One can then verify that

$$\det\left((\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^{*})_{\boldsymbol{r}}\right) = \det\left(\boldsymbol{B}_{\boldsymbol{r},\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{r},\boldsymbol{s}}^{*}\right) = \sum_{\tilde{\boldsymbol{r}} \in \binom{\boldsymbol{s}}{l}} \det\left(\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}^{*}\right),$$

where the last equality arises from the Cauchy-Binet formula (e.g. [127]). Some algebraic manipulation yields that for any $l \leq k$:

$$\begin{aligned}
\sum_{\boldsymbol{s} \in \binom{[n]}{k}} E_l\left(\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^{*}\right) &= \sum_{\boldsymbol{s} \in \binom{[n]}{k}} \sum_{\boldsymbol{r} \in \binom{[m]}{l}} \det\left((\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^{*})_{\boldsymbol{r}}\right) = \sum_{\boldsymbol{s} \in \binom{[n]}{k}} \sum_{\boldsymbol{r} \in \binom{[m]}{l}} \sum_{\tilde{\boldsymbol{r}} \in \binom{\boldsymbol{s}}{l}} \det\left(\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}^{*}\right) \\
&= \sum_{\boldsymbol{r} \in \binom{[m]}{l}} \sum_{\tilde{\boldsymbol{r}} \in \binom{[n]}{l}} \sum_{\boldsymbol{s}:\tilde{\boldsymbol{r}} \subseteq \boldsymbol{s}} \det\left(\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}^{*}\right) \\
&\overset{(a)}{=} \sum_{\boldsymbol{r} \in \binom{[m]}{l}} \sum_{\tilde{\boldsymbol{r}} \in \binom{[n]}{l}} \binom{n-l}{k-l} \det\left(\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}^{*}\right) \\
&\overset{(b)}{=} \sum_{\boldsymbol{r} \in \binom{[m]}{l}} \binom{n-l}{k-l} = \binom{n-l}{k-l}\binom{m}{l},
\end{aligned} \tag{B.5}$$

where (a) follows from the fact that the number of k-combinations (out of $[n]$) containing $\tilde{\boldsymbol{r}}$ (an l-combination) is $\binom{n-l}{k-l}$, and (b) follows from the Cauchy-Binet formula and the fact that $\boldsymbol{B}\boldsymbol{B}^{*} = \boldsymbol{I}_m$, i.e.

$$\sum_{\tilde{\boldsymbol{r}} \in \binom{[n]}{l}} \det\left(\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}\boldsymbol{B}_{\boldsymbol{r},\tilde{\boldsymbol{r}}}^{*}\right) = \det\left(\boldsymbol{B}_{\boldsymbol{r},[n]}\boldsymbol{B}_{\boldsymbol{r},[n]}^{*}\right) = \det\left(\boldsymbol{I}_l\right) = 1.$$

Substituting (B.5) into (B.4) yields

$$\begin{aligned}
\sum_{\boldsymbol{s} \in \binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I}_m + \boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^{*}\right) &= \epsilon^m \binom{n}{k} + \sum_{l=1}^{k} \epsilon^{m-l} \sum_{\boldsymbol{s} \in \binom{[n]}{k}} E_l\left(\boldsymbol{B}_{\boldsymbol{s}}\boldsymbol{B}_{\boldsymbol{s}}^{*}\right) \\
&= \sum_{l=0}^{k} \binom{n-l}{k-l}\binom{m}{l} \epsilon^{m-l}.
\end{aligned}$$

By observing that $\boldsymbol{B_s}$ is a tall $m \times k$ matrix, one has

$$\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{B}_{\boldsymbol{s}}^* \boldsymbol{B}_{\boldsymbol{s}}\right) = \epsilon^{k-m} \det\left(\epsilon \boldsymbol{I}_m + \boldsymbol{B}_{\boldsymbol{s}} \boldsymbol{B}_{\boldsymbol{s}}^*\right)$$

$$\Rightarrow \quad \sum_{\boldsymbol{s} \in \binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{B}_{\boldsymbol{s}}^* \boldsymbol{B}_{\boldsymbol{s}}\right) = \sum_{l=0}^{k} \binom{n-l}{k-l} \binom{m}{l} \epsilon^{k-l}.$$

The above expression allows us to derive a crude bound as

$$\begin{aligned}
\binom{n}{k} \min_{\boldsymbol{s} \in \binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I} + \boldsymbol{B}_{\boldsymbol{s}}^* \boldsymbol{B}_{\boldsymbol{s}}\right) &\leq \sum_{\boldsymbol{s} \in \binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I} + \boldsymbol{B}_{\boldsymbol{s}}^* \boldsymbol{B}_{\boldsymbol{s}}\right) = \sum_{l=0}^{k} \binom{n-l}{k-l} \binom{m}{l} \epsilon^{k-l} \\
&\leq \sum_{l=0}^{k} \binom{n}{k-l} \binom{m}{l} \epsilon^{k-l} = \sum_{l=0}^{k} \binom{n}{l} \binom{m}{m-k+l} \epsilon^{l} \\
&= \binom{m}{k} \sum_{l=0}^{k} \binom{n}{l} \frac{\binom{m}{m-k+l}}{\binom{m}{k}} \epsilon^{l}.
\end{aligned}$$

Since the term $\binom{m}{m-k+l} / \binom{m}{k}$ can be bounded above as

$$\frac{\binom{m}{m-k+l}}{\binom{m}{k}} = \frac{\frac{m!}{(m-k+l)!(k-l)!}}{\frac{m!}{k!(m-k)!}} = \frac{k!}{(k-l)! \frac{(m-k+l)!}{(m-k)!}} = \frac{\binom{k}{l}}{\binom{m-k+l}{l}} \leq \binom{k}{l},$$

we obtain

$$\begin{aligned}
\binom{n}{k} \min_{\boldsymbol{s} \in \binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I} + \boldsymbol{B}_{\boldsymbol{s}}^* \boldsymbol{B}_{\boldsymbol{s}}\right) &\leq \sum_{\boldsymbol{s} \in \binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I} + \boldsymbol{B}_{\boldsymbol{s}}^* \boldsymbol{B}_{\boldsymbol{s}}\right) \leq \binom{m}{k} \sum_{l=0}^{k} \binom{n}{l} \frac{\binom{m}{m-k+l}}{\binom{m}{k}} \epsilon^{l} \\
&\leq \binom{m}{k} \sum_{l=0}^{k} \binom{n}{l} \binom{k}{l} \epsilon^{l} \leq \binom{m}{k} \sum_{l=0}^{k} \binom{n+k}{2l} \left(\sqrt{\epsilon}\right)^{2l} \\
&\leq \binom{m}{k} \sum_{i=0}^{n+k} \binom{n+k}{i} \left(\sqrt{\epsilon}\right)^{i} \\
&= \binom{m}{k} \left(1 + \sqrt{\epsilon}\right)^{n+k}, \qquad \text{(B.6)}
\end{aligned}$$

where the last equality follows from the binomial theorem.

(2) Since $\boldsymbol{B}_i$ has orthonormal rows, applying the inequality of arithmetic and geometric means yields

$$
\begin{aligned}
\sum_{\boldsymbol{s}\in\binom{[n]}{k}} \left(\prod_{i=1}^{p} \det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right)\right)^{\frac{1}{p}} &\leq \sum_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{\sum_{i=1}^{p} \det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right)}{p} \\
&= \frac{1}{p}\sum_{i=1}^{p}\left\{\sum_{\boldsymbol{s}\in\binom{[n]}{k}} \det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right)\right\} \\
&\leq \binom{m}{k}\left(1+\sqrt{\epsilon}\right)^{n+k},
\end{aligned}
$$

where the second inequality follows from (B.6). Since $\boldsymbol{M}_\alpha$ has orthonormal rows, applying (B.6) yields

$$
\begin{aligned}
\binom{n}{k}\min_{\boldsymbol{s}\in\binom{[n]}{k}} \left(\prod_{i=1}^{p} \det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right)\right)^{\frac{1}{p}} &\leq \sum_{\boldsymbol{s}\in\binom{[n]}{k}} \left(\prod_{i=1}^{p} \det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right)\right)^{\frac{1}{p}} \\
&\leq \binom{m}{k}\left(1+\sqrt{\epsilon}\right)^{n+k}. \qquad \text{(B.7)}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\min_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{1}{np}\sum_{i=1}^{p} \log\det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right) &= \min_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{1}{n}\log\left(\prod_{i=1}^{p} \det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right)\right)^{\frac{1}{p}} \\
&\leq \frac{1}{n}\log\left(\frac{\binom{m}{k}}{\binom{n}{k}}\left(1+\sqrt{\epsilon}\right)^{n+k}\right) \\
&= \frac{1}{n}\log\binom{m}{k} - \frac{1}{n}\log\binom{n}{k} + \frac{n+k}{n}\log\left(1+\sqrt{\epsilon}\right) \\
&\leq \frac{1}{n}\log\binom{m}{k} - \frac{1}{n}\log\binom{n}{k} + 2\sqrt{\epsilon}.
\end{aligned}
$$

When (n, k, m) are all large numbers, the entropy approximation (B.3) allows us to approximate the above bound as

$$\begin{aligned}\min_{\boldsymbol{s}\in\binom{[n]}{k}} \frac{1}{np}\sum_{i=1}^{p}\log\det\left(\epsilon \boldsymbol{I}_k + (\boldsymbol{B}_i)_{\boldsymbol{s}}^{*}(\boldsymbol{B}_i)_{\boldsymbol{s}}\right) &\le \frac{m}{n}\mathcal{H}\left(\frac{k}{m}\right) - \mathcal{H}\left(\frac{k}{n}\right) + \frac{\log(n+1)}{n} + 2\sqrt{\epsilon} \\ &= \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) - \mathcal{H}(\beta) + 2\sqrt{\epsilon} + \frac{\log(n+1)}{n}.\end{aligned} \tag{B.8}$$

B.3 Proof of Theorem 3.7

The quantity of interest can be bounded from below by

$$\begin{aligned}\log\det\left(\epsilon \boldsymbol{I}_k + \left(\boldsymbol{M}\boldsymbol{M}^{\top}\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right) &= \log\det\left(\epsilon\boldsymbol{M}\boldsymbol{M}^{\top} + \boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right) - \log\det\left(\boldsymbol{M}\boldsymbol{M}^{\top}\right) \\ \ge\ \log\det\left(\frac{\epsilon}{k}\sigma_{\min}\left(\boldsymbol{M}\boldsymbol{M}^{\top}\right)\boldsymbol{I}_k + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\right) &- \log\det\left(\frac{1}{k}\boldsymbol{M}\boldsymbol{M}^{\top}\right),\end{aligned}$$

which helps separate $\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^{\top}$ and $\boldsymbol{M}\boldsymbol{M}^{\top}$ if $\sigma_{\min}\left(\boldsymbol{M}\boldsymbol{M}^{\top}\right)$ is a constant bounded away from zero. The behavior of the least singular value of a rectangular random matrix with independent sub-Gaussian entries has been largely studied in the random matrix literature (e.g. [134, Theorem 3.1] and [135, Corollary 5.35]), which we cite as follows.

Lemma B.1. *Suppose that $m = (1-\delta)n$ for some absolute constant $\delta \in (0,1)$. Let $\boldsymbol{M}$ be an $m \times n$ real-valued random matrix whose entries are jointly independent symmetric sub-Gaussian random variables with zero mean and unit variance. Then there exist universal constants $C, c > 0$ such that*

$$\sigma_{\min}\left(\boldsymbol{M}\boldsymbol{M}^{*}\right) > Cm \tag{B.9}$$

with probability at least $1 - 2\exp(-cn)$.

In particular, if the entries of $\boldsymbol{M}$ are i.i.d. standard Gaussian random variables, then for any constant $0 < \xi < \sqrt{n} - \sqrt{m}$,

$$\sigma_{\min}\left(\boldsymbol{M}\boldsymbol{M}^{*}\right) > \left(\sqrt{n} - \sqrt{m} - \xi\right)^2$$

with probability at least[1] $1-\exp\left(-\xi^2/2\right)$.

Setting $m = k$, we can derive that with probability exceeding $1-2\exp\left(-cn\right)$,

$$\log\det\left(\epsilon \boldsymbol{I}_k + \left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right) \geq \log\det\left(\epsilon C\boldsymbol{I}_k + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right) - \log\det\left(\frac{1}{k}\boldsymbol{M}\boldsymbol{M}^\top\right) \tag{B.10}$$

holds for general independent symmetric sub-Gaussian matrices. Also, with probability at least $1-2e^{-2}$,

$$\begin{aligned}
&\log\det\left(\epsilon \boldsymbol{I}_k + \left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right) \\
&\quad \geq \log\det\left(\epsilon\frac{\left(1-\sqrt{\beta}-\frac{2}{\sqrt{n}}\right)^2}{\beta}\boldsymbol{I}_k + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right) - \log\det\left(\frac{1}{k}\boldsymbol{M}\boldsymbol{M}^\top\right)
\end{aligned} \tag{B.11}$$

holds for i.i.d. standard Gaussian matrices and any constant $\xi \in (0,1)$.

The next step is to quantify the term $\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right)$ for some small $0<\epsilon<1$. This has been characterized in Lemma 3.4, indicating that

$$\forall \boldsymbol{s} \in \binom{[n]}{k}, \quad \log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right) \geq -1 - \mathcal{O}\left(\min\left\{\frac{1}{\sqrt{n\epsilon}}\right\}, \frac{\log\frac{1}{\epsilon}+\log n}{n^{1/4}}\right) \tag{B.12}$$

with exponentially high probability. In addition, since $\boldsymbol{M}$ satisfies the assumptions of Lemma 3.5 with $\beta = \alpha$, simple manipulation gives

$$\begin{aligned}
\frac{1}{n}\log\det\left(\frac{1}{k}\boldsymbol{M}\boldsymbol{M}^\top\right) &= \frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top\right) + \frac{k}{n}\log\frac{n}{k} \\
&\leq (1-\beta)\log\frac{1}{1-\beta} - \beta + \beta\log\frac{1}{\beta} + \frac{1}{\sqrt{n}}
\end{aligned} \tag{B.13}$$

with probability exceeding $1 - C_7\exp\left(-c_7 n\right)$. The above results (B.9), (B.10), (B.12) and (B.13) taken collectively yield the following: under the condition of Theorem 3.7,

[1] Note that this follows from [135, Proposition 5.34 and Theorem 5.32] by observing that $\sigma_{\min}\left(\boldsymbol{M}\right)$ is a 1-Lipschitz function.

one has

$$\forall \boldsymbol{s}: \quad \frac{1}{n}\log\det\left(\epsilon \boldsymbol{I}_k + \left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right) \geq -\mathcal{H}\left(\beta\right) + \frac{\log k}{2k} - \mathcal{O}\left(\frac{1}{\sqrt{n\epsilon}}\right)$$

with probability exceeding $1 - C\exp\left(-cn\right)$ for some absolute constants $c, C > 0$. Combining this lower bound with the upper bound developed in Theorem 3.6 (with $\alpha = \beta$) concludes the proof.

B.4 Proof of Theorem 3.8

Our goal is to evaluate $\frac{1}{n}\log\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{M}_{\boldsymbol{s}}^\top\left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right)$ for some small $\epsilon > 0$. We first define two Wishart matrices $\Xi_{\backslash \boldsymbol{s}} := \frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top - \frac{1}{n}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top$ and $\Xi_{\boldsymbol{s}} := \frac{1}{n}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top$. Apparently, $\Xi_{\boldsymbol{s}} \sim \mathcal{W}_m\left(k, \frac{1}{n}\boldsymbol{I}_m\right)$ and $\Xi_{\backslash \boldsymbol{s}} \sim \mathcal{W}_m\left(n-k, \frac{1}{n}\boldsymbol{I}_m\right)$. When $1-\alpha > \beta$, i.e. $n-k > m$, the Wishart matrix $\Xi_{\backslash \boldsymbol{s}}$ is invertible with probability 1.

One difficulty in evaluating $\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{M}_{\boldsymbol{s}}^\top\left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right)$ is that $\boldsymbol{M}_{\boldsymbol{s}}$ and $\boldsymbol{M}\boldsymbol{M}^\top$ are not independent. This motivates us to decouple them first as follows

$$\begin{aligned}
\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{M}_{\boldsymbol{s}}^\top\left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right) &= \epsilon^{k-m}\det\left(\epsilon \boldsymbol{I}_m + \left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\frac{1}{n}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right) \\
&= \epsilon^{k-m}\det\left(\epsilon\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top + \frac{1}{n}\boldsymbol{M}_{\boldsymbol{s}}\boldsymbol{M}_{\boldsymbol{s}}^\top\right)\det\left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1} \\
&= \epsilon^{k-m}\det\left(\epsilon\Xi_{\backslash \boldsymbol{s}} + \left(1+\epsilon\right)\Xi_{\boldsymbol{s}}\right)\det\left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1} \\
&= \epsilon^{k-m}\det\left(\epsilon \boldsymbol{I}_m + \left(1+\epsilon\right)\Xi_{\boldsymbol{s}}\Xi_{\backslash \boldsymbol{s}}^{-1}\right)\det\left(\Xi_{\backslash \boldsymbol{s}}\right)\det\left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1} \\
&= \det\left(\epsilon \boldsymbol{I}_k + \left(1+\epsilon\right)\frac{1}{n}\boldsymbol{M}_{\boldsymbol{s}}^\top\Xi_{\backslash \boldsymbol{s}}^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right)\det\left(\Xi_{\backslash \boldsymbol{s}}\right)\det\left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}
\end{aligned}$$

or, equivalently,

$$\frac{1}{n}\log\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{M}_{\boldsymbol{s}}^{\top}\left(\boldsymbol{M}\boldsymbol{M}^{\top}\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right) = \frac{1}{n}\log\det\left(\epsilon \boldsymbol{I}_k + (1+\epsilon)\frac{1}{n}\boldsymbol{M}_{\boldsymbol{s}}^{\top}\Xi_{\backslash \boldsymbol{s}}^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right) + \frac{1}{n}\log\det\left(\Xi_{\backslash \boldsymbol{s}}\right) - \frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^{\top}\right). \tag{B.14}$$

The point of developing this identity (B.14) is to decouple the left-hand side of (B.14) through 3 matrices $\boldsymbol{M}_{\boldsymbol{s}}^{\top}\Xi_{\backslash \boldsymbol{s}}^{-1}\boldsymbol{M}_{\boldsymbol{s}}$, $\Xi_{\backslash \boldsymbol{s}}$ and $\boldsymbol{M}\boldsymbol{M}^{\top}$. In particular, since $\boldsymbol{M}_{\boldsymbol{s}}$ and $\Xi_{\backslash \boldsymbol{s}}$ are jointly independent, we can examine the concentration of measure for $\boldsymbol{M}_{\boldsymbol{s}}$ and $\Xi_{\backslash \boldsymbol{s}}$ separately when evaluating $\boldsymbol{M}_{\boldsymbol{s}}^{\top}\Xi_{\backslash \boldsymbol{s}}^{-1}\boldsymbol{M}_{\boldsymbol{s}}$.

The second and third terms of (B.14) can be evaluated through Lemma 3.6. Specifically, Lemma 3.6 indicates that

$$\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{M}\boldsymbol{M}^{\top}\right) \leq -(1-\alpha)\log(1-\alpha) - \alpha + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \tag{B.15}$$

with probability at least $1 - C_6\exp(-2n)$ for some constant $C_6 > 0$, and that for all $\boldsymbol{s} \in \binom{[n]}{k}$,

$$\begin{aligned}
\frac{1}{n}\log\det\left(\Xi_{\backslash \boldsymbol{s}}\right) &= \frac{n-k}{n}\frac{1}{n-k}\log\det\left(\frac{n}{n-k}\Xi_{\backslash \boldsymbol{s}}\right) + \frac{1}{n}\log\det\left(\frac{n-k}{n}\boldsymbol{I}\right) \\
&\geq (1-\beta)\left\{-\left(1-\frac{\alpha}{1-\beta}\right)\log\left(1-\frac{\alpha}{1-\beta}\right) - \frac{\alpha}{1-\beta}\right\} + \alpha\log(1-\beta) + \mathcal{O}\left(\frac{\log n}{n^{1/3}}\right) \\
&\geq -(1-\alpha-\beta)\log\left(1-\frac{\alpha}{1-\beta}\right) - \alpha + \alpha\log(1-\beta) + \mathcal{O}\left(\frac{\log n}{n^{1/3}}\right)
\end{aligned} \tag{B.16}$$

hold simultaneously with probability exceeding $1 - C_9\exp(-2n)$.

Our main task then amounts to quantifying $\log\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{M}_{\boldsymbol{s}}^{\top}\Xi_{\backslash \boldsymbol{s}}^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right)$, which can be lower bounded via Lemma 3.7. This together with (B.15), (B.16) and (B.14) yields that

$$
\begin{aligned}
&\frac{1}{n}\log\det\left(\epsilon \boldsymbol{I}_k + \boldsymbol{M}_{\boldsymbol{s}}^\top \left(\boldsymbol{M}\boldsymbol{M}^\top\right)^{-1}\boldsymbol{M}_{\boldsymbol{s}}\right) \\
&\quad \geq -(\alpha-\beta)\log(\alpha-\beta) + \alpha\log\alpha + (1-\alpha-\beta)\log\left(1-\frac{\beta}{1-\alpha}\right) - \beta\log(1-\alpha) \\
&\qquad -(1-\alpha-\beta)\log\left(1-\frac{\alpha}{1-\beta}\right) - \alpha + \alpha\log(1-\beta) + (1-\alpha)\log(1-\alpha) + \alpha - \mathcal{O}\left(\frac{\log n}{n^{1/3}}\right) \\
&\quad = \alpha\mathcal{H}\left(\frac{\beta}{\alpha}\right) - \mathcal{H}(\beta) - \mathcal{O}\left(\frac{\log n}{n^{1/3}}\right)
\end{aligned}
$$

with probability exceeding $1 - C_9\exp(-2n)$ for some constants $C_9 > 0$.

Since there are at most $\binom{n}{k} < e^n$ different states $\boldsymbol{s}$, applying the union bound over all states completes the proof.

B.5 Proof of Lemma 3.1

Set $f(x) = \log(\epsilon + x)$. Observe that the Lipschitz constant of $g(x) := f(x^2) = \log(\epsilon + x^2)$ is upper bounded by

$$
\forall x \geq 0, \quad |g'(x)| = \left|\frac{2x}{\epsilon + x^2}\right| = \left|\frac{2}{\frac{\epsilon}{x} + x}\right| \leq \frac{1}{\sqrt{\epsilon}}.
$$

When $\boldsymbol{A}_{ij}$'s are bounded in magnitude by D, define the following *concave* function

$$
g_\epsilon(x) := \begin{cases} \log(\epsilon + x^2), & \text{if } x \geq \sqrt{\epsilon}, \\ \frac{1}{\sqrt{\epsilon}}(x-\epsilon) + \log(2\epsilon), & \text{if } 0 \leq x < \sqrt{\epsilon}, \end{cases}
$$

whose Lipschitz constant is bounded by $\frac{1}{\sqrt{\epsilon}}$. This function obeys the following interlacing bound

$$
\log\left(\frac{2}{e}\epsilon + x\right) \leq g_\epsilon\left(\sqrt{x}\right) \leq \log(\epsilon + x), \quad \forall x \geq 0. \tag{B.17}
$$

Define the function

$$f_{\epsilon}(\boldsymbol{A}) := \frac{1}{k}\sum_{i=1}^{k} g_{\epsilon}\left(\sqrt{\lambda_i\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)}\right), \tag{B.18}$$

then it follows from (B.17) that

$$\frac{1}{k}\log\det\left(\frac{2\epsilon}{e}\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \le f_{\epsilon}(\boldsymbol{A}) \le \frac{1}{k}\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right). \tag{B.19}$$

One can then apply [28, Corollary 1.8(a)] to derive that for any $\delta > \frac{8D\sqrt{\pi}}{\sqrt{\epsilon}}$,

$$f_{\epsilon}(\boldsymbol{A}) \ge \mathbb{E}\left[f_{\epsilon}(\boldsymbol{A})\right] - \frac{1}{k}\delta \ge \mathbb{E}\left[\frac{1}{k}\log\det\left(\frac{2\epsilon}{e}\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \frac{1}{k}\delta$$

with probability at least $1-4\exp\left(-\frac{\epsilon}{4D^2}\left(\delta-\frac{4D\sqrt{\pi}}{\sqrt{\epsilon}}\right)^2\right)$. This together with (B.19) yields that for any $\delta > \frac{8D\sqrt{\pi}}{\sqrt{\epsilon}}$,

$$\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \ge k f_{\epsilon}(\boldsymbol{A}) \ge \mathbb{E}\left[\log\det\left(\frac{2\epsilon}{e}\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \delta$$

with probability at least $1-4\exp\left(-\frac{\epsilon}{16D^2}\delta^2\right)$. A similar argument indicates that for any $\delta > \frac{8D\sqrt{\pi}}{\sqrt{\frac{e\epsilon}{2}}}$,

$$\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \le k f_{\frac{e}{2}\epsilon}(\boldsymbol{A}) \le \mathbb{E}\left[\log\det\left(\frac{e\epsilon}{2}\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \delta$$

with probability at least $1-4\exp\left(-\frac{e\epsilon}{32D^2}\delta^2\right)$.

If $\boldsymbol{A}_{ij}$ satisfies the LSI with uniformly bounded constant c_{LS}, then applying [28, Corollary 1.8(b)] leads to

$$\left|\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]\right| > \delta$$

with probability at most $2\exp\left(-\frac{\epsilon\delta^2}{2c_{\mathrm{LS}}}\right)$, as claimed.

B.6 Proof of Lemma 3.2

(1) By Jensen's inequality,

$$\begin{aligned}\frac{1}{k}\mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] &\leq \frac{1}{k}\log\mathbb{E}\left[\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]\\ &\leq -1+\frac{1.5\log(ek)}{k}+2\sqrt{\epsilon}\log\frac{1}{\epsilon},\end{aligned}$$

where the last inequality follows from [72, Lemma 3].

(2) The lower bound follows from the concentration inequality. Define

$$Y := k\left(f_\epsilon\left(\boldsymbol{A}\right)-\mathbb{E}\left[f_\epsilon\left(\boldsymbol{A}\right)\right]\right),$$

where $f_\epsilon(\boldsymbol{A})$ is defined in (B.18). Similar to the proof of Lemma 3.1, applying [28, Corollary 1.8] indicates that

$$\mathbb{P}\left(|Y|>\delta\right)\leq 4\exp\left(-\tilde{c}\epsilon\delta^2\right)$$

for some constant $\tilde{c}$. If we denote by $f_Y(\cdot)$ the probability density function of Y, then

$$\begin{aligned}\mathbb{E}\left(e^Y\right)\leq\mathbb{E}\left(e^{|Y|}\right)&=\int_0^\infty e^y f_{|Y|}(y)\,\mathrm{d}y=-e^y\mathbb{P}\left(|Y|>y\right)|_0^\infty+\int_0^\infty e^y\mathbb{P}\left(|Y|>y\right)\mathrm{d}y\\ &\leq 1+\int_0^\infty 4\exp\left(y-\tilde{c}\epsilon y^2\right)\mathrm{d}y\ <\ 1+4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right).\end{aligned}$$

Taking the logarithms of both sides and plugging in the expression of Y yields

$$\begin{aligned}\log\left(\mathbb{E}\left[e^{kf_\epsilon(\boldsymbol{A})}\right]\right)&\leq k\mathbb{E}\left[f_\epsilon\left(\boldsymbol{A}\right)\right]+\log\left[1+4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right)\right]\\ &\leq\mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I}+\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]+\log\left[1+4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right)\right].\end{aligned}\tag{B.20}$$

Also, the inequality $kf_\epsilon(\boldsymbol{A}) \geq \log\det\left(\frac{2\epsilon}{e}\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)$ allows us to derive

$$\begin{aligned}\mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] &\geq \log\left(\mathbb{E}\left[e^{kf_\epsilon(\boldsymbol{A})}\right]\right) - \log\left[1 + 4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right)\right] \\ &\geq \log\mathbb{E}\left[\det\left(\frac{2\epsilon}{e}\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] - \log\left[1 + 4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right)\right]. \qquad \text{(B.21)}\end{aligned}$$

Furthermore, if we denote by $\prod_k$ the permutation group of k elements, then the Leibniz formula for the determinant gives

$$\det(\boldsymbol{A}) = \sum_{\sigma\in\prod_k}\mathrm{sgn}(\sigma)\prod_{i=1}^{k}\boldsymbol{A}_{i,\sigma(i)}.$$

Taking advantage of the joint independence hypothesis yields

$$\mathbb{E}\left[\det\left(\boldsymbol{A}\boldsymbol{A}^\top\right)\right] = \mathbb{E}\left[(\det(\boldsymbol{A}))^2\right] = \sum_{\sigma\in\prod_k}\mathbb{E}\left[\prod_{i=1}^{k}\left|\boldsymbol{A}_{i,\sigma(i)}\right|^2\right] = \sum_{\sigma\in\prod_k}\prod_{i=1}^{k}\mathbb{E}\left[\left|\boldsymbol{A}_{i,\sigma(i)}\right|^2\right] = k!$$

This taken collectively with (B.21) yields that

$$\begin{aligned}\frac{1}{k}\mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^T\right)\right] &\geq \frac{1}{k}\log\mathbb{E}\left[\det\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] - \frac{1}{k}\log\left[1 + 4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right)\right] \\ &= \frac{1}{k}\log\frac{k!}{k^k} - \frac{1}{k}\log\left[1 + 4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right)\right] \\ &= -1 + \frac{\log k}{2k} - \frac{1}{k}\log\left[1 + 4\sqrt{\frac{\pi}{\tilde{c}\epsilon}}\exp\left(\frac{1}{4\tilde{c}\epsilon}\right)\right],\end{aligned}$$

where the last inequality makes uses of the well-known Stirling-type inequality

$$k! \geq k^{k+\frac{1}{2}}e^{-k}, \quad \Rightarrow \quad \frac{1}{k}\log\frac{k!}{k^k} \geq \frac{\log k}{2k} - 1.$$

This indicates that $\frac{1}{k}\mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] \geq -1 + \frac{\log k}{2k} - \mathcal{O}\left(\frac{1}{k\epsilon}\right).$

In particular, for i.i.d. Gaussian ensemble $\boldsymbol{A}_{ij} \sim \mathcal{N}(0,1)$, one has $\tilde{c} = \frac{1}{2}$, and hence for any $\epsilon \leq 0.8$:

$$\begin{aligned}\frac{1}{k}\mathbb{E}\left[\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] &\geq -1 + \frac{\log k}{2k} - \frac{1}{k}\log\left[1 + 4\sqrt{\frac{\pi}{2\epsilon}}\exp\left(\frac{1}{2\epsilon}\right)\right] \\ &\geq -1 + \frac{\log k}{2k} - \frac{1}{k}\log\left[\exp\left(\frac{2}{\epsilon}\right)\right] \\ &= -1 + \frac{\log k}{2k} - \frac{2}{k\epsilon},\end{aligned}$$

where the second inequality uses the inequality that $1 + 4\sqrt{\frac{\pi}{2\epsilon}}\exp\left(\frac{1}{2\epsilon}\right) \leq \exp\left(\frac{2}{\epsilon}\right)$ for any $0 < \epsilon \leq 0.8$.

B.7 Proof of Lemma 3.3

We first attempt to estimate the number of eigenvalues of $\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^\top$ not exceeding ϵ. To this end, we consider the following function

$$f_{2,\xi}(x) := \begin{cases} -\sqrt{\frac{x}{\xi}} + 2, & \text{if } x \leq 4\xi, \\ 0 & \text{else,} \end{cases}$$

as well as the convex function

$$g_{2,\xi}(x) := f_{2,\xi}\left(x^2\right) = \begin{cases} -\frac{x}{\sqrt{\xi}} + 2, & \text{if } x \leq 2\sqrt{\xi}, \\ 0 & \text{else.} \end{cases}$$

Thus,

$$\mathbf{1}_{(0,\xi)}(x) \leq f_{2,\xi}(x) \leq 2 \cdot \mathbf{1}_{(0,4\xi)}(x), \quad \forall x \tag{B.22}$$

and the Lischiptz constant of $g_{2,\xi}(\cdot)$ is bounded above by $\frac{1}{\sqrt{\xi}}$.

By our assumptions, $\boldsymbol{A}_{ij}$'s are symmetric and sub-Gaussian. If we generate the random matrix $\tilde{\boldsymbol{A}}$ such that

$$\tilde{\boldsymbol{A}}_{ij} := \boldsymbol{A}_{ij}\mathbf{1}_{\left\{|\boldsymbol{A}_{ij}| \leq k^{1/4}\right\}},$$

then one can easily verify that

$$\mathbb{P}\left(\tilde{\boldsymbol{A}}=\boldsymbol{A}\right)=1-o\left(1\right),\quad \mathbb{E}\left[\tilde{\boldsymbol{A}}_{ij}\right]=0,\quad \mathbb{E}\left[\left|\tilde{\boldsymbol{A}}_{ij}\right|^{2}\right]=1-o\left(1\right).$$

If we denote by

$$N_{I}\left(\boldsymbol{A}\right):=\mathrm{card}\left\{i\mid\lambda_{i}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\in I\right\}$$

for any interval I, then it follows from the theory on local spectral statistics [136, Theorem 4.1] that

$$N_{(0,\xi)}\left(\tilde{\boldsymbol{A}}\right)\leq k\int_{0}^{\xi}\rho\left(x\right)\mathrm{d}x+\frac{1}{4}k\xi$$

with probability at least $1-c_5/k^3$ for some constant $c_5>0$, where $\rho(x)$ represents the Marchenko–Pastur law

$$\rho(x):=\frac{1}{2\pi x}\sqrt{\left(4-x\right)x}\cdot\mathbf{1}_{[0,1]}\left(x\right)\leq\frac{1}{\pi\sqrt{x}}.$$

This immediately implies that

$$\frac{1}{k}\mathbb{E}\left[N_{(0,\xi)}\left(\boldsymbol{A}\right)\right]\leq\frac{c_5}{k^3}+\left(1-\frac{c_5}{k^3}\right)\left(\int_{0}^{\xi}\rho\left(x\right)\mathrm{d}x+\frac{1}{4}\epsilon\right)\leq\frac{2}{\pi}\sqrt{\xi}+\frac{1}{4}\xi+\frac{c_5}{k^3}. \tag{B.23}$$

The concentration inequality [28, Corollary 1.8] ensures that for any $\delta=\Omega\left(\frac{1}{k\sqrt{\xi}}\right)$,

$$\frac{1}{k}N_{(0,\xi)}\left(\boldsymbol{A}\right)\leq\ \frac{1}{k}\sum_{i=1}^{k}f_{2,\xi}\left(\lambda_{i}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right)\leq\ \frac{1}{k}\sum_{i=1}^{k}\mathbb{E}\left[f_{2,\xi}\left(\lambda_{i}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right)\right]+\delta$$

with probability exceeding $1-4\exp\left(-\tilde{c}\xi\delta^{2}k^{2}\right)$ for some constant $\tilde{c}>0$. The inequality (B.22) suggests that

$$\frac{1}{k}\sum_{i=1}^{k}\mathbb{E}\left[f_{2,\xi}\left(\lambda_{i}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right)\right]\leq\frac{2}{k}\cdot\mathbb{E}\left[N_{(0,4\xi)}\left(\boldsymbol{A}\right)\right]\leq\frac{8}{\pi}\sqrt{\xi}+2\xi+\frac{2c_5}{k^3},$$

and hence for any $\tau \geq \frac{1}{\sqrt{k}}$,

$$\frac{1}{k}N_{(0,\xi)}\left(\boldsymbol{A}\right) \leq \frac{8}{\pi}\sqrt{\xi} + 2\xi + \frac{2c_5}{k^3} + \sqrt{\frac{\tau}{\tilde{c}k\xi}}$$

with probability at least $1 - 4\exp\left(-\tau k\right)$. By setting $\xi = \sqrt{\frac{\tau}{k}}$, one can derive that with probability exceeding $1 - 4\exp\left(-\tau k\right)$,

$$\frac{1}{k}N_{(0,\xi)}\left(\boldsymbol{A}\right) \leq \frac{c_{11}\tau^{1/4}}{k^{1/4}}, \quad \forall \frac{1}{k} \leq \tau < k$$

holds for some universal constant $c_{11} > 0$.

The above estimates on the eigenvalue concentration allow us to derive

$$\begin{aligned}\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) &\geq \frac{1}{k}\log\det{}^{\epsilon+\xi}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \frac{N_{(0,\xi)}\left(\boldsymbol{A}\right)}{k}\log\left(\frac{\epsilon+\xi}{\epsilon}\right)\\ &\geq \frac{1}{k}\log\det{}^{\xi}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \frac{c_{12}\tau^{1/4}}{k^{1/4}}\left(\log\frac{1}{\epsilon} + \log k\right) \quad \text{(B.24)}\end{aligned}$$

with exponential high probability, where $c_{12} > 0$ is another constant. Here, the function $\det^{\xi}\left(\cdot\right)$ is defined in (B.27). Now it suffices to obtain a lower estimate on $\log\det^{\xi}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right)$. Putting the bounds (B.29) and (B.31) together indicates that there exists a constant $c_{13} > 0$ such that

$$\begin{aligned}\frac{1}{k}\log\det{}^{\xi}\left(\frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) &\geq -1 + \frac{\log\left(2\pi k\right)}{2k} - \frac{c_{13}}{k\xi} - \frac{c_{13}\sqrt{\tau}}{\sqrt{k\xi}}\\ &\geq -1 + \frac{\log\left(2\pi k\right)}{2k} - \frac{c_{13}}{k^{1/2}\tau^{1/2}} - \frac{c_{13}\tau^{1/4}}{k^{1/4}}\end{aligned}$$

with probability exceeding $1 - 4\exp\left(-\tau k\right)$. This combined with (B.24) yields that for any $\frac{1}{\sqrt{k}} \leq \tau < k$,

$$\frac{1}{k}\log\det\left(\epsilon\boldsymbol{I} + \frac{1}{k}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \geq -1 + \frac{\log\left(2\pi k\right)}{2k} - \frac{c_{14}\tau^{1/4}}{k^{1/4}}\left(\log\frac{1}{\epsilon} + \log k\right)$$

with probability exceeding $1 - 4\exp\left(-\tau k\right)$, as claimed.

B.8 Proof of Lemma 3.5

We first develop the upper bound. The Cauchy-Binet formula indicates that

$$\mathbb{E}\left[\det\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] = \sum_{\boldsymbol{s}\in\binom{[n]}{m}} \mathbb{E}\left[\det\left(\boldsymbol{A}_{\boldsymbol{s}}\boldsymbol{A}_{\boldsymbol{s}}^{\top}\right)\right],$$

where $\boldsymbol{s}$ ranges over all m-combinations of $\{1,\cdots,n\}$, and $\boldsymbol{A}_{\boldsymbol{s}}$ is the $m\times m$ minor of $\boldsymbol{A}$ whose columns are the columns of $\boldsymbol{A}$ at indices from $\boldsymbol{s}$. It has been shown in [72, Equation (74)] that for each jointly independent $m\times m$ ensemble $\boldsymbol{A}_{\boldsymbol{s}}$, the determinant satisfies

$$\mathbb{E}\left[\det\left(\boldsymbol{A}_{\boldsymbol{s}}\boldsymbol{A}_{\boldsymbol{s}}^{\top}\right)\right] = m!,$$

which immediately leads to

$$\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] = \frac{1}{n^m}\sum_{\boldsymbol{s}\in\binom{[n]}{m}} \mathbb{E}\left[\det\left(\boldsymbol{A}_{\boldsymbol{s}}\boldsymbol{A}_{\boldsymbol{s}}^{\top}\right)\right] = \frac{m!}{n^m}\binom{n}{m}.$$

Besides, using the Stirling type inequality $\sqrt{2\pi}m^{m+\frac{1}{2}}e^{-m} \leq m! \leq em^{m+\frac{1}{2}}e^{-m}$, one can obtain

$$\log(m!) \leq \log\left(em^{m+\frac{1}{2}}e^{-m}\right) = \left(m+\frac{1}{2}\right)\log m - m + 1$$

and similarly

$$\log(m!) \geq \left(m+\frac{1}{2}\right)\log m - m + \frac{1}{2}\log(2\pi).$$

These further give rise to

$$\begin{aligned}\frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] &\leq -\frac{m}{n}\log n + \frac{\left(m+\frac{1}{2}\right)\log m}{n} - \frac{m}{n} + \frac{1}{n} + \mathcal{H}\left(\frac{m}{n}\right)\\ &= -\frac{m}{n}\log n + \frac{m\log m}{n} - \frac{m}{n} + \frac{2+\log m}{2n} + \mathcal{H}\left(\frac{m}{n}\right)\\ &= (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{\log\left(e^2 m\right)}{2n} \qquad \text{(B.25)}\end{aligned}$$

and, similarly,

$$\frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]\geq(1-\alpha)\log\frac{1}{1-\alpha}-\alpha-\frac{\log(n+1)}{2n}. \tag{B.26}$$

Define

$$f_{1,\epsilon}(x):=\begin{cases}\frac{2}{\sqrt{\epsilon}}\left(\sqrt{x}-\sqrt{\epsilon}\right)+\log\epsilon, & 0<x<\epsilon,\\ \log x, & x\geq\epsilon,\end{cases}$$

and

$$\det{}^{\epsilon}(\boldsymbol{X}):=\prod_{i=1}^{m}e^{f_{1,\epsilon}(\lambda_i(\boldsymbol{X}))}. \tag{B.27}$$

One can easily justify that the Lischiptz constant of the function

$$g_{1,\epsilon}(x):=f_{1,\epsilon}\left(x^2\right)=\begin{cases}\frac{2}{\sqrt{\epsilon}}\left(x-\sqrt{\epsilon}\right)+\log\epsilon, & 0<x<\sqrt{\epsilon},\\ 2\log x & x\geq\sqrt{\epsilon},\end{cases}$$

is bounded above by $\frac{2}{\sqrt{\epsilon}}$, and that $g_{1,\epsilon}(x)$ is a *concave* function. Also,

$$\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)=\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)$$

holds in the event that $\left\{\lambda_{\min}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\geq\epsilon\right\}$.

Set

$$Z:=\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)-\mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]. \tag{B.28}$$

Applying [28, Corollary 1.8] then gives

$$\mathbb{P}\left(|Z|>\tau\right)\leq 4\exp\left(-\frac{\tilde{c}\epsilon\tau^2}{\alpha}\right) \tag{B.29}$$

for some constant $\tilde{c}$ (in particular, $\tilde{c} = \frac{1}{8}$ when $\boldsymbol{A}_{ij} \sim \mathcal{N}(0,1)$). This allows us to derive

$$\begin{aligned}
\mathbb{E}\left[e^{Z}\right] \leq \mathbb{E}\left[e^{|Z|}\right] &= -e^{z}\mathbb{P}\left(|Z|>z\right)|_{z=0}^{\infty} + \int_{0}^{\infty} e^{z}\mathbb{P}\left(|Z|>z\right)\mathrm{d}z \\
&\leq 1 + \int_{0}^{\infty} 4\exp\left(z - \frac{\tilde{c}\epsilon z^{2}}{\alpha}\right)\mathrm{d}z \\
&< 1 + 4\sqrt{\frac{\pi\alpha}{\tilde{c}\epsilon}}\exp\left(\frac{\alpha}{4\tilde{c}\epsilon}\right). \qquad \text{(B.30)}
\end{aligned}$$

Taking the logarithm of both sides of (B.30) and plugging in the expression of Z yields

$$\log\mathbb{E}\left[\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] \leq \mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] + \log\left[1 + 4\sqrt{\frac{\pi\alpha}{\tilde{c}\epsilon}}\exp\left(\frac{\alpha}{4\tilde{c}\epsilon}\right)\right],$$

leading to

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] &\geq \frac{1}{n}\log\mathbb{E}\left[\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \mathcal{O}\left(\frac{1}{n\epsilon}\right) \\
&\geq \frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \mathcal{O}\left(\frac{1}{n\epsilon}\right) \\
&= (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{\log(2\pi m)}{2n} - \mathcal{O}\left(\frac{1}{n\epsilon}\right). \qquad \text{(B.31)}
\end{aligned}$$

In the Gaussian case, this can be more explicitly expressed as follows: for any $\epsilon \leq \alpha$,

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] &\geq \frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \frac{1}{n}\log\left[1 + 8\sqrt{\frac{2\pi\alpha}{\epsilon}}\exp\left(\frac{2\alpha}{\epsilon}\right)\right] \\
&\geq \frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \frac{1}{n}\log\left[\exp\left(\frac{6\alpha}{\epsilon}\right)\right] \\
&\geq (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{\log(2\pi m)}{2n} - \frac{6\alpha}{n\epsilon}. \qquad \text{(B.32)}
\end{aligned}$$

where we have made use of the fact that $1 + 8\sqrt{\pi x}\exp(x) \leq \exp(3x)$ for any $x \geq 2$.

On the other hand, in order to develop an upper bound on $\mathbb{E}\left[\log\det^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]$, we set $\tau = \sqrt{\frac{\alpha\log n}{\tilde{c}\epsilon}}$ in the inequality (B.29), which reveals that with probability at least $1-4n^{-1}$,

$$|Z| = \left|\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]\right| \leq \sqrt{\frac{\alpha\log n}{\tilde{c}\epsilon}}$$

or, equivalently,

$$\frac{1}{e^{\sqrt{\frac{\alpha\log n}{\tilde{c}\epsilon}}}} e^{\mathbb{E}\left[\log\det^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]} \leq \det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \leq e^{\sqrt{\frac{\alpha\log n}{\tilde{c}\epsilon}}} \cdot e^{\mathbb{E}\left[\log\det^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]}$$

with probability exceeding $1-4n^{-1}$. In addition, the tail distribution for $\sigma_{\min}\left(\boldsymbol{A}\right)$ satisfies [137, Theorem 1.1]

$$\mathbb{P}\left\{\sigma_{\min}\left(\frac{1}{\sqrt{n}}\boldsymbol{A}\right) \leq \tilde{\epsilon}\left(1-\sqrt{\alpha}\right)\right\} \leq \left(C\tilde{\epsilon}\right)^{(1-\alpha)n} + e^{-cn}$$

for some constants $C, c > 0$. For sufficiently small $\epsilon > 0$, one can simply write

$$\mathbb{P}\left\{\sigma_{\min}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \leq \epsilon\right\} \leq C_{\mathrm{m}}\exp\left(-c_{\mathrm{m}}n\right)$$

for some constants $c_{\mathrm{m}}, C_{\mathrm{m}} > 0$, implying that

$$\mathbb{P}\left\{\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) = \det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right\} \geq 1 - C_{\mathrm{m}}\exp\left(-c_{\mathrm{m}}n\right). \tag{B.33}$$

Hence, the union bound suggests that

$$\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) = \det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \quad \text{and} \quad \det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) > e^{-1}e^{\mathbb{E}\left[\log\det^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]}$$

hold simultaneously with probability exceeding $1-\frac{\check{c}}{n}$ for some constant $\check{c} > 0$. Consequently,

$$\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] \geq \left(1-\frac{\check{c}}{n}\right)\frac{1}{e^{\sqrt{\frac{\alpha\log n}{\tilde{c}\epsilon}}}} e^{\mathbb{E}\left[\log\det^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]},$$

indicating that for sufficiently large n,

$$
\begin{aligned}
\frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] &\geq \frac{1}{n}\mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]+\frac{\log\left(1-\frac{\check{c}}{n}\right)}{n}-\frac{\sqrt{\frac{\alpha\log n}{\check{c}\epsilon}}}{n} \\
&\geq \frac{1}{n}\mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]-\mathcal{O}\left(\frac{\sqrt{\log n}}{\sqrt{\epsilon}n}\right)
\end{aligned}
$$

Equivalently, we have, for sufficiently large n and sufficiently small constant $\epsilon > 0$,

$$
\begin{aligned}
\frac{1}{n}\mathbb{E}\left[\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] &\leq \frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]+\mathcal{O}\left(\frac{\sqrt{\log n}}{\sqrt{\epsilon}n}\right) \\
&\leq (1-\alpha)\log\frac{1}{1-\alpha}-\alpha+\frac{\log\left(e^{2}m\right)}{2n}+\mathcal{O}\left(\frac{\sqrt{\log n}}{\sqrt{\epsilon}n}\right),
\end{aligned}
\tag{B.34}
$$

where the last inequality follows from (B.25).

Now we are ready to develop a high-probability bound on $\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)$. By picking $\epsilon > 0$ to be a sufficiently small *constant*, one has

$$
\begin{aligned}
&\mathbb{P}\left\{\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\geq(1-\alpha)\log\frac{1}{1-\alpha}-\alpha+\frac{1}{\sqrt{n}}\right\} \\
&\leq\ \mathbb{P}\left\{\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\geq(1-\alpha)\log\frac{1}{1-\alpha}-\alpha+\frac{1}{\sqrt{n}}\right\} \\
&=\ \mathbb{P}\left\{\frac{1}{n}Z\geq\frac{1}{\sqrt{n}}-\mathcal{O}\left(\frac{\log n}{n\sqrt{\epsilon}}\right)\right\} \\
&\leq\ C_{7}\exp\left(-c_{7}n\right)
\end{aligned}
$$

for some absolute constants $c_7, C_7 > 0$. Similarly, for a sufficiently small *constant* $\epsilon > 0$,

$$\begin{aligned}
&\mathbb{P}\left\{\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) \leq (1-\alpha)\log\frac{1}{1-\alpha} - \alpha - \frac{1}{\sqrt{n}}\right\} \leq \mathbb{P}\left\{\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) \neq \det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right\} \\
&\qquad\qquad + \mathbb{P}\left\{\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) \leq (1-\alpha)\log\frac{1}{1-\alpha} - \alpha - \frac{1}{\sqrt{n}}\right\} \\
&\leq\ C_{\mathrm{m}}\exp\left(-c_{\mathrm{m}}n\right) + \mathbb{P}\left\{\frac{1}{n}Z \leq -\frac{1}{\sqrt{n}} + \mathcal{O}\left(\frac{\log n}{n} + \frac{1}{n\epsilon}\right)\right\} \qquad \text{(B.35)} \\
&\leq\ \tilde{C}_7\exp\left(-\tilde{c}_7 n\right)
\end{aligned}$$

for some absolute constants $\tilde{c}_7, \tilde{C}_7 > 0$. Here, (B.35) is a consequence of (B.33) and (B.31). This establishes Part (1) of the lemma.

B.9 Proof of Lemma 3.6

(1) In order to obtain a lower bound on $\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)$, we first attempt to estimate the number of eigenvalues of $\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top$ that are smaller than some value $\epsilon > 0$, i.e. $\sum_{i=1}^{n}\mathbf{1}_{[0,\epsilon]}\left(\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right)$. Since the indicator function $\mathbf{1}_{[0,\epsilon]}(\cdot)$ entails discontinuous points, we define an upper bound on $\mathbf{1}_{[0,\epsilon]}(\cdot)$ as

$$f_{2,\epsilon}(x) := \begin{cases} 1, & \text{if } 0 \leq x \leq \epsilon; \\ -x/\epsilon + 2, & \text{if } \epsilon < x \leq 2\epsilon; \\ 0, & \text{else.} \end{cases}$$

For any $\epsilon > 0$, one can easily verify that[2]

$$f_{2,\epsilon}(x) \leq \frac{\epsilon}{x}, \quad \forall x \geq 0,$$

[2]This follows from the inequality $\frac{\epsilon}{x} \geq 1$ $(0 \leq x \leq \epsilon)$ as well as the bound that $\frac{\epsilon}{x} + \frac{x}{\epsilon} - 2 \geq 2\sqrt{\frac{\epsilon}{x}\cdot\frac{x}{\epsilon}} - 2 = 0$.

and hence

$$\sum_{i=1}^{n} f_{2,\epsilon}\left(\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right) \leq \sum_{i=1}^{m} \frac{\epsilon}{\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)} = \epsilon \mathrm{tr}\left(\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)^{-1}\right).$$

This further gives

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} f_{2,\epsilon}\left(\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right)\right] \leq \frac{\epsilon}{n}\mathbb{E}\left[\mathrm{tr}\left(\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)^{-1}\right)\right] = \frac{m\epsilon}{n-m-1} = \frac{\alpha}{1-\alpha-\frac{1}{n}}\epsilon,$$

which follows from the property of inverse Wishart matrices (e.g. [138, Theorem 2.2.8]).

Clearly, the Lipschitz constant of the function

$$g_{2,\epsilon}(x) := f_{2,\epsilon}(x^2) = \begin{cases} 1, & \text{if } 0 \leq x \leq \sqrt{\epsilon}; \\ -x^2/\epsilon + 2, & \text{if } \sqrt{\epsilon} < x \leq \sqrt{2\epsilon}; \\ 0, & \text{else.} \end{cases}$$

is bounded above by $\sqrt{8/\epsilon}$. Applying [28, Corollary 1.8(b)] then yields that for any $\delta > 0$

$$\begin{aligned} &\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{[0,\epsilon]}\left(\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right) > \frac{\alpha}{1-\alpha-\frac{1}{n}}\epsilon + \delta\right) \\ &\leq \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} f_{2,\epsilon}\left(\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right) > \frac{\alpha}{1-\alpha-\frac{1}{n}}\epsilon + \delta\right) \\ &\leq \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} f_{2,\epsilon}\left(\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right) > \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} f_{2,\epsilon}\left(\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right)\right] + \delta\right) \\ &\leq 2\exp\left(-\frac{\epsilon}{16\alpha}\delta^2 n^2\right). \end{aligned}$$

In other words, this immediately implies that for any $\tau > 0$,

$$\frac{\mathrm{card}\left\{i \mid \lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) < \epsilon\right\}}{n} < \frac{\alpha}{1-\alpha-\frac{1}{n}}\epsilon + \frac{4\sqrt{\alpha\tau}}{\sqrt{n\epsilon}} \tag{B.36}$$

with probability exceeding $1-2\exp(-\tau n)$, as claimed. By setting $\epsilon = n^{-1/3}$, one has

$$\frac{\mathrm{card}\left\{i \mid \lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) < \frac{1}{n^{1/3}}\right\}}{n} < \frac{\frac{\alpha}{1-\alpha-\frac{1}{n}} + 4\sqrt{\alpha\tau}}{n^{1/3}} \tag{B.37}$$

with probability at least $1-2\exp(-\tau n)$.

(2) Lemma B.1 asserts that if $n > \frac{2}{1-\sqrt{\alpha}}$, then

$$\lambda_{\min}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) \geq \left(1-\sqrt{\alpha}-\frac{2}{n}\right)^2$$

with probability at least $1-\exp\left(-\frac{2}{n}\right)$. Conditional on this event, we have

$$\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) = \det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)$$

for the numerical value $\epsilon = \left(1-\sqrt{\alpha}-\frac{2}{n}\right)^2$. Additionally, the bound (B.29) under Gaussian ensembles can be explicitly written as

$$\mathbb{P}\left(|Z| > \tau\right) \leq 2\exp\left(-\frac{\epsilon\tau^2}{8\alpha}\right), \tag{B.38}$$

indicating that

$$\mathbb{P}\left\{\left|\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) - \mathbb{E}\left[\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]\right| > \sqrt{\frac{4\alpha}{\epsilon n}}\right\} < 2\exp\left(-\frac{n}{2}\right) \tag{B.39}$$

and

$$\mathbb{P}\left\{\left|\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) - \mathbb{E}\left[\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]\right| > \frac{4\tau\sqrt{\frac{\alpha}{\epsilon}}}{n}\right\} < 2\exp\left(-2\tau^2\right). \tag{B.40}$$

Putting these together yields that for any $n > \max\left\{\frac{2}{1-\sqrt{\alpha}}, 6\right\}$,

$$\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) = \det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) \quad \text{and} \quad \det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) > \frac{1}{e^{\sqrt{\frac{4\alpha}{\epsilon n}}}} e^{\mathbb{E}\left[\log\det^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]}$$

simultaneously hold with probability exceeding $1 - 2\exp\left(-\frac{n}{2}\right) - \exp\left(-\frac{2}{n}\right) > 1/n$. Since $\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)$ is non-negative, this gives

$$\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] \geq \frac{1}{n} \cdot \frac{1}{e^{\sqrt{\frac{4\alpha}{\epsilon n}}}} e^{\mathbb{E}\left[\log\det^\epsilon\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right]},$$

and therefore for any $n > \max\left\{\frac{2}{1-\sqrt{\alpha}}, 7\right\}$,

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\left[\log\det^\epsilon\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] &\leq \frac{1}{n}\log\mathbb{E}\left[\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right)\right] + \frac{\log n}{n} + \frac{\sqrt{\frac{4\alpha}{\epsilon}}}{n^{1.5}} \\
&\leq (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{\log\left(e^2 m\right) + 2\log n}{2n} + \frac{2\sqrt{\alpha}}{\left(1-\sqrt{\alpha}-\frac{2}{n}\right)}\frac{1}{n^{1.5}} \\
&< (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{3\log n}{n} + \frac{2\sqrt{\alpha}}{\left(1-\sqrt{\alpha}-\frac{2}{n}\right)}\frac{1}{n^{1.5}},
\end{aligned}$$

where the second inequality follows from (B.25) and the value $\epsilon = \left(1-\sqrt{\alpha}-\frac{2}{n}\right)^2$.

Putting this and (B.40) together gives that for any $n \geq 4$,

$$\begin{aligned}
\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) &\leq \frac{4\sqrt{\alpha}}{\left(1-\sqrt{\alpha}-\frac{2}{n}\right)}\frac{\tau}{n} + (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{3\log n}{n} + \frac{2\sqrt{\alpha}}{\left(1-\sqrt{\alpha}-\frac{2}{n}\right)}\frac{1}{n^{1.5}} \\
&< (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{3\log n}{n} + \frac{4\sqrt{\alpha}}{\left(1-\sqrt{\alpha}-\frac{2}{n}\right)n}\left(\tau + \frac{1}{2\sqrt{n}}\right) \\
&\leq (1-\alpha)\log\frac{1}{1-\alpha} - \alpha + \frac{3\log n}{n} + \frac{5\sqrt{\alpha}}{\left(1-\sqrt{\alpha}-\frac{2}{n}\right)n}\tau
\end{aligned}$$

with probability exceeding $1 - 2\exp\left(-2\tau^2\right)$.

(3) On the other hand, the inequality (B.37) indicates that for any $n > \left(\frac{\alpha}{1-\alpha-\frac{1}{n}} + 4\sqrt{\alpha\tau}\right)^3$,

$$\mathbb{P}\left\{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^\top\right) \leq \frac{1}{n^{1/3}}\right\} \leq \mathbb{P}\left\{\frac{\mathrm{card}\left\{i \mid \lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^\top\right) < \frac{1}{n^{1/3}}\right\}}{n} > \frac{\frac{\alpha}{1-\alpha-\frac{1}{n}} + 4\sqrt{\alpha\tau}}{n^{1/3}}\right\} \leq 2e^{-\tau n}.$$

Also, it follows from [139, Theorem 4.5] that for any constant $n \geq \frac{6.414}{1-\alpha} \cdot e^{\frac{\tau}{1-\alpha}}$,

$$\mathbb{P}\left\{\frac{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right)}{\lambda_{\min}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right)} > n^2\right\} = \mathbb{P}\left\{\frac{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right)}{\lambda_{\min}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right)} > \frac{n^2}{(n-m+1)^2} \cdot (n-m+1)^2\right\}$$
$$\leq \frac{1}{\sqrt{2\pi}}\left(\frac{6.414}{(1-\alpha)n}\right)^{(1-\alpha)n} \leq \frac{1}{\sqrt{2\pi}} e^{-\tau n}.$$

These bounds taken collectively imply that $\forall n > \max\left\{\frac{6.414 e^{\frac{\tau}{1-\alpha}}}{1-\alpha}, \left(\frac{\alpha}{1-\alpha-\frac{1}{n}} + 4\sqrt{\alpha\tau}\right)^3\right\}$,

$$\begin{aligned}
\mathbb{P}\left\{\lambda_{\min}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) < \frac{1}{n^{7/3}}\right\} \leq\ & \mathbb{P}\left\{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right) \leq \frac{1}{n^{1/3}}\right\} + \\
& \mathbb{P}\left\{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right) > \frac{1}{n^{1/3}} \text{ and } \lambda_{\min}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) < \frac{1}{n^{7/3}}\right\} \\
\leq\ & \mathbb{P}\left\{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right) \leq \frac{1}{n^{1/3}}\right\} + \mathbb{P}\left\{\frac{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right)}{\lambda_{\min}\left(\boldsymbol{A}\boldsymbol{A}^{\top}\right)} > n^2\right\} \\
\leq\ & 2e^{-\tau n} + \frac{1}{\sqrt{2\pi}} e^{-\tau n} < 3e^{-\tau n}.
\end{aligned}$$

Consequently, when $\epsilon = n^{-1/3}$,

$$\begin{aligned}
\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) \geq\ & \frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \sum_{i:\lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)<\frac{1}{n^{1/3}}}\left(\log\frac{1}{n^{1/3}} - \log\frac{1}{n^{7/3}}\right) \\
\geq\ & \frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \frac{\mathrm{card}\left\{i \mid \lambda_i\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) < \frac{1}{n^{1/3}}\right\}}{n} \cdot \log\left(n^2\right) \\
>\ & \frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \left(\frac{2\alpha}{1-\alpha-\frac{1}{n}} + 8\sqrt{\alpha\tau}\right) \cdot \frac{\log n}{n^{1/3}}
\end{aligned}$$

with probability exceeding $1 - 3\exp(-\tau n)$. Making use of (B.40) yields that for $\epsilon = n^{-1/3}$,

$$\mathbb{P}\left\{\left|\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \mathbb{E}\left[\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right]\right| > \frac{\sqrt{8\tau\alpha}}{n^{1/3}}\right\} < 2\exp(-\tau n).$$

Putting the above two bounds together implies that with probability exceeding $1 - 5\exp(-\tau n)$,

$$\begin{aligned}
\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) &\geq \frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right) - \left(\frac{2\alpha}{1-\alpha-\frac{1}{n}} + 8\sqrt{\alpha\tau}\right)\frac{\log n}{n^{1/3}} \\
&\geq \mathbb{E}\left[\frac{1}{n}\log\det{}^{\epsilon}\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \frac{\sqrt{8\tau\alpha}}{n^{1/3}} - \left(\frac{2\alpha}{1-\alpha-\frac{1}{n}} + 8\sqrt{\alpha\tau}\right)\frac{\log n}{n^{1/3}} \\
&> \mathbb{E}\left[\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}\boldsymbol{A}^{\top}\right)\right] - \left(\sqrt{8\tau\alpha} + \frac{2\alpha}{1-\alpha-\frac{1}{n}} + 8\sqrt{\alpha\tau}\right)\cdot\frac{\log n}{n^{1/3}} \\
&> (1-\alpha)\log\frac{1}{1-\alpha} - \alpha - \frac{6\alpha}{n^{2/3}} - \left(\frac{2\alpha}{1-\alpha-\frac{1}{n}} + 11\sqrt{\alpha\tau}\right)\frac{\log n}{n^{1/3}},
\end{aligned}$$

where the last inequality follows from (B.32).

B.10 Proof of Lemma 3.7

Suppose that the singular value decomposition of the real-valued $\boldsymbol{A}$ is given by $\boldsymbol{A} = \boldsymbol{U}_{\boldsymbol{A}}\begin{bmatrix}\boldsymbol{\Sigma}_{\boldsymbol{A}} \\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}_{\boldsymbol{A}}^{\top}$, where $\boldsymbol{\Sigma}_{\boldsymbol{A}}$ is a diagonal matrix containing all k singular values of $\boldsymbol{A}$. One can then write

$$\begin{aligned}
\log\det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{A}^{\top}\boldsymbol{B}^{-1}\boldsymbol{A}\right) &= \log\det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{V}_{\boldsymbol{A}}\begin{bmatrix}\boldsymbol{\Sigma}_{\boldsymbol{A}} & \boldsymbol{0}\end{bmatrix}\boldsymbol{U}_{\boldsymbol{A}}^{\top}\boldsymbol{B}^{-1}\boldsymbol{U}_{\boldsymbol{A}}\begin{bmatrix}\boldsymbol{\Sigma}_{\boldsymbol{A}} \\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}_{\boldsymbol{A}}^{\top}\right) \\
&= \log\det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{\Sigma}_{\boldsymbol{A}}\left(\tilde{\boldsymbol{B}}^{-1}\right)_{[k]}\boldsymbol{\Sigma}_{\boldsymbol{A}}\right) \\
&\geq \log\det\left(\frac{1}{n}\boldsymbol{\Sigma}_{\boldsymbol{A}}^{2}\right) - \log\det\left\{\frac{1}{n}\left(\tilde{\boldsymbol{B}}^{-1}\right)_{[k]}\right\} \qquad \text{(B.41)}
\end{aligned}$$

where $\tilde{\boldsymbol{B}} = \boldsymbol{U}_{\boldsymbol{A}}^{\top}\boldsymbol{B}\boldsymbol{U}_{\boldsymbol{A}} \sim \mathcal{W}_m\left(n-k, \boldsymbol{U}_{\boldsymbol{A}}^{\top}\boldsymbol{U}_{\boldsymbol{A}}\right) = \mathcal{W}_m\left(n-k, \boldsymbol{I}_m\right)$ from the property of Wishart distribution. Here, $\left(\tilde{\boldsymbol{B}}^{-1}\right)_{[k]}$ denotes the leading $k \times k$ minor consisting of matrix elements of $\tilde{\boldsymbol{B}}^{-1}$ in rows and columns from 1 to k, which is independent of $\boldsymbol{A}$ by Gaussianality.

Note that $\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{\Sigma}_{\boldsymbol{A}}^2\right) = \frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}^\top\boldsymbol{A}\right)$. Then Lemma 3.6 implies that for any $\tau > \frac{1}{n^{1/3}}$,

$$
\begin{aligned}
\frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{\Sigma}_{\boldsymbol{A}}^2\right) &= \frac{1}{n}\log\det\left(\frac{1}{n}\boldsymbol{A}^\top\boldsymbol{A}\right) = \frac{m}{n}\frac{1}{m}\log\det\left(\frac{1}{m}\boldsymbol{A}^\top\boldsymbol{A}\right) + \frac{1}{n}\log\det\left(\frac{m}{n}\boldsymbol{I}_k\right) \\
&\geq \alpha\left(-\left(1-\frac{\beta}{\alpha}\right)\log\left(1-\frac{\beta}{\alpha}\right) - \frac{\beta}{\alpha}\right) + \beta\log\alpha - \frac{c_8\tau\log n}{n^{1/3}} \\
&= -\left(\alpha-\beta\right)\log\left(\frac{\alpha-\beta}{\alpha}\right) - \beta + \beta\log\alpha - \frac{c_8\tau\log n}{n^{1/3}} \\
&= -\left(\alpha-\beta\right)\log\left(\alpha-\beta\right) - \beta + \alpha\log\alpha - \frac{c_8\tau\log n}{n^{1/3}},
\end{aligned}
\tag{B.42}
$$

with probability exceeding $1 - C_8\exp\left(-\tau^2 n\right)$ for some absolute constants $C_8, c_8 > 0$. On the other hand, it is well known (e.g. [138, Theorem 2.3.3]) that for a Wishart matrix $\tilde{\boldsymbol{B}} \sim \mathcal{W}_m\left(n-k, \boldsymbol{I}_m\right)$, $\left(\tilde{\boldsymbol{B}}^{-1}\right)_{[k]}^{-1}$ also follows the Wishart distribution, that is, $\left(\tilde{\boldsymbol{B}}^{-1}\right)_{[k]}^{-1} \sim \mathcal{W}_k\left(n-m, \boldsymbol{I}_k\right)$. Applying Lemma 3.6 again yields that

$$
\begin{aligned}
&\frac{1}{n}\log\det\left(\frac{1}{n}\left(\tilde{\boldsymbol{B}}^{-1}\right)_{[k]}^{-1}\right) = \frac{n-m}{n}\frac{1}{n-m}\log\det\left(\frac{1}{n-m}\left(\tilde{\boldsymbol{B}}^{-1}\right)_{[k]}^{-1}\right) + \frac{1}{n}\log\det\left(\frac{n-m}{n}\boldsymbol{I}_k\right) \\
&\quad\leq (1-\alpha)\left\{-\left(1-\frac{\beta}{1-\alpha}\right)\log\left(1-\frac{\beta}{1-\alpha}\right) - \frac{\beta}{1-\alpha}\right\} + \beta\log\left(1-\alpha\right) + \frac{c_9\tau}{\sqrt{n}} \\
&\quad= -\left(1-\alpha-\beta\right)\log\left(1-\frac{\beta}{1-\alpha}\right) - \beta + \beta\log\left(1-\alpha\right) + \frac{c_9\tau}{\sqrt{n}}
\end{aligned}
\tag{B.43}
$$

holds with probability exceeding $1 - C_9\exp\left(-\tau^2 n\right)$ for some universal constants $C_9, c_9 > 0$.

Combining (B.41), (B.42) and (B.41) suggests that for any $\tau > \frac{1}{n^{1/3}}$,

$$
\begin{aligned}
\frac{1}{n}\log\det\left(\epsilon\boldsymbol{I}_k + \boldsymbol{A}^\top\boldsymbol{B}^{-1}\boldsymbol{A}\right) \geq& -\left(\alpha-\beta\right)\log\left(\alpha-\beta\right) + \alpha\log\alpha + \left(1-\alpha-\beta\right)\log\left(1-\frac{\beta}{1-\alpha}\right) \\
&- \beta\log\left(1-\alpha\right) - \frac{c_{10}\tau\log n}{n^{1/3}}
\end{aligned}
$$

with probability exceeding $1 - C_{10}\exp\left(-\tau^2 n\right)$ for some constants $C_{10}, c_{10} > 0$.

Appendix C

Proofs of Theorems and Lemmas in Chapter 4

C.1 Bernstein Inequality

To simplify presentation, we state below a user-friendly version of Bernstein inequality, which is an immediate consequence of [140, Theorem 1.6].

Lemma C.1. *Consider m independent random matrices $\boldsymbol{M}_l$ ($1 \leq l \leq m$) of dimension $d_1 \times d_2$, each satisfying $\mathbb{E}[\boldsymbol{M}_l] = 0$ and $\|\boldsymbol{M}_l\| \leq B$. Define*

$$\sigma^2 := \max\left\{\left\|\sum_{l=1}^{m}\mathbb{E}\left[\boldsymbol{M}_l\boldsymbol{M}_l^*\right]\right\|, \left\|\sum_{l=1}^{m}\mathbb{E}\left[\boldsymbol{M}_l^*\boldsymbol{M}_l\right]\right\|\right\}.$$

Then there exists a universal constant $c_0 > 0$ such that for any integer $a \geq 2$,

$$\left\|\sum_{l=1}^{m}\boldsymbol{M}_l\right\| \leq c_0\left(\sqrt{a\sigma^2\log(d_1+d_2)} + aB\log(d_1+d_2)\right) \tag{C.1}$$

with probability at least $1 - (d_1+d_2)^{-a}$.

C.2 Proof of Lemma 4.1

Consider any valid perturbation $\boldsymbol{H}$ obeying $\mathcal{P}_{\Omega}(\boldsymbol{X}+\boldsymbol{H})=\mathcal{P}_{\Omega}(\boldsymbol{X})$, and denote by $\boldsymbol{H}_{\mathrm{e}}$ the enhanced form of $\boldsymbol{H}$. We note that the constraint requires $\mathcal{A}_{\Omega}'(\boldsymbol{H}_{\mathrm{e}})=0$ (or $\mathcal{A}_{\Omega}(\boldsymbol{H}_{\mathrm{e}})=0$) and $\mathcal{A}^{\perp}(\boldsymbol{H}_{\mathrm{e}})=0$. In addition, set $\boldsymbol{Z}_0=\mathcal{P}_{T^{\perp}}(\boldsymbol{B})$ for any $\boldsymbol{B}$ that satisfies $\langle \boldsymbol{B},\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\rangle=\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{*}$ and $\|\boldsymbol{B}\|\leq 1$. Therefore, $\boldsymbol{Z}_0\in T^{\perp}$ and $\|\boldsymbol{Z}_0\|\leq 1$, and hence $\boldsymbol{U}\boldsymbol{V}^{*}+\boldsymbol{Z}_0$ is a sub-gradient of the nuclear norm at $\boldsymbol{X}_{\mathrm{e}}$. We will establish this lemma by considering two scenarios separately.

(1) Consider first the case in which $\boldsymbol{H}_{\mathrm{e}}$ satisfies

$$\|\mathcal{P}_{T}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}\leq\frac{n_1^2n_2^2}{2}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}\,. \tag{C.2}$$

Since $\boldsymbol{U}\boldsymbol{V}^{*}+\boldsymbol{Z}_0$ is a sub-gradient of the nuclear norm at $\boldsymbol{X}_{\mathrm{e}}$, it follows that

$$\begin{aligned}
&\|\boldsymbol{X}_{\mathrm{e}}+\boldsymbol{H}_{\mathrm{e}}\|_{*}\geq\|\boldsymbol{X}_{\mathrm{e}}\|_{*}+\langle\boldsymbol{U}\boldsymbol{V}^{*}+\boldsymbol{Z}_0,\boldsymbol{H}_{\mathrm{e}}\rangle\\
&\quad=\|\boldsymbol{X}_{\mathrm{e}}\|_{*}+\langle\boldsymbol{W},\boldsymbol{H}_{\mathrm{e}}\rangle+\langle\boldsymbol{Z}_0,\boldsymbol{H}_{\mathrm{e}}\rangle-\langle\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^{*},\boldsymbol{H}_{\mathrm{e}}\rangle\\
&\quad=\|\boldsymbol{X}_{\mathrm{e}}\|_{*}+\left\langle\left(\mathcal{A}_{\Omega}'+\mathcal{A}^{\perp}\right)(\boldsymbol{W}),\boldsymbol{H}_{\mathrm{e}}\right\rangle+\langle\boldsymbol{Z}_0,\boldsymbol{H}_{\mathrm{e}}\rangle-\langle\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^{*},\boldsymbol{H}_{\mathrm{e}}\rangle && \text{(C.3)}\\
&\quad\geq\|\boldsymbol{X}_{\mathrm{e}}\|_{*}+\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{*}-\langle\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^{*},\boldsymbol{H}_{\mathrm{e}}\rangle && \text{(C.4)}
\end{aligned}$$

where (C.3) holds from (4.21), and (C.4) follows from the property of $\boldsymbol{Z}_0$ and the fact that $\left(\mathcal{A}_{\Omega}'+\mathcal{A}^{\perp}\right)(\boldsymbol{H}_{\mathrm{e}})=0$. The last term of (C.4) can be bounded as

$$\begin{aligned}
\langle\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^{*},\boldsymbol{H}_{\mathrm{e}}\rangle&=\langle\mathcal{P}_{T}(\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^{*}),\boldsymbol{H}_{\mathrm{e}}\rangle+\langle\mathcal{P}_{T^{\perp}}(\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^{*}),\boldsymbol{H}_{\mathrm{e}}\rangle\\
&\leq\|\mathcal{P}_{T}(\boldsymbol{W}-\boldsymbol{U}\boldsymbol{V}^{*})\|_{\mathrm{F}}\|\mathcal{P}_{T}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}+\|\mathcal{P}_{T^{\perp}}(\boldsymbol{W})\|\,\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{*}\\
&\leq\frac{1}{2n_1^2n_2^2}\|\mathcal{P}_{T}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}+\frac{1}{2}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{*},
\end{aligned}$$

where the last inequality follows from the assumptions (4.22) and (4.23). Plugging this into (C.4) yields

$$\begin{aligned}\|\boldsymbol{X}_{\mathrm{e}}+\boldsymbol{H}_{\mathrm{e}}\|_{*} &\geq \|\boldsymbol{X}_{\mathrm{e}}\|_{*}-\frac{1}{2n_1^2n_2^2}\|\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}+\frac{1}{2}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{*}\\ &\geq \|\boldsymbol{X}_{\mathrm{e}}\|_{*}-\frac{1}{4}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}+\frac{1}{2}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}} \qquad \text{(C.5)}\\ &\geq \|\boldsymbol{X}_{\mathrm{e}}\|_{*}+\frac{1}{4}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}\end{aligned}$$

where (C.5) follows from the inequality $\|\boldsymbol{M}\|_{*}\geq\|\boldsymbol{M}\|_{\mathrm{F}}$ and (C.2). Therefore, $\boldsymbol{X}_{\mathrm{e}}$ is the minimizer of EMaC.

We still need to prove the uniqueness of the minimizer. The inequality (C.5) implies that $\|\boldsymbol{X}_{\mathrm{e}}+\boldsymbol{H}_{\mathrm{e}}\|_{*}=\|\boldsymbol{X}_{\mathrm{e}}\|_{*}$ holds only when $\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}=0$. If $\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}=0$, then $\|\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}\leq\frac{n_1^2n_2^2}{2}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}=0$, and hence $\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})=\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})=0$, which only occurs when $\boldsymbol{H}_{\mathrm{e}}=0$. Hence, $\boldsymbol{X}_{\mathrm{e}}$ is the unique minimizer in this situation.

(2) On the other hand, consider the complement scenario where the following holds

$$\|\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}\geq\frac{n_1^2n_2^2}{2}\|\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}. \qquad \text{(C.6)}$$

We would first like to bound $\left\|\left(\frac{n_1n_2}{m}\mathcal{A}_{\Omega}+\mathcal{A}^{\perp}\right)\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\|_{\mathrm{F}}$ and $\left\|\left(\frac{n_1n_2}{m}\mathcal{A}_{\Omega}+\mathcal{A}^{\perp}\right)\mathcal{P}_{T^{\perp}}(\boldsymbol{H}_{\mathrm{e}})\right\|_{\mathrm{F}}$. The former term can be lower bounded by

$$\begin{aligned}&\left\|\left(\frac{n_1n_2}{m}\mathcal{A}_{\Omega}+\mathcal{A}^{\perp}\right)\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\|_{\mathrm{F}}^2\\ &=\left\langle\frac{n_1n_2}{m}\mathcal{A}_{\Omega}\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}}),\frac{n_1n_2}{m}\mathcal{A}_{\Omega}\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\rangle+\left\langle\mathcal{A}^{\perp}\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}}),\mathcal{A}^{\perp}\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\rangle\\ &\geq\left\langle\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}}),\frac{n_1n_2}{m}\mathcal{A}_{\Omega}\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\rangle+\left\langle\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}}),\mathcal{A}^{\perp}\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\rangle\\ &=\left\langle\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}}),\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\rangle+\left\langle\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}}),\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_{\Omega}\mathcal{P}_T-\mathcal{P}_T\mathcal{A}\mathcal{P}_T\right)\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\right\rangle\\ &\geq\|\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}^2-\left\|\mathcal{P}_T\mathcal{A}\mathcal{P}_T-\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_{\Omega}\mathcal{P}_T\right\|\|\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}^2\\ &\geq\left(1-\left\|\mathcal{P}_T\mathcal{A}\mathcal{P}_T-\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_{\Omega}\mathcal{P}_T\right\|\right)\|\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}^2\frac{1}{2}\|\mathcal{P}_T(\boldsymbol{H}_{\mathrm{e}})\|_{\mathrm{F}}^2. \qquad \text{(C.7)}\end{aligned}$$

On the other hand, since the operator norm of any projection operator is bounded above by 1, one can verify that

$$\left\| \frac{n_1 n_2}{m} \mathcal{A}_\Omega + \mathcal{A}^\perp \right\| \leq \frac{n_1 n_2}{m} \left(\left\| \mathcal{A}_{a_1} + \mathcal{A}^\perp \right\| + \sum_{i=2}^{m} \left\| \mathcal{A}_{a_i} \right\| \right) \leq n_1 n_2,$$

where a_i $(1 \leq i \leq m)$ are m uniform random indices that form Ω. This implies the following bound:

$$\left\| \left(\frac{n_1 n_2}{m} \mathcal{A}_\Omega + \mathcal{A}^\perp \right) \mathcal{P}_{T^\perp} \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} \leq n_1 n_2 \left\| \mathcal{P}_{T^\perp} \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} \leq \frac{2}{n_1 n_2} \left\| \mathcal{P}_T \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}},$$

where the last inequality arises from our assumption. Combining this with the above two bounds yields

$$\begin{aligned}
0 &= \left\| \left(\frac{n_1 n_2}{m} \mathcal{A}_\Omega + \mathcal{A}^\perp \right) \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} \\
&\geq \left\| \left(\frac{n_1 n_2}{m} \mathcal{A}_\Omega + \mathcal{A}^\perp \right) \mathcal{P}_T \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} - \left\| \left(\frac{n_1 n_2}{m} \mathcal{A}_\Omega + \mathcal{A}^\perp \right) \mathcal{P}_{T^\perp} \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} \\
&\geq \sqrt{\frac{1}{2}} \left\| \mathcal{P}_T \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} - \frac{2}{n_1 n_2} \left\| \mathcal{P}_T \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} \geq \frac{1}{2} \left\| \mathcal{P}_T \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} \geq \frac{n_1^2 n_2^2}{4} \left\| \mathcal{P}_{T^\perp} \left(\boldsymbol{H}_{\mathrm{e}} \right) \right\|_{\mathrm{F}} \geq 0,
\end{aligned}$$

which immediately indicates $\mathcal{P}_{T^\perp} \left(\boldsymbol{H}_{\mathrm{e}} \right) = 0$ and $\mathcal{P}_T \left(\boldsymbol{H}_{\mathrm{e}} \right) = 0$. Hence, (C.6) can only hold when $\boldsymbol{H}_{\mathrm{e}} = 0$.

C.3 Proof of Lemma 4.2

Since $\boldsymbol{U}$ (resp. $\boldsymbol{V}$) and $\boldsymbol{E}_{\mathrm{L}}$ (resp. $\boldsymbol{E}_{\mathrm{R}}$) determine the same column (resp. row) space, we can write

$$\boldsymbol{U}\boldsymbol{U}^* \boldsymbol{V}\boldsymbol{V}^* = \boldsymbol{E}_{\mathrm{L}} \left(\boldsymbol{E}_{\mathrm{L}}^* \boldsymbol{E}_{\mathrm{L}} \right)^{-1} \boldsymbol{E}_{\mathrm{L}}^*, \quad = \boldsymbol{E}_{\mathrm{R}}^* \left(\boldsymbol{E}_{\mathrm{R}} \boldsymbol{E}_{\mathrm{R}}^* \right)^{-1} \boldsymbol{E}_{\mathrm{R}},$$

and thus

$$\left\|\mathcal{P}_U\left(\boldsymbol{A}_{(k,l)}\right)\right\|_{\mathrm{F}}^2 \leq \left\|\boldsymbol{E}_{\mathrm{L}}\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)^{-1}\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{A}_{(k,l)}\right\|_{\mathrm{F}}^2 \leq \frac{1}{\sigma_{\min}\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)}\left\|\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{A}_{(k,l)}\right\|_{\mathrm{F}}^2,$$
$$\left\|\mathcal{P}_V\left(\boldsymbol{A}_{(k,l)}\right)\right\|_{\mathrm{F}}^2 \leq \left\|\boldsymbol{A}_{(k,l)}\boldsymbol{E}_{\mathrm{R}}^*\left(\boldsymbol{E}_{\mathrm{R}}\boldsymbol{E}_{\mathrm{R}}^*\right)^{-1}\boldsymbol{E}_{\mathrm{R}}\right\|_{\mathrm{F}}^2 \leq \frac{1}{\sigma_{\min}\left(\boldsymbol{E}_{\mathrm{R}}\boldsymbol{E}_{\mathrm{R}}^*\right)}\left\|\boldsymbol{A}_{(k,l)}\boldsymbol{E}_{\mathrm{R}}^*\right\|_{\mathrm{F}}^2.$$

Note that $\sqrt{\omega_{k,l}}\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{A}_{(k,l)}$ consists of $\omega_{k,l}$ columns of $\boldsymbol{E}_{\mathrm{L}}^*$ (and hence it contains $r\omega_{k,l}$ nonzero entries in total). Owing to the fact that each entry of $\boldsymbol{E}_{\mathrm{L}}^*$ has magnitude $\frac{1}{\sqrt{k_1k_2}}$, one can derive

$$\left\|\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{A}_{(k,l)}\right\|_{\mathrm{F}}^2 = \frac{1}{\omega_{k,l}}\cdot r\omega_{k,l}\cdot\frac{1}{k_1k_2} = \frac{r}{k_1k_2} \leq \frac{rc_{\mathrm{s}}}{n_1n_2}.$$

A similar argument yields $\left\|\boldsymbol{A}_{(k,l)}\boldsymbol{E}_{\mathrm{R}}^*\right\|_{\mathrm{F}}^2 \leq \frac{c_{\mathrm{s}}r}{n_1n_2}$. Combining $\sigma_{\min}\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right) \geq \frac{1}{\mu_1}$ and $\sigma_{\min}\left(\boldsymbol{E}_{\mathrm{R}}\boldsymbol{E}_{\mathrm{R}}^*\right) \geq \frac{1}{\mu_1}$, (4.24) follows by plugging these facts into the above equations.

To show (4.25), since $\left|\langle\boldsymbol{A}_{\boldsymbol{b}},\mathcal{P}_T\left(\boldsymbol{A}_{\boldsymbol{a}}\right)\rangle\right| = \left|\langle\mathcal{P}_T\left(\boldsymbol{A}_{\boldsymbol{b}}\right),\boldsymbol{A}_{\boldsymbol{a}}\rangle\right|$, we only need to examine the situation where $\omega_{\boldsymbol{b}} < \omega_{\boldsymbol{a}}$. Observe that

$$\left|\langle\boldsymbol{A}_{\boldsymbol{b}},\mathcal{P}_T\boldsymbol{A}_{\boldsymbol{a}}\rangle\right| \leq \left|\langle\boldsymbol{A}_{\boldsymbol{b}},\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}_{\boldsymbol{a}}\rangle\right| + \left|\langle\boldsymbol{A}_{\boldsymbol{b}},\boldsymbol{A}_{\boldsymbol{a}}\boldsymbol{V}\boldsymbol{V}^*\rangle\right| + \left|\langle\boldsymbol{A}_{\boldsymbol{b}},\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}_{\boldsymbol{a}}\boldsymbol{V}\boldsymbol{V}^*\rangle\right|.$$

Owing to the multi-fold Hankel structure of $\boldsymbol{A}_{\boldsymbol{a}}$, the matrix $\boldsymbol{U}\boldsymbol{U}^*\sqrt{\omega_{\boldsymbol{a}}}\boldsymbol{A}_{\boldsymbol{a}}$ consists of $\omega_{\boldsymbol{a}}$ columns of $\boldsymbol{U}\boldsymbol{U}^*$. Since there are only $\omega_{\boldsymbol{b}}$ nonzero entries in $\boldsymbol{A}_{\boldsymbol{b}}$ each of magnitude $\frac{1}{\sqrt{\omega_{\boldsymbol{b}}}}$, we can derive

$$\begin{aligned}\left|\langle\boldsymbol{A}_{\boldsymbol{b}},\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}_{\boldsymbol{a}}\rangle\right| &\leq \left\|\boldsymbol{A}_{\boldsymbol{b}}\right\|_1\left\|\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}_{\boldsymbol{a}}\right\|_\infty = \frac{\omega_{\boldsymbol{b}}}{\sqrt{\omega_{\boldsymbol{b}}}}\cdot\max_{\alpha,\beta}\left|\left(\boldsymbol{U}\boldsymbol{U}^*\boldsymbol{A}_{\boldsymbol{a}}\right)_{\alpha,\beta}\right| \\ &\leq \sqrt{\frac{\omega_{\boldsymbol{b}}}{\omega_{\boldsymbol{a}}}}\max_{\alpha,\beta}\left|\left(\boldsymbol{U}\boldsymbol{U}^*\right)_{\alpha,\beta}\right|.\end{aligned}$$

Each entry of $\boldsymbol{U}\boldsymbol{U}^*$ is bounded in magnitude by

$$\begin{aligned}\left|\left(\boldsymbol{U}\boldsymbol{U}^*\right)_{k,l}\right| &= \left|\boldsymbol{e}_k^\top\boldsymbol{E}_{\mathrm{L}}\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)^{-1}\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{e}_l\right| \leq \left\|\boldsymbol{e}_k^\top\boldsymbol{E}_{\mathrm{L}}\right\|_{\mathrm{F}}\left\|\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)^{-1}\right\|\left\|\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{e}_l\right\|_{\mathrm{F}} \\ &\leq \frac{r}{k_1k_2}\frac{1}{\sigma_{\min}\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)} \leq \frac{\mu_1c_{\mathrm{s}}r}{n_1n_2},\end{aligned} \tag{C.8}$$

which immediately implies that

$$|\langle \boldsymbol{A_b}, \boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A_a}\rangle| \leq \sqrt{\frac{\omega_{\boldsymbol{b}}}{\omega_{\boldsymbol{a}}}} \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{C.9}$$

Similarly, one can derive

$$|\langle \boldsymbol{A_b}, \boldsymbol{A_a}\boldsymbol{V}\boldsymbol{V}^*\rangle| \leq \sqrt{\frac{\omega_{\boldsymbol{b}}}{\omega_{\boldsymbol{a}}}} \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{C.10}$$

We still need to bound the magnitude of $\langle \boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A_a} \boldsymbol{V}\boldsymbol{V}^*, \boldsymbol{A_b}\rangle$. One can observe that for the kth row of $\boldsymbol{U}\boldsymbol{U}^*$:

$$\left\|\boldsymbol{e}_k^\top \boldsymbol{U}\boldsymbol{U}^*\right\|_{\mathrm{F}} \leq \left\|\boldsymbol{e}_k^\top \boldsymbol{E}_{\mathrm{L}} \left(\boldsymbol{E}_{\mathrm{L}}^* \boldsymbol{E}_{\mathrm{L}}\right)^{-1} \boldsymbol{E}_{\mathrm{L}}^*\right\|_{\mathrm{F}} \leq \left\|\boldsymbol{e}_k^\top \boldsymbol{E}_{\mathrm{L}}\right\|_{\mathrm{F}} \left\|\left(\boldsymbol{E}_{\mathrm{L}}^* \boldsymbol{E}_{\mathrm{L}}\right)^{-1} \boldsymbol{E}_{\mathrm{L}}^*\right\| \leq \sqrt{\frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}}.$$

Similarly, for the lth column of $\boldsymbol{V}\boldsymbol{V}^*$, one has $\|\boldsymbol{V}\boldsymbol{V}^* \boldsymbol{e}_l\|_{\mathrm{F}} \leq \sqrt{\frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}}$. The magnitude of the entries of $\boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A_a} \boldsymbol{V}\boldsymbol{V}^*$ can now be bounded by

$$\left|\left(\boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A_a} \boldsymbol{V}\boldsymbol{V}^*\right)_{k,l}\right| \leq \|\boldsymbol{A_a}\| \left\|\boldsymbol{e}_k^\top \boldsymbol{U}\boldsymbol{U}^*\right\|_{\mathrm{F}} \|\boldsymbol{V}\boldsymbol{V}^* \boldsymbol{e}_l\|_{\mathrm{F}} \leq \frac{1}{\sqrt{\omega_{\boldsymbol{a}}}} \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2},$$

where we used $\|\boldsymbol{A_a}\| = 1/\sqrt{\omega_{\boldsymbol{a}}}$. Since $\boldsymbol{A_b}$ has only $\omega_{\boldsymbol{b}}$ nonzero entries each having magnitude $\frac{1}{\sqrt{\omega_{\boldsymbol{b}}}}$, one can verify that

$$|\langle \boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A_a} \boldsymbol{V}\boldsymbol{V}^*, \boldsymbol{A_b}\rangle| \leq \left(\max_{k,l} \left|\left(\boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A_a} \boldsymbol{V}\boldsymbol{V}^*\right)_{k,l}\right|\right) \cdot \frac{\omega_{\boldsymbol{b}}}{\sqrt{\omega_{\boldsymbol{b}}}} = \sqrt{\frac{\omega_{\boldsymbol{b}}}{\omega_{\boldsymbol{a}}}} \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{C.11}$$

The above bounds (C.9), (C.10) and (C.11) taken together lead to (4.25).

C.4 Proof of Lemma 4.3

Define a family of operators

$$\mathcal{Z}_{(k,l)} := \frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{A}_{(k,l)} \mathcal{P}_T - \frac{1}{m} \mathcal{P}_T \mathcal{A} \mathcal{P}_T.$$

for any $(k,l) \in [n_1] \times [n_2]$. For any matrix $\boldsymbol{M}$, we can compute

$$\mathcal{P}_T \mathcal{A}_{(k,l)} \mathcal{P}_T (\boldsymbol{M}) = \mathcal{P}_T \left(\left\langle \boldsymbol{A}_{(k,l)}, \mathcal{P}_T \boldsymbol{M} \right\rangle \boldsymbol{A}_{(k,l)} \right) = \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right) \left\langle \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right), \boldsymbol{M} \right\rangle \tag{C.12}$$

and hence

$$\begin{aligned} \left(\mathcal{P}_T \mathcal{A}_{(k,l)} \mathcal{P}_T \right)^2 (\boldsymbol{M}) &= \left[\mathcal{P}_T \mathcal{A}_{(k,l)} \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right) \right] \left\langle \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right), \boldsymbol{M} \right\rangle \\ &= \left\langle \boldsymbol{A}_{(k,l)}, \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right) \right\rangle \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right) \left\langle \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right), \boldsymbol{M} \right\rangle \\ &= \left\| \mathcal{P}_T \left(\boldsymbol{A}_{(k,l)} \right) \right\|_{\mathrm{F}}^2 \mathcal{P}_T \mathcal{A}_{(k,l)} \mathcal{P}_T (\boldsymbol{M}) \leq \frac{2\mu_1 c_{\mathrm{s}} r}{n_1 n_2} \mathcal{P}_T \mathcal{A}_{(k,l)} \mathcal{P}_T (\boldsymbol{M}), \end{aligned}$$

where the last inequality follows from (4.26). This further gives

$$\left\| \mathcal{P}_T \mathcal{A}_{(k,l)} \mathcal{P}_T \right\| \leq \frac{2\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{C.13}$$

Let $\boldsymbol{a}_i$ $(1 \leq i \leq m)$ be m independent indices uniformly drawn from $[n_1] \times [n_2]$, then we have $\mathbb{E}\left[\mathcal{Z}_{\boldsymbol{a}_i}\right] = 0$ and

$$\left\| \mathcal{Z}_{\boldsymbol{a}_i} \right\| \leq 2 \max_{(k,l) \in [n_1] \times [n_2]} \frac{n_1 n_2}{m} \left\| \mathcal{P}_T \boldsymbol{A}_{(k,l)} \mathcal{P}_T \right\| \leq \frac{4\mu_1 c_{\mathrm{s}} r}{m}$$

following from (C.13). Further,

$$\mathbb{E}\left[\mathcal{Z}_{\boldsymbol{a}_i}^2\right] = \mathbb{E}\left(\frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{A}_{\boldsymbol{a}_i} \mathcal{P}_T \right)^2 - \left(\mathbb{E}\left[\frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{A}_{\boldsymbol{a}_i} \mathcal{P}_T \right] \right)^2 = \frac{n_1^2 n_2^2}{m^2} \mathbb{E}\left(\mathcal{P}_T \mathcal{A}_{\boldsymbol{a}_i} \mathcal{P}_T \right)^2 - \frac{1}{m^2} \left(\mathcal{P}_T \mathcal{A} \mathcal{P}_T \right)^2,$$

We can then bound the operator norm as

$$\begin{aligned} \sum_{i=1}^m \left\| \mathbb{E}\left[\mathcal{Z}_{\boldsymbol{a}_i}^2\right] \right\| &\leq \sum_{i=1}^m \frac{n_1^2 n_2^2}{m^2} \left\| \mathbb{E}\left(\mathcal{P}_T \mathcal{A}_{\boldsymbol{a}_i} \mathcal{P}_T \right)^2 \right\| + \frac{1}{m} \left\| \left(\mathcal{P}_T \mathcal{A} \mathcal{P}_T \right)^2 \right\| \\ &\leq \frac{n_1^2 n_2^2}{m} \frac{2\mu_1 c_{\mathrm{s}} r}{n_1 n_2} \left\| \mathbb{E}\left[\mathcal{P}_T \mathcal{A}_{\boldsymbol{a}_i} \mathcal{P}_T \right] \right\| + \frac{1}{m} \qquad \text{(C.14)} \\ &= \frac{2\mu_1 c_{\mathrm{s}} r n_1 n_2}{m} \frac{1}{n_1 n_2} \left\| \mathcal{P}_T \mathcal{A} \mathcal{P}_T \right\| + \frac{1}{m^2} \leq \frac{4\mu_1 c_{\mathrm{s}} r}{m}, \qquad \text{(C.15)} \end{aligned}$$

where (C.14) uses (C.13). Applying Lemma C.1 yields that there exists some constant $0 < \epsilon \leq \frac{1}{2}$ such that $\|\sum_{i=1}^{m} \mathcal{Z}_{\boldsymbol{a}_i}\| \leq \epsilon$ with probability exceeding $1-(n_1n_2)^{-4}$, provided that $m > c_1\mu_1 c_{\mathrm{s}} r \log(n_1 n_2)$ for some universal constant $c_1 > 0$.

C.5 Proof of Lemma 4.4

Suppose that $\mathcal{A}_\Omega = \sum_{i=1}^{m} \mathcal{A}_{\boldsymbol{a}_i}$, where $\boldsymbol{a}_i$, $1 \leq i \leq m$, are m independent indices drawn uniformly at random from $[n_1] \times [n_2]$. Define

$$\boldsymbol{S}_{(k,l)} := \frac{n_1 n_2}{m} \mathcal{A}_{(k,l)}(\boldsymbol{M}) - \frac{1}{m}\mathcal{A}(\boldsymbol{M}), \quad (k,l) \in [n_1] \times [n_2],$$

which obeys $\mathbb{E}[\boldsymbol{S}_{\boldsymbol{a}_i}] = \boldsymbol{0}$ and $\left(\frac{n_1n_2}{m}\mathcal{A}_\Omega - \mathcal{A}\right)(\boldsymbol{M}) := \sum_{i=1}^{m} \boldsymbol{S}_{\boldsymbol{a}_i}$.

In order to apply Lemma C.1, one needs to bound $\left\|\mathbb{E}\left[\sum_{i=1}^{m} \boldsymbol{S}_{\boldsymbol{a}_i}\boldsymbol{S}_{\boldsymbol{a}_i}^*\right]\right\|$ and $\|\boldsymbol{S}_{\boldsymbol{a}_i}\|$, which we tackle separately in the sequel. Observe that

$$\begin{aligned}
\boldsymbol{0} \preceq \boldsymbol{S}_{(k,l)}\boldsymbol{S}_{(k,l)}^* &= \left(\frac{n_1n_2}{m}\mathcal{A}_{(k,l)}(\boldsymbol{M}) - \frac{1}{m}\mathcal{A}(\boldsymbol{M})\right) \cdot \left(\frac{n_1n_2}{m}\mathcal{A}_{(k,l)}(\boldsymbol{M}) - \frac{1}{m}\mathcal{A}(\boldsymbol{M})\right)^* \\
&\preceq \left(\frac{n_1n_2}{m}\right)^2 \mathcal{A}_{(k,l)}(\boldsymbol{M})\left(\mathcal{A}_{(k,l)}(\boldsymbol{M})\right)^* = \left(\frac{n_1n_2}{m}\right)^2 \left|\left\langle \boldsymbol{A}_{(k,l)}, \boldsymbol{M}\right\rangle\right|^2 \boldsymbol{A}_{(k,l)} \cdot \boldsymbol{A}_{(k,l)}^\top \\
&\preceq \left(\frac{n_1n_2}{m}\right)^2 \frac{\left|\left\langle \boldsymbol{A}_{(k,l)}, \boldsymbol{M}\right\rangle\right|^2}{\omega_{k,l}} \boldsymbol{I},
\end{aligned}$$

where the first inequality follows since $\frac{1}{m}\sum_{k,l}\mathcal{A}_{(k,l)}(\boldsymbol{M}) = \frac{1}{m}\mathcal{A}(\boldsymbol{M})$, and the last inequality arises from the fact that all non-zero entries of $\boldsymbol{A}_{(k,l)} \cdot \boldsymbol{A}_{(k,l)}^\top$ lie on its diagonal and are bounded in magnitude by $\frac{1}{\omega_{k,l}}$. This immediately suggests that

$$\begin{aligned}
&\left\|\mathbb{E}\left[\sum_{i=1}^{m} \boldsymbol{S}_{\boldsymbol{a}_i}\boldsymbol{S}_{\boldsymbol{a}_i}^*\right]\right\| = \frac{m}{n_1n_2}\left\|\sum_{(k,l)\in[n_1]\times[n_2]} \boldsymbol{S}_{(k,l)}\boldsymbol{S}_{(k,l)}^*\right\| \\
&\quad \leq \frac{m}{n_1n_2}\left\|\left(\frac{n_1n_2}{m}\right)^2 \left(\sum_{(k,l)\in[n_1]\times[n_2]} \frac{\left|\left\langle \boldsymbol{A}_{(k,l)}, \boldsymbol{M}\right\rangle\right|^2}{\omega_{k,l}}\right)\boldsymbol{I}\right\| = \frac{n_1n_2}{m}\|\boldsymbol{M}\|_{\mathcal{A},2}^2, \quad \text{(C.16)}
\end{aligned}$$

where the last equality follows from the definition of $\|\boldsymbol{M}\|_{\mathcal{A},2}$. Following the same argument, one can derive the same bound for $\left\|\mathbb{E}\left[\sum_{i=1}^{m}\boldsymbol{S}_{\boldsymbol{a}_i}^{*}\boldsymbol{S}_{\boldsymbol{a}_i}\right]\right\|$ as well.

On the other hand, the operator norm of each $\boldsymbol{S}_{(k,l)}$ can be bounded as follows

$$\begin{aligned}
\left\|\boldsymbol{S}_{(k,l)}\right\| \leq \left\|\frac{n_1 n_2}{m}\mathcal{A}_{(k,l)}\left(\boldsymbol{M}\right)\right\| + \left\|\frac{1}{m}\mathcal{A}\left(\boldsymbol{M}\right)\right\| &\leq 2\max_{(k,l)\in[n_1]\times[n_2]}\left\|\frac{n_1 n_2}{m}\mathcal{A}_{(k,l)}\left(\boldsymbol{M}\right)\right\| \\
&= \frac{2n_1 n_2}{m}\max_{(k,l)\in[n_1]\times[n_2]}\left\|\left\langle \boldsymbol{A}_{(k,l)},\boldsymbol{M}\right\rangle \boldsymbol{A}_{(k,l)}\right\| \qquad\qquad \text{(C.17)}\\
&= \frac{2n_1 n_2}{m}\max_{(k,l)\in[n_1]\times[n_2]}\left|\frac{\left\langle \boldsymbol{A}_{(k,l)},\boldsymbol{M}\right\rangle}{\sqrt{\omega_{k,l}}}\right| = \frac{2n_1 n_2}{m}\left\|\boldsymbol{M}\right\|_{\mathcal{A},\infty},
\end{aligned}$$

where (C.17) holds since $\left\|\boldsymbol{A}_{(k,l)}\right\| = \frac{1}{\sqrt{\omega_{k,l}}}$ and the last equality follows by applying the definition of $\|\cdot\|_{\mathcal{A},\infty}$.

Finally, we combine the above two bounds together with Bernstein inequality (Lemma C.1) to obtain

$$\left\|\left(\frac{n_1 n_2}{m}\mathcal{A}_{\Omega}-\mathcal{A}\right)\left(\boldsymbol{M}\right)\right\| \leq c_2\sqrt{\frac{n_1 n_2\log\left(n_1 n_2\right)}{m}}\left\|\boldsymbol{M}\right\|_{\mathcal{A},2} + c_2\frac{2n_1 n_2\log\left(n_1 n_2\right)}{m}\left\|\boldsymbol{M}\right\|_{\mathcal{A},\infty}$$

with high probability, where $c_2 > 0$ is some absolute constant.

C.6 Proof of Lemma 4.5

Write $\mathcal{A}_{\Omega} = \sum_{i=1}^{m}\mathcal{A}_{\boldsymbol{a}_i}$, where $\boldsymbol{a}_i$ $(1\leq i\leq m)$ are m independent indices uniformly drawn from $[n_1]\times[n_2]$. By the definition of $\|\boldsymbol{M}\|_{\mathcal{A},2}$, we need to examine the components

$$\frac{1}{\sqrt{\omega_{k,l}}}\left\langle \boldsymbol{A}_{(k,l)},\left(\frac{n_1 n_2}{m}\mathcal{P}_T\mathcal{A}_{\Omega}-\mathcal{P}_T\mathcal{A}\right)\left(\boldsymbol{M}\right)\right\rangle, \quad \forall (k,l)\in[n_1]\times[n_2].$$

Define a set of variables $z_{(\alpha,\beta)}$'s to be

$$z_{(\alpha,\beta)}^{(k,l)} := \frac{1}{\sqrt{\omega_{k,l}}}\left\langle \boldsymbol{A}_{(k,l)},\frac{n_1 n_2}{m}\mathcal{P}_T\mathcal{A}_{(\alpha,\beta)}\left(\boldsymbol{M}\right)-\frac{1}{m}\mathcal{P}_T\mathcal{A}\left(\boldsymbol{M}\right)\right\rangle, \qquad \text{(C.18)}$$

thus resulting in

$$\frac{1}{\sqrt{\omega_{k,l}}}\left\langle \boldsymbol{A}_{(k,l)}, \left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega - \mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\rangle := \sum_{i=1}^{m} z_{\boldsymbol{a}_i}^{(k,l)}.$$

The definition of $\|\boldsymbol{M}\|_{\mathcal{A},2}$ allows us to express

$$\left\|\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega - \mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},2} = \left\|\sum_{i=1}^{m}\boldsymbol{z}_{\boldsymbol{a}_i}\right\|_2, \tag{C.19}$$

where $\boldsymbol{z}_{(\alpha,\beta)}$'s are defined to be n_1n_2-dimensional vectors

$$\boldsymbol{z}_{(\alpha,\beta)} := \left[z_{(\alpha,\beta)}^{(k,l)}\right]_{(k,l)\in[n_1]\times[n_2]}, \quad (\alpha,\beta)\in[n_1]\times[n_2].$$

For any random vector $\boldsymbol{v}\in\mathcal{V}$, one can easily bound $\|\boldsymbol{v}-\mathbb{E}\boldsymbol{v}\|_2 \leq 2\sup_{\tilde{\boldsymbol{v}}\in\mathcal{V}}\|\tilde{\boldsymbol{v}}\|_2$. Observing that $\mathbb{E}\left[\boldsymbol{z}_{(\alpha,\beta)}\right]=\boldsymbol{0}$, we can bound

$$\begin{aligned}
\left\|\boldsymbol{z}_{(\alpha,\beta)}\right\|_2 &\leq 2\sqrt{\sum_{k,l}\frac{1}{\omega_{k,l}}\left|\left\langle \boldsymbol{A}_{(k,l)}, \frac{2n_1n_2}{m}\mathcal{P}_T\mathcal{A}_{(\alpha,\beta)}(\boldsymbol{M})\right\rangle\right|^2} \\
&= \frac{2n_1n_2}{m}\sqrt{\sum_{k,l}\frac{1}{\omega_{k,l}}\left|\left\langle \boldsymbol{A}_{(k,l)}, \mathcal{P}_T\left(\boldsymbol{A}_{(\alpha,\beta)}\right)\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right\rangle\right|^2} \\
&= \frac{2n_1n_2}{m}\frac{\left|\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right|}{\sqrt{\omega_{\alpha,\beta}}}\sqrt{\sum_{k,l}\frac{\omega_{\alpha,\beta}\left|\left\langle \boldsymbol{A}_{(k,l)}, \mathcal{P}_T\left(\boldsymbol{A}_{(\alpha,\beta)}\right)\right\rangle\right|^2}{\omega_{k,l}}} \\
&\leq \frac{2n_1n_2}{m}\frac{\left|\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right|}{\sqrt{\omega_{\alpha,\beta}}}\sqrt{\frac{\mu_5 r}{n_1n_2}} = 2\sqrt{\frac{n_1n_2}{m}\cdot\frac{\mu_5 r}{m}}\frac{\left|\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right|}{\sqrt{\omega_{\alpha,\beta}}},
\end{aligned} \tag{C.20}$$

where (C.20) follows from the definition of μ_5 in (4.34). Now it follows that

$$\|\boldsymbol{z}_{\boldsymbol{a}_i}\|_2 \leq \max_{\alpha,\beta}\left\|\boldsymbol{z}_{(\alpha,\beta)}\right\|_2 \leq \max_{\alpha,\beta} 2\sqrt{\frac{n_1n_2}{m}\cdot\frac{\mu_5 r}{m}}\frac{\left|\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right|}{\sqrt{\omega_{\alpha,\beta}}} \leq 2\sqrt{\frac{n_1n_2}{m}\cdot\frac{\mu_5 r}{m}}\left\|\boldsymbol{M}\right\|_{\mathcal{A},\infty}, \tag{C.21}$$

where (C.21) follows from (4.31). On the other hand,

$$\left|\mathbb{E}\left[\sum_{i=1}^{m}\boldsymbol{z}_{\boldsymbol{a}_i}^{*}\boldsymbol{z}_{\boldsymbol{a}_i}\right]\right| = \frac{m}{n_1n_2}\sum_{\alpha,\beta}\|\boldsymbol{z}_{(\alpha,\beta)}\|_2^2 \le \frac{m}{n_1n_2}\sum_{\alpha,\beta}4\frac{n_1n_2}{m}\cdot\frac{\mu_5 r}{m}\frac{\left|\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right|^2}{\omega_{\alpha,\beta}} = \frac{4\mu_5 r}{m}\left\|\boldsymbol{M}\right\|_{\mathcal{A},2}^2,$$

which again follows from (4.32). Since $\boldsymbol{z}_{\boldsymbol{a}_i}$'s are vectors, we immediately obtain $\left\|\mathbb{E}\left[\sum_{i=1}^{m}\boldsymbol{z}_{\boldsymbol{a}_i}\boldsymbol{z}_{\boldsymbol{a}_i}^{*}\right]\right\| = \left|\mathbb{E}\left[\sum_{i=1}^{m}\boldsymbol{z}_{\boldsymbol{a}_i}^{*}\boldsymbol{z}_{\boldsymbol{a}_i}\right]\right|$. Applying Lemma C.1 then suggests that

$$\begin{aligned}\left\|\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega - \mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},2} &\le c_3\sqrt{\frac{\mu_5 r\log(n_1n_2)}{m}}\left\|\boldsymbol{M}\right\|_{\mathcal{A},2}\\ &+ c_3\sqrt{\frac{n_1n_2}{m}\cdot\frac{\mu_5 r}{m}}\log(n_1n_2)\left\|\boldsymbol{M}\right\|_{\mathcal{A},\infty}\end{aligned}$$

with high probability for some numerical constant $c_3 > 0$, which completes the proof.

C.7 Proof of Lemma 4.6

From Appendix C.6, it is straightforward that

$$\left\|\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega - \mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},\infty} = \max_{k,l}\left|\sum_{i=1}^{m} z_{\boldsymbol{a}_i}^{(k,l)}\right|, \tag{C.22}$$

where $z_{\boldsymbol{a}_i}^{(k,l)}$'s are defined as (C.18). Using similar techniques as (C.20), we can obtain

$$\begin{aligned}\left|z_{(\alpha,\beta)}^{(k,l)}\right| &\le 2\max_{k,l}\frac{\left|\left\langle \boldsymbol{A}_{(k,l)},\frac{n_1n_2}{m}\mathcal{P}_T\left(\boldsymbol{A}_{(\alpha,\beta)}\right)\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right\rangle\right|}{\sqrt{\omega_{k,l}}}\\ &\le 2\max_{k,l}\left(\frac{1}{\sqrt{\omega_{k,l}}}\sqrt{\frac{\omega_{k,l}}{\omega_{\alpha,\beta}}}\frac{3\mu_1 c_s r}{n_1n_2}\right)\frac{n_1n_2}{m}\left|\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right| = \frac{6\mu_1 c_s r}{m}\frac{1}{\sqrt{\omega_{\alpha,\beta}}}\left|\left\langle \boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right|,\end{aligned}$$

where we have made use of the fact (4.25). As a result, one has

$$\left|z_{(\alpha,\beta)}^{(k,l)}\right| \le \frac{6\mu_1 c_s r}{m}\|\boldsymbol{M}\|_{\mathcal{A},\infty}, \quad \text{and}$$

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{m}|z_{\boldsymbol{a}_i}^{(k,l)}|^2\right] &= \frac{m}{n_1n_2}\sum_{\alpha,\beta}\left|z_{(\alpha,\beta)}^{(k,l)}\right|^2 \leq \frac{m}{n_1n_2}\left(\frac{6\mu_1c_sr}{m}\right)^2\sum_{\alpha,\beta}\frac{1}{\omega_{\alpha,\beta}}\left|\left\langle\boldsymbol{A}_{(\alpha,\beta)},\boldsymbol{M}\right\rangle\right|^2 \\ &= \frac{36\mu_1^2c_s^2r^2}{mn_1n_2}\|\boldsymbol{M}\|_{\mathcal{A},2}^2.\end{aligned}$$

The Bernstein inequality in Lemma C.1 taken collectively with the union bound yields that

$$\begin{aligned}\left\|\left(\frac{n_1n_2}{m}\mathcal{P}_T\mathcal{A}_\Omega-\mathcal{P}_T\mathcal{A}\right)(\boldsymbol{M})\right\|_{\mathcal{A},\infty} \leq{}& c_4\sqrt{\frac{\mu_1c_{\mathrm{s}}r\log(n_1n_2)}{m}}\cdot\sqrt{\frac{\mu_1c_{\mathrm{s}}r}{n_1n_2}}\|\boldsymbol{M}\|_{\mathcal{A},2} \\ &+c_4\frac{\mu_1c_{\mathrm{s}}r\log(n_1n_2)}{m}\|\boldsymbol{M}\|_{\mathcal{A},\infty}\end{aligned}$$

with high probability for some constant $c_4>0$, completing the proof.

C.8 Proof of Lemma 4.7

To bound $\|\boldsymbol{U}\boldsymbol{V}^*\|_{\mathcal{A},\infty}$, observe that there exists a unitary matrix $\boldsymbol{B}$ such that

$$\boldsymbol{U}\boldsymbol{V}^*=\boldsymbol{E}_{\mathrm{L}}\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)^{-\frac{1}{2}}\boldsymbol{B}\left(\boldsymbol{E}_{\mathrm{R}}\boldsymbol{E}_{\mathrm{R}}^*\right)^{-\frac{1}{2}}\boldsymbol{E}_{\mathrm{R}}.$$

For any $(k,l)\in[n_1]\times[n_2]$, we can then bound

$$\begin{aligned}\left|(\boldsymbol{U}\boldsymbol{V}^*)_{k,l}\right| &\leq \left\|\boldsymbol{e}_k^\top\boldsymbol{E}_{\mathrm{L}}\right\|_{\mathrm{F}}\left\|\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)^{-\frac{1}{2}}\right\|\|\boldsymbol{B}\|\left\|\left(\boldsymbol{E}_{\mathrm{R}}^*\boldsymbol{E}_{\mathrm{R}}\right)^{-\frac{1}{2}}\right\|\|\boldsymbol{E}_{\mathrm{R}}\boldsymbol{e}_l\|_{\mathrm{F}} \\ &\leq \sqrt{\frac{r}{k_1k_2}}\mu_1\sqrt{\frac{r}{(n_1-k_1+1)(n_2-k_2+1)}} \leq \frac{\mu_1c_{\mathrm{s}}r}{n_1n_2}.\end{aligned}$$

Since $\boldsymbol{A}_{(k,l)}$ has only $\omega_{k,l}$ nonzero entries each of magnitude $\frac{1}{\sqrt{\omega_{k,l}}}$, this leads to

$$\|\boldsymbol{U}\boldsymbol{V}^*\|_{\mathcal{A},\infty}=\frac{1}{\omega_{k,l}}\left|\sum_{(\alpha,\beta)\in\Omega_{\mathrm{e}}(k,l)}(\boldsymbol{U}\boldsymbol{V}^*)_{\alpha,\beta}\right| \leq \max_{k,l}\left|(\boldsymbol{U}\boldsymbol{V}^*)_{k,l}\right| \leq \frac{\mu_1c_{\mathrm{s}}r}{n_1n_2}.$$

The rest is to bound $\|\boldsymbol{U}\boldsymbol{V}^*\|_{\mathcal{A},2}$ and $\left\|\mathcal{P}_T\left(\sqrt{\omega_{k,l}}\boldsymbol{A}_{(k,l)}\right)\right\|_{\mathcal{A},2}$. Observe that the ith row of $\boldsymbol{U}\boldsymbol{V}^*$ obeys

$$\begin{aligned}\left\|\boldsymbol{e}_i^\top\boldsymbol{U}\boldsymbol{V}^*\right\|_{\mathrm{F}}^2 = \left\|\boldsymbol{e}_i^\top\boldsymbol{U}\right\|_{\mathrm{F}}^2 = \left\|\boldsymbol{e}_i^\top\boldsymbol{E}_{\mathrm{L}}\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)^{-\frac{1}{2}}\right\|_{\mathrm{F}}^2 &\le \left\|\boldsymbol{e}_i^\top\boldsymbol{E}_{\mathrm{L}}\right\|_{\mathrm{F}}^2\left\|\left(\boldsymbol{E}_{\mathrm{L}}^*\boldsymbol{E}_{\mathrm{L}}\right)^{-1}\right\| \\ &\le \mu_1\left\|\boldsymbol{e}_i^\top\boldsymbol{E}_{\mathrm{L}}\right\|_{\mathrm{F}}^2 \le \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \end{aligned} \tag{C.23}$$

That said, the total energy allocated to any row of $\boldsymbol{U}\boldsymbol{V}^*$ cannot exceed $\frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}$.

Moreover, the matrix $\mathcal{P}_T\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)$ enjoys similar properties as well, which we briefly reason as follows. First, the matrix $\boldsymbol{U}\boldsymbol{U}^*\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)$ obeys

$$\left\|\boldsymbol{e}_i^\top\boldsymbol{U}\boldsymbol{U}^*\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\right\|_{\mathrm{F}}^2 \le \left\|\boldsymbol{e}_i^\top\boldsymbol{U}\right\|_{\mathrm{F}}^2\left\|\boldsymbol{U}^*\right\|^2\left\|\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right\|^2 \le \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2},$$

since the operator norm of $\boldsymbol{U}$ and $\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}$ are both bounded by 1. The same bound for $\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\boldsymbol{V}\boldsymbol{V}^*$ can be demonstrated via the same argument as for $\boldsymbol{U}\boldsymbol{U}^*\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)$. Additionally, for $\boldsymbol{U}\boldsymbol{U}^*\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\boldsymbol{V}\boldsymbol{V}^*$ one has

$$\left\|\boldsymbol{e}_i^\top\boldsymbol{U}\boldsymbol{U}^*\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\boldsymbol{V}\boldsymbol{V}^*\right\|_{\mathrm{F}}^2 \le \left\|\boldsymbol{e}_i^\top\boldsymbol{U}\right\|_{\mathrm{F}}^2\left\|\boldsymbol{U}^*\right\|^2\left\|\boldsymbol{V}\boldsymbol{V}^*\right\|^2\left\|\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right\|^2 \le \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2}.$$

By definition of $\mathcal{P}_T$,

$$\begin{aligned}\left\|\boldsymbol{e}_i^\top\mathcal{P}_T\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\right\|_{\mathrm{F}}^2 \le 3\left\|\boldsymbol{e}_i^\top\boldsymbol{U}\boldsymbol{U}^*\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\right\|_{\mathrm{F}}^2 &+ 3\left\|\boldsymbol{e}_i^\top\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\boldsymbol{V}\boldsymbol{V}^*\right\|_{\mathrm{F}}^2 \\ &+ 3\left\|\boldsymbol{e}_i^\top\boldsymbol{U}\boldsymbol{U}^*\left(\sqrt{\omega_{\alpha,\beta}}\boldsymbol{A}_{(\alpha,\beta)}\right)\boldsymbol{V}\boldsymbol{V}^*\right\|_{\mathrm{F}}^2 \le \frac{9\mu_1 c_{\mathrm{s}} r}{n_1 n_2}.\end{aligned}$$

Now our task boils down to bounding $\|\boldsymbol{M}\|_{\mathcal{A},2}$ for some matrix $\boldsymbol{M}$ satisfying some energy constraints per row, which subsumes $\|\boldsymbol{U}\boldsymbol{V}^*\|_{\mathcal{A},2}$ and $\left\|\mathcal{P}_T\left(\sqrt{\omega_{k,l}}\boldsymbol{A}_{(k,l)}\right)\right\|_{\mathcal{A},2}$ as special cases. We can then conclude the proof by applying the following lemma.

Lemma C.2. *Denote by the set $\mathcal{M}$ of feasible matrices satisfying*

$$\max_i\left\|\boldsymbol{e}_i^\top\boldsymbol{M}\right\|_{\mathrm{F}}^2 \le \frac{9\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{C.24}$$

Then there exists some universal constant $c_3 > 0$ such that

$$\max_{\boldsymbol{M}\in\mathcal{M}} \|\boldsymbol{M}\|_{\mathcal{A},2}^2 \leq c_3 \frac{\mu_1 c_{\mathrm{s}} r}{n_1 n_2} \log^2 (n_1 n_2) . \tag{C.25}$$

Proof. For ease of presentation, we split any matrix $\boldsymbol{M}$ into 4 parts, defined as follows:

- $\boldsymbol{M}^{(1)}$ (resp. $\boldsymbol{M}^{(2)}$, $\boldsymbol{M}^{(3)}$, $\boldsymbol{M}^{(4)}$): the matrix containing all upper (resp. lower, upper, lower) triangular components of all upper (resp. upper, lower, lower) triangular blocks of $\boldsymbol{M}$;

Here, we use the term "upper triangular" and "lower triangular" in short for "left upper triangular" and "right lower triangular", which are more natural for Hankel matrices. Instead of maximizing $\|\boldsymbol{M}\|_{\mathcal{A},2}$ directly, we will handle $\max_{\boldsymbol{M}\in\mathcal{M}} \|\boldsymbol{M}^{(l)}\|_{\mathcal{A},2}^2$ for each $1 \leq l \leq 4$ separately, owing to the fact that

$$\max_{\boldsymbol{M}\in\mathcal{M}} \|\boldsymbol{M}\|_{\mathcal{A},2}^2 \leq 4 \max_{\boldsymbol{M}:\, \boldsymbol{M}^{(l)}\in\mathcal{M}} \left\|\boldsymbol{M}^{(l)}\right\|_{\mathcal{A},2}^2 . \tag{C.26}$$

In the sequel, we only demonstrate how to control $\|\boldsymbol{M}^{(1)}\|_{\mathcal{A},2}$. Similar bounds can be derived for $\|\boldsymbol{M}^{(l)}\|_{\mathcal{A},2}$ $(2 \leq l \leq 4)$ via very similar arguments.

To facilitate analysis, we divide the entire index set into several subsets $\mathcal{W}_{i,j}$ such that for all $1 \leq i \leq \lceil \log (n_1) \rceil$ and $1 \leq j \leq \lceil \log (n_2) \rceil$,

$$\mathcal{W}_{i,j} := \bigcup \left\{ \Omega_{\mathrm{e}} (k,l) \mid (k,l) \in \left[2^{i-1}, 2^i\right] \times \left[2^{j-1}, 2^j\right] \right\} . \tag{C.27}$$

Consequently, for each $\Omega_{\mathrm{e}} (k,l) \subseteq \mathcal{W}_{i,j}$, one has $2^{i-1} \cdot 2^{j-1} \leq \omega_{k,l} \leq 2^{i+j}$. This allows us to derive for each $\mathcal{W}_{i,j}$ that

$$\sum_{(k,l)\in\mathcal{W}_{i,j}} \frac{1}{\omega_{k,l}^2} \left| \sum\nolimits_{(\alpha,\beta)\in\Omega_{\mathrm{e}}(k,l)} \boldsymbol{M}_{\alpha,\beta}^{(1)} \right|^2 \leq \sum_{(k,l)\in\mathcal{W}_{i,j}} \frac{1}{\omega_{k,l}} \sum\nolimits_{(\alpha,\beta)\in\Omega_{\mathrm{e}}(k,l)} \left| \boldsymbol{M}_{\alpha,\beta}^{(1)} \right|^2 \tag{C.28}$$

$$\leq \frac{1}{2^{i+j-2}} \sum_{(k,l)\in\mathcal{W}_{i,j}} \sum\nolimits_{(\alpha,\beta)\in\Omega_{\mathrm{e}}(k,l)} \left| \boldsymbol{M}_{\alpha,\beta}^{(1)} \right|^2 , \tag{C.29}$$

where (C.28) follows from the root-mean square v.s. arithmetic mean inequality.

Observe that the indices contained in $\mathcal{W}_{i,j}$ reside within no more than $2^i \cdot 2^j$ rows. By assumption (C.24), the total energy allocated to $\mathcal{W}_{i,j}$ must be bounded above by

$$\sum_{(k,l)\in\mathcal{W}_{i,j}} \sum\nolimits_{(\alpha,\beta)\in\Omega_{\mathrm{e}}(k,l)} \left| \boldsymbol{M}^{(1)}_{\alpha,\beta} \right|^2 \leq 2^i \cdot 2^j \max_i \left\| \boldsymbol{e}_i^\top \boldsymbol{M} \right\|_{\mathrm{F}}^2 \leq 2^{i+j} \cdot \frac{9\mu_1 c_{\mathrm{s}} r}{n_1 n_2}.$$

Substituting this into (C.29) immediately leads to

$$\sum_{(k,l)\in\mathcal{W}_{i,j}} \frac{1}{\omega_{k,l}^2} \left| \sum\nolimits_{(\alpha,\beta)\in\Omega_{\mathrm{e}}(k,l)} \boldsymbol{M}^{(1)}_{\alpha,\beta} \right|^2 \leq \frac{36\mu_1 c_{\mathrm{s}} r}{n_1 n_2}. \tag{C.30}$$

By definition, $\|\boldsymbol{M}\|_{\mathcal{A},2}^2 = \sum_{1\leq i\leq\lceil\log n_1\rceil, 1\leq j\leq\lceil\log n_2\rceil} \sum_{(k,l)\in\mathcal{W}_{i,j}} \frac{\left|\sum_{(\alpha,\beta)\in\Omega_{\mathrm{e}}(k,l)} \boldsymbol{M}_{\alpha,\beta}\right|^2}{\omega_{k,l}^2}$. Combining the above bounds over all $\mathcal{W}_{i,j}$ then gives

$$\left\| \boldsymbol{M}^{(1)} \right\|_{\mathcal{A},2}^2 \leq \frac{36\mu_1 c_{\mathrm{s}} r \left\lceil \log\left(n_1\right) \right\rceil \cdot \left\lceil \log\left(n_2\right) \right\rceil}{n_1 n_2}$$

as claimed. $\square$

Appendix D

Proofs of Theorems and Lemmas in Chapter 5

D.1 Proof of Proposition 5.1

To prove Proposition 5.1, we will first derive an upper bound and a lower bound on $\mathbb{E}\left[|\langle \boldsymbol{B}_i, \boldsymbol{X}\rangle|\right]$, and then apply the Bernstein-type inequality [135, Proposition 5.16] to establish the large deviation bound.

In order to derive an upper bound on $\mathbb{E}\left[|\langle \boldsymbol{B}_i, \boldsymbol{X}\rangle|\right]$, the key step is to apply the Hanson-Wright inequality [141, 142], which characterizes the concentration of measure for quadratic forms in sub-Gaussian random variables. We adopt the version in [142] and repeat it below for completeness.

Lemma D.1 (**Hanson-Wright Inequality**). *Let $\boldsymbol{X} = (X_1, \ldots, X_n) \in \mathbb{R}^n$ be a random vector with independent components X_i which satisfy $\mathbb{E}\left[X_i\right] = 0$ and $\|X_i\|_{\psi_2} \leq K$. Let $\boldsymbol{A}$ be an $n \times n$ matrix. Then for any $t > 0$,*

$$\mathbb{P}\left\{\left|\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} - \mathbb{E}\left[\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}\right]\right| > t\right\} \leq 2\exp\left[-c\min\left(\frac{t^2}{K^4\|\boldsymbol{A}\|_F^2}, \frac{t}{K^2\|\boldsymbol{A}\|}\right)\right] \tag{D.1}$$

Observe that $\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle$ can be written as a symmetric quadratic form in $2n$ i.i.d. sub-Gaussian random variables:

$$\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle = \begin{bmatrix} \boldsymbol{a}_{2i-1}^\top & \boldsymbol{a}_{2i}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & \\ & -\boldsymbol{X} \end{bmatrix} \begin{bmatrix} \boldsymbol{a}_{2i-1} \\ \boldsymbol{a}_{2i} \end{bmatrix}.$$

The Hanson-Wright inequality (D.1) then asserts that: there exists an absolute constant $c > 0$ such that for any matrix $\boldsymbol{X}$, $|\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle| \leq t$ with probability at least

$$1 - 2\exp\left[-c\min\left(\frac{t^2}{4K^4 \|\boldsymbol{X}\|_{\mathrm{F}}^2}, \frac{t}{K^2\|\boldsymbol{X}\|}\right)\right].$$

This indicates that $\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle$ is a sub-exponential random variable [135] satisfying

$$\mathbb{E}\left[|\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle|\right] \leq c_1 \|\boldsymbol{X}\|_{\mathrm{F}} \tag{D.2}$$

for some positive constant c_1.

On the other hand, to derive a lower bound on $\mathbb{E}\left[|\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle|\right]$, we notice that for a random variable ξ, repeatedly applying the Cauchy-Schwartz inequality yields

$$\left(\mathbb{E}\left[\xi^2\right]\right)^2 \leq \mathbb{E}\left[|\xi|\right]\mathbb{E}\left[|\xi|^3\right] \leq \mathbb{E}\left[|\xi|\right]\sqrt{\mathbb{E}\left[\xi^2\right]\mathbb{E}\left[\xi^4\right]},$$

which further leads to

$$\mathbb{E}\left[|\xi|\right] \geq \sqrt{\frac{\left(\mathbb{E}\left[\xi^2\right]\right)^3}{\mathbb{E}\left[\xi^4\right]}}. \tag{D.3}$$

Let $\xi := \langle \boldsymbol{B}_i, \boldsymbol{X} \rangle$, of which the second moment can be expressed as

$$\mathbb{E}\left[\xi^2\right] = \mathbb{E}\left[|\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle|^2\right] = \langle \boldsymbol{X}, \mathbb{E}\left[\left(\mathcal{B}_i^*\mathcal{B}_i\right)\boldsymbol{X}\right] \rangle.$$

Simple algebraic manipulation yields

$$\mathbb{E}\left[\left(\mathcal{B}_i^*\mathcal{B}_i\right)(\boldsymbol{X})\right] = 4\boldsymbol{X} + 2\left(\mu_4 - 3\right)\mathrm{diag}(\boldsymbol{X}),$$

and hence

$$\mathbb{E}\left[\xi^2\right] = 4\left\|\boldsymbol{X}\right\|_{\mathrm{F}}^2 + 2\left(\mu_4 - 3\right)\sum_{i=1}^{n}\left|\boldsymbol{X}_{ii}\right|^2 \geq \min\{4, 2(\mu_4 - 1)\}\left\|\boldsymbol{X}\right\|_{\mathrm{F}}^2 = c_2\left\|\boldsymbol{X}\right\|_{\mathrm{F}}^2, \tag{D.4}$$

where $c_2 := \min\{4, 2(\mu_4 - 1)\}$. Furthermore, since $\xi := \langle \boldsymbol{B}_i, \boldsymbol{X} \rangle$ has been shown to be sub-exponential with sub-exponential norm $O\left(\|\boldsymbol{X}\|_{\mathrm{F}}\right)$, one can derive [135]

$$\mathbb{E}\left[\xi^4\right] = \left(4\left\|\xi\right\|_{\psi_1}\right)^4 \leq c_3\left\|\boldsymbol{X}\right\|_{\mathrm{F}}^4 \tag{D.5}$$

for some constant $c_7 > 0$. This taken collectively with (D.3) and (D.4) gives rise to

$$\mathbb{E}\left[\left|\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle\right|\right] \geq \sqrt{\frac{c_2^3\left\|\boldsymbol{X}\right\|_{\mathrm{F}}^6}{c_3\left\|\boldsymbol{X}\right\|_{\mathrm{F}}^4}} = c_4\left\|\boldsymbol{X}\right\|_{\mathrm{F}}$$

for some constant $c_4 > 0$.

Now, we are ready to characterize the concentration of $\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle$, which is a simple consequence of the following sub-exponential variant of Bernstein inequality.

Lemma D.2. *[135, Proposition 5.16] Let $X_1, \ldots, X_m$ be independent sub-exponential random variables with $\mathbb{E}[X_i] = 0$ and $K = \max_i \|X_i\|_{\psi_1}$. Then for every $t > 0$, we have*

$$\mathbb{P}\left\{\frac{1}{m}\left|\sum_{i=1}^{m} X_i\right| \geq t\right\} \leq 2\exp\left[-cm\min\left(\frac{t^2}{K^2}, \frac{t}{K}\right)\right] \tag{D.6}$$

where c is an absolute constant.

For any matrix $\boldsymbol{X}$, let $X_i = \left|\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle\right| - \mathbb{E}\left[\left|\langle \boldsymbol{B}_i, \boldsymbol{X} \rangle\right|\right]$, $i = 1, \ldots, m$. It is apparent that X_i's satisfy the conditions in Lemma D.2, therefore for any $\epsilon > 0$, we have

$$\left|\frac{1}{m}\left\|\mathcal{B}(\boldsymbol{X})\right\|_1 - \frac{1}{m}\mathbb{E}\left[\left\|\mathcal{B}\left(\boldsymbol{X}\right)\right\|_1\right]\right| \leq \epsilon\|\boldsymbol{X}\|_{\mathrm{F}}$$

with probability exceeding $1 - 2\exp(-cm\epsilon)$ for some absolute constant $c > 0$. This yields

$$\frac{1}{m}\left\|\mathcal{B}(\boldsymbol{X})\right\|_1 \leq \frac{1}{m}\mathbb{E}\left[\left\|\mathcal{B}\left(\boldsymbol{X}\right)\right\|_1\right] + \epsilon\|\boldsymbol{X}\|_{\mathrm{F}} \leq (c_1 + \epsilon)\|\boldsymbol{X}\|_{\mathrm{F}}$$

and

$$\frac{1}{m}\left\|\mathcal{B}(\boldsymbol{X})\right\|_1 \geq \frac{1}{m}\mathbb{E}\left[\left\|\mathcal{B}\left(\boldsymbol{X}\right)\right\|_1\right] - \epsilon\|\boldsymbol{X}\|_{\mathrm{F}} \geq (c_4-\epsilon)\|\boldsymbol{X}\|_{\mathrm{F}}$$

with probability at least $1-2\exp(-cm\epsilon)$, where the constants c, c_1 and c_4 depend only on the sub-Gaussian norm of a_i. Renaming the universal constants establishes Proposition 5.1.

D.2 Proof of Theorem 5.4

The proof of Theorem 5.4 follows the entropy method introduced in [60]. Specifically, the RIP-ℓ_2/ℓ_2 constant can be bounded by

$$\begin{aligned}
\delta_r &= \sup_{\|\boldsymbol{X}\|_{\mathrm{F}}\leq 1,\mathrm{rank}(\boldsymbol{X})\leq r}\left|\frac{1}{m}\sum_{i=1}^m \left|\langle \boldsymbol{B}_i,\boldsymbol{X}\rangle\right|^2 - \|\boldsymbol{X}\|_{\mathrm{F}}^2\right| \\
&= \sup_{T\in\mathcal{M}_r^2,\boldsymbol{X}\in T,\|\boldsymbol{X}\|_{\mathrm{F}}\leq 1}\left|\left\langle \boldsymbol{X},\left(\frac{1}{m}\sum_{i=1}^m \mathcal{B}_i^*\mathcal{B}_i-\mathcal{I}\right)\boldsymbol{X}\right\rangle\right| \leq \sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left(\frac{1}{m}\sum_{i=1}^m \mathcal{B}_i^*\mathcal{B}_i-\mathcal{I}\right)\mathcal{P}_T\right\| \\
&\leq \sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left\{\frac{1}{m}\sum_{i=1}^m\left(\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right)\right\}\mathcal{P}_T\right\| + \frac{c_5}{n}, \qquad \text{(D.7)}
\end{aligned}$$

where $\mathcal{M}_r^2 := \{\text{tangent space with respect to } \boldsymbol{M} \mid \forall \boldsymbol{M}:\ \mathrm{rank}\left(\boldsymbol{M}\right)\leq r\}$, and the last inequality follows from the near-isotropic assumption of $\mathcal{B}_i$ (i.e. (5.22)).

The first step is to prove that $\mathbb{E}\left[\delta_r\right] \leq \epsilon$ for some small constant $\epsilon > 0$. For sufficiently large n, it suffices to prove that

$$E := \mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left\{\frac{1}{m}\sum_{i=1}^m\left(\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\mathcal{B}_i^*\mathcal{B}_i\right)\right\}\mathcal{P}_T\right\|\right] \leq \delta. \qquad \text{(D.8)}$$

This can be established by a Gaussian process approach as follows.

Observe that $\frac{1}{m}\sum_{i=1}^m\left(\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right)$ is a zero-mean operator, which can be reduced to symmetric operators via the symmetrization argument (see, e.g. [127]).

Specifically, let $\tilde{\mathcal{B}}_i$ be an independent copy of $\mathcal{B}_i$. Conditioning on $\mathcal{B}_i$ we have

$$\mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\frac{1}{m}\sum_{i=1}^{m}\tilde{\mathcal{B}}_i^*\tilde{\mathcal{B}}_i\right|\left.\mathcal{B}_i\ (1\leq i\leq m)\right]=\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right].$$

Since the function $f(\mathcal{X}):=\sup_{T\in\mathcal{M}_r^2}\|\mathcal{P}_T\mathcal{X}\mathcal{P}_T\|$ is convex in $\mathcal{X}$, applying Jensen's inequality yields

$$\begin{aligned}
&\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left\{\frac{1}{m}\sum_{i=1}^{m}\left(\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right)\right\}\mathcal{P}_T\right\|\\
&\quad=\sup_{T\in\mathcal{M}_r^2}\left\|\mathbb{E}\left[\mathcal{P}_T\left(\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\frac{1}{m}\sum_{i=1}^{m}\tilde{\mathcal{B}}_i^*\tilde{\mathcal{B}}_i\right)\mathcal{P}_T\right|\mathcal{B}_i\ (1\leq i\leq m)\right]\right\|\\
&\quad\leq\ \mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left(\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\frac{1}{m}\sum_{i=1}^{m}\tilde{\mathcal{B}}_i^*\tilde{\mathcal{B}}_i\right)\mathcal{P}_T\right\|\right|\left.\mathcal{B}_i\ (1\leq i\leq m)\right].
\end{aligned}$$

By undoing conditioning over $\mathcal{B}_i$ we derive

$$\begin{aligned}
&\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left\{\frac{1}{m}\sum_{i=1}^{m}\left(\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right)\right\}\mathcal{P}_T\right\|\right]\\
&\quad\leq\ \mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left(\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\frac{1}{m}\sum_{i=1}^{m}\tilde{\mathcal{B}}_i^*\tilde{\mathcal{B}}_i\right)\mathcal{P}_T\right\|\right]\\
&\quad\leq\ 2\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|\right],
\end{aligned}\tag{D.9}$$

where ϵ_i's are i.i.d. symmetric Bernoulli random variables. Moreover, if we generate a set of i.i.d. random variables $g_i\sim\mathcal{N}(0,1)$, then the conditional expectation obeys

$$\mathbb{E}\left[\frac{1}{m}\sum_{i=1}^{m}|g_i|\,\epsilon_i\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right|\left.\epsilon_i,\mathcal{B}_i\ (1\leq i\leq m)\right]=\sqrt{\frac{2}{\pi}}\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T.$$

Similarly, by convexity of $f(\mathcal{X}) := \sup_{T\in\mathcal{M}_r^2} \|\mathcal{P}_T\mathcal{X}\mathcal{P}_T\|$, one can obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\frac{1}{m}\sum_{i=1}^m \epsilon_i\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|\right] \\
&\quad = \sqrt{\frac{\pi}{2}}\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathbb{E}\left[\frac{1}{m}\sum_{i=1}^m |g_i|\,\epsilon_i\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\,\middle|\,\epsilon_i,\mathcal{B}_i\ (1\le i\le m)\right]\right\|\right] && \text{(D.10)}\\
&\quad \le \sqrt{\frac{\pi}{2}}\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\frac{1}{m}\sum_{i=1}^m g_i\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|\right]. && \text{(D.11)}
\end{aligned}$$

Putting (D.9) and (D.11) together we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left\{\frac{1}{m}\sum_{i=1}^m\left(\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right)\right\}\mathcal{P}_T\right\|\right] \le \sqrt{2\pi}\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\frac{1}{m}\sum_{i=1}^m g_i\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|\right]\\
&= \sqrt{2\pi}\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2,\boldsymbol{X}\in T,\|\boldsymbol{X}\|_{\mathrm{F}}=1}\left|g_i\frac{1}{m}\sum_{i=1}^m\left|\mathcal{B}_i(\boldsymbol{X})\right|^2\right|\right], && \text{(D.12)}
\end{aligned}$$

which converts the problem into bounding the supremum of a Gaussian process.

We now prove the following lemma.

Lemma D.3. *Suppose that $g_i \sim \mathcal{N}(0,1)$ are i.i.d. random variables, and that $K \le n^2$. Conditional on $\mathcal{B}_i$'s, we have*

$$\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left(\sum_{i=1}^m g_i\mathcal{B}_i^*\mathcal{B}_i\right)\mathcal{P}_T\right\|\right] \le C_{14}\sqrt{r}K\log^3 n \sup_{T:T\in\mathcal{M}_r^2}\sqrt{\left\|\sum_{i=1}^m\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|}. \tag{D.13}$$

Combining Lemma D.3 with (D.12) and undoing the conditioning on $\mathcal{B}_i$'s yield

$$\begin{aligned}
&\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\left\{\frac{1}{m}\sum_{i=1}^{m}\left(\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right)\right\}\mathcal{P}_T\right\|\right]\\
&\quad\leq\ \frac{C_{15}\sqrt{r}K\log^3 n}{m}\left(\mathbb{E}\left[\sqrt{\sup_{T:T\in\mathcal{M}_r^2}\left\|\sum_{i=1}^{m}\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|}\right]\right)\\
&\quad\leq\ \frac{C_{15}\sqrt{r}K\log^3 n}{\sqrt{m}}\sqrt{\mathbb{E}\left[\sup_{T:T\in\mathcal{M}_r^2}\left\|\frac{1}{m}\sum_{i=1}^{m}\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|\right]}
\end{aligned}$$

for some numerical constant $C_{15}>0$, where the last inequality follows from Jensen's inequality. Recall the definition of E in (D.8), then the above inequality implies

$$E\leq C_{15}\left(\frac{\sqrt{r}K\log^3 n}{\sqrt{m}}\right)\sqrt{E+1},$$

or more concretely,

$$\mathbb{E}\left[\delta_r\right]\leq E\leq 2C_{15}\frac{\sqrt{r}K\log^3 n}{\sqrt{m}}<1 \tag{D.14}$$

as soon as $m>\left(2C_{15}\sqrt{r}K\log^3 n\right)^2$.

Now that we have established that $\mathbb{E}\left[\delta_r\right]$ can be a small constant if $m>4C_{15}^2 rK^2\log^6 n$, it remains to show that δ_r sharply concentrates around $\mathbb{E}\left[\delta_r\right]$. To this end, consider the Banach space Υ of operators $\mathcal{H}:\mathbb{R}^{n\times n}\mapsto\mathbb{R}^{n\times n}$ equipped with the norm

$$\left\|\mathcal{H}\right\|_{\Upsilon}:=\sup_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\mathcal{H}\mathcal{P}_T\right\|.$$

Let ε_i's be i.i.d. symmetric Bernoulli variables, then the symmetrization trick yields

$$\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}\right]\leq\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\varepsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}\right]\leq 2\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}\right],$$

and

$$\mathbb{P}\left\{\left\|\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right\|_{\Upsilon}>2\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i-\mathbb{E}\left[\mathcal{B}_i^*\mathcal{B}_i\right]\right\|_{\Upsilon}\right]+u\right\}$$
$$\leq\mathbb{P}\left\{\left\|\frac{1}{m}\sum_{i=1}^{m}\left(\mathcal{B}_i^*\mathcal{B}_i-\tilde{\mathcal{B}}_i^*\tilde{\mathcal{B}}_i\right)\right\|_{\Upsilon}>u\right\}\leq 2\mathbb{P}\left\{\left\|\frac{1}{m}\sum_{i=1}^{m}\varepsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}>\frac{u}{2}\right\},$$

where $\tilde{\mathcal{B}}_i$ is an independent copy of $\mathcal{B}_i$. Note that $\varepsilon_i\mathcal{B}_i^*\mathcal{B}_i$'s are i.i.d. zero-mean random operators.

In addition, for any $1\leq i\leq m$, we know that

$$\left\|\epsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}=\max_{T\in\mathcal{M}_r^2}\left\|\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|=\max_{T\in\mathcal{M}_r^2,\|\boldsymbol{X}\|_{\mathrm{F}}=1}\left|\left\langle\boldsymbol{B}_i,\mathcal{P}_T\left(\boldsymbol{X}\right)\right\rangle\right|^2$$
$$\leq\max_{T\in\mathcal{M}_r^2,\|\boldsymbol{X}\|_{\mathrm{F}}=1}\left\|\boldsymbol{B}_i\right\|^2\left\|\mathcal{P}_T\left(\boldsymbol{X}\right)\right\|_*^2\leq K^2r.$$

[60, Theorem 3.10] asserts that there is a universal constant $C_{12}>0$ such that

$$\mathbb{P}\left\{\left\|\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}>8q\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}\right]+\frac{2K^2r}{m}l+t\right\}$$
$$\leq\left(\frac{C_{12}}{q}\right)^l+2\exp\left(-\frac{t^2}{256q\left(\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}\right]\right)^2}\right)$$

If we take $q=2C_{12}$, $l=C_{13}\log n$ and $t=C_{14}\sqrt{\log n}\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}\right]$, then for sufficiently large C_{13} and C_{14}, there exists an absolute constant $C_{20}>0$ such that if $m>C_{20}rK^2\log^7 n$, then for any small positive constant δ we have

$$\left\|\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}<C_{15}\sqrt{\log n}\mathbb{E}\left[\left\|\frac{1}{m}\sum_{i=1}^{m}\epsilon_i\mathcal{B}_i^*\mathcal{B}_i\right\|_{\Upsilon}\right]<\delta$$

with probability exceeding $1-n^{-2}$.

Now that we have ensured a small RIP-ℓ_2/ℓ_2 constant, repeating the argument as in [119, 89] implies

$$\|\hat{\boldsymbol{\Sigma}}-\boldsymbol{\Sigma}\|_{\mathrm{F}}\leq C_2\frac{\epsilon_2}{\sqrt{m}}\tag{D.15}$$

for all $\mathbf{\Sigma}$ of rank at most r. This concludes the proof.

D.3 Proof of Lemma 5.1

We first introduce a few mathematical notations before proceeding to the proof. Let the singular value decomposition of a rank-r matrix $\mathbf{\Sigma}$ be $\mathbf{\Sigma} = \boldsymbol{U}\mathbf{\Lambda}\boldsymbol{V}^\top$, then the tangent space T at the point $\mathbf{\Sigma}$ is defined as $T := \left\{\boldsymbol{U}\boldsymbol{M}_1 + \boldsymbol{M}_2\boldsymbol{V}^\top \mid \boldsymbol{M}_1 \in \mathbb{R}^{r\times n}, \boldsymbol{M}_2 \in \mathbb{R}^{n\times r}\right\}$. We denote by $\mathcal{P}_T$ and $\mathcal{P}_{T^\perp}$ the orthogonal projection onto T and its orthogonal complement, respectively. For notational simplicity, we denote $\boldsymbol{H}_T := \mathcal{P}_T(\boldsymbol{H})$ and $\boldsymbol{H}_{T^\perp} := \boldsymbol{H} - \mathcal{P}_T(\boldsymbol{H})$ for any matrices $\boldsymbol{H} \in \mathbb{R}^{n\times n}$.

Write $\mathbf{\Sigma} := \mathbf{\Sigma}_r + \mathbf{\Sigma}_{\mathrm{c}}$, where $\mathbf{\Sigma}_r$ represents the best rank-r approximation of $\mathbf{\Sigma}$. Denote by T the tangent space with respect to $\mathbf{\Sigma}_r$. Suppose that the solution to (5.5) is given by $\hat{\mathbf{\Sigma}} = \mathbf{\Sigma} + \boldsymbol{H}$ for some matrix $\boldsymbol{H}$. The optimality of $\hat{\mathbf{\Sigma}}$ yields

$$\begin{aligned}
0 &\geq \|\mathbf{\Sigma} + \boldsymbol{H}\|_* - \|\mathbf{\Sigma}\|_* \geq \|\mathbf{\Sigma}_r + \boldsymbol{H}\|_* - \|\mathbf{\Sigma}_c\|_* - \|\mathbf{\Sigma}\|_* \\
&\geq \|\mathbf{\Sigma}_r + \boldsymbol{H}_{T^\perp}\|_* - \|\boldsymbol{H}_T\|_* - \|\mathbf{\Sigma}_r\|_* - 2\|\mathbf{\Sigma}_{\mathrm{c}}\|_* \\
&= \|\mathbf{\Sigma}_r\|_* + \|\boldsymbol{H}_{T^\perp}\|_* - \|\boldsymbol{H}_T\|_* - \|\mathbf{\Sigma}_r\|_* - 2\|\mathbf{\Sigma}_{\mathrm{c}}\|_*,
\end{aligned}$$

which leads to

$$\|\boldsymbol{H}_{T^\perp}\|_* \leq \|\boldsymbol{H}_T\|_* + 2\|\mathbf{\Sigma}_{\mathrm{c}}\|_*. \tag{D.16}$$

We then divide $\boldsymbol{H}_{T^\perp}$ into $M = \left\lceil \frac{n-r}{K_1} \right\rceil$ orthogonal matrices $\boldsymbol{H}_1, \boldsymbol{H}_2, \cdots, \boldsymbol{H}_M$ satisfying the following: (i) the largest singular value of $\boldsymbol{H}_{i+1}$ does not exceed the smallest non-zero singular value of $\boldsymbol{H}_i$, and (ii)

$$\|\boldsymbol{H}_{T^\perp}\|_* = \sum_{i=1}^{M} \|\boldsymbol{H}_i\|_* \tag{D.17}$$

and rank$(\boldsymbol{H}_i) = K_1$ for $1 \leq i \leq M-1$. Along with the bound (D.16), this yields that

$$\sum_{i\geq 2} \|\boldsymbol{H}_i\|_{\mathrm{F}} \leq \frac{1}{\sqrt{K_1}} \sum_{i\geq 2} \|\boldsymbol{H}_{i-1}\|_* \leq \frac{1}{\sqrt{K_1}} \|\boldsymbol{H}_{T^\perp}\|_* \leq \frac{1}{\sqrt{K_1}} \left(\|\boldsymbol{H}_T\|_* + 2\|\boldsymbol{\Sigma}_{\mathrm{c}}\|_*\right). \tag{D.18}$$

It then follows that $\frac{1}{m}\|\mathcal{B}(\boldsymbol{H})\|_1 \leq \frac{2}{m}\|\mathcal{A}(\boldsymbol{\Sigma}) - \boldsymbol{y}\|_1 \leq \frac{2\epsilon_1}{m}$, and that

$$\begin{aligned}
\frac{2\epsilon_1}{m} &\geq \frac{1}{m}\|\mathcal{B}(\boldsymbol{H})\|_1 \geq \frac{1}{m}\|\mathcal{B}(\boldsymbol{H}_T + \boldsymbol{H}_1)\|_1 - \sum_{i\geq 2} \frac{1}{m}\|\mathcal{B}(\boldsymbol{H}_i)\|_1 \\
&\geq \left(1 - \delta^{\mathrm{lb}}_{2r+K_1}\right)\|\boldsymbol{H}_T + \boldsymbol{H}_1\|_{\mathrm{F}} - \left(1 + \delta^{\mathrm{ub}}_{K_1}\right)\sum_{i\geq 2}\|\boldsymbol{H}_i\|_{\mathrm{F}} \\
&\geq \frac{\left(1 - \delta^{\mathrm{lb}}_{2r+K_1}\right)}{\sqrt{2}}\left(\|\boldsymbol{H}_T\|_{\mathrm{F}} + \|\boldsymbol{H}_1\|_{\mathrm{F}}\right) - \frac{\left(1 + \delta^{\mathrm{ub}}_{K_1}\right)}{\sqrt{K_1}}\left(\|\boldsymbol{H}_T\|_* + 2\|\boldsymbol{\Sigma}_{\mathrm{c}}\|_*\right).
\end{aligned}$$

By reorganizing the terms and and using the fact that $\|\boldsymbol{H}_T\|_* \leq \sqrt{2r}\|\boldsymbol{H}_T\|_{\mathrm{F}}$, one can derive

$$\left[\frac{\left(1 - \delta^{\mathrm{lb}}_{2r+K_1}\right)}{\sqrt{2}} - \frac{\left(1 + \delta^{\mathrm{ub}}_{K_1}\right)\sqrt{2r}}{\sqrt{K_1}}\right]\|\boldsymbol{H}_T\|_{\mathrm{F}} + \frac{\left(1 - \delta^{\mathrm{lb}}_{2r+K_1}\right)}{\sqrt{2}}\|\boldsymbol{H}_1\|_{\mathrm{F}} \leq \frac{2\left(1 + \delta^{\mathrm{ub}}_{K_1}\right)}{\sqrt{K_1}}\|\boldsymbol{\Sigma}_{\mathrm{c}}\|_* + \frac{2\epsilon_1}{m}. \tag{D.19}$$

The bound (D.19) allows us to see that if $\frac{1-\delta^{\mathrm{lb}}_{2r+K_1}}{\sqrt{2}} - \left(1 + \delta^{\mathrm{ub}}_{K_1}\right)\sqrt{\frac{2r}{K_1}} \geq \beta_1 > 0$ for some absolute constant β_1, then one has

$$\|\boldsymbol{H}_T\|_{\mathrm{F}} + \|\boldsymbol{H}_1\|_{\mathrm{F}} \leq \frac{2}{\beta_1}\left(\frac{\left(1 + \delta^{\mathrm{ub}}_{K_1}\right)}{\sqrt{K_1}}\|\boldsymbol{\Sigma}_{\mathrm{c}}\|_* + \frac{\epsilon_1}{m}\right). \tag{D.20}$$

On the other hand, (D.18) allows us to bound

$$\sum_{i\geq 2}\|\boldsymbol{H}_i\|_{\mathrm{F}} \leq \frac{1}{\sqrt{K_1}}\left(\|\boldsymbol{H}_T\|_* + 2\|\boldsymbol{\Sigma}_{\mathrm{c}}\|_*\right) \leq \sqrt{\frac{2r}{K_1}}\|\boldsymbol{H}_T\|_{\mathrm{F}} + \frac{2}{\sqrt{K_1}}\|\boldsymbol{\Sigma}_{\mathrm{c}}\|_*. \tag{D.21}$$

This taken collectively with (D.20) establishes

$$\|\boldsymbol{H}\|_{\mathrm{F}} \leq C_1 \frac{\|\boldsymbol{\Sigma}_{\mathrm{c}}\|_*}{\sqrt{K_1}} + C_2 \frac{\epsilon_1}{m}$$

for some absolute constants C_1 and C_2.

D.4 Proof of Lemma 5.2

Before proceeding to the proof, we introduce a few notations for convenience of presentation. Let $\boldsymbol{X} := \boldsymbol{x}\boldsymbol{x}^\top$, $\boldsymbol{X}_\Omega := \boldsymbol{x}_\Omega \boldsymbol{x}_\Omega^\top$ and $\boldsymbol{X}_{\mathrm{c}} := \boldsymbol{X} - \boldsymbol{X}_\Omega$, where $\boldsymbol{x}_\Omega$ denotes the k-sparse approximation of $\boldsymbol{x}$ whose support is denoted by Ω. We set $\boldsymbol{u} := \frac{1}{\|\boldsymbol{x}_\Omega\|_2}\boldsymbol{x}_\Omega$, and hence the tangent space T with respect to $\boldsymbol{X}_\Omega$ and its orthogonal complement $T^\perp$ are characterized by

$$\begin{aligned} T &:= \left\{ \boldsymbol{u}\boldsymbol{z}^\top + \boldsymbol{z}\boldsymbol{u}^\top \mid \boldsymbol{z} \in \mathbb{R}^n \right\}, \\ T^\perp &:= \left\{ \left(\boldsymbol{I} - \boldsymbol{u}\boldsymbol{u}^\top\right) \boldsymbol{M} \left(\boldsymbol{I} - \boldsymbol{u}\boldsymbol{u}^\top\right) \mid \boldsymbol{M} \in \mathbb{R}^{n\times n} \right\}. \end{aligned}$$

We adopt the notations introduced in [59] as follows: let Ω denote the support of $\boldsymbol{X}_\Omega$, and decompose the entire matrix space into the direct sum of 3 subspaces as

$$\left(T \cap \Omega\right) \oplus \left(T^\perp \cap \Omega\right) \oplus \left(\Omega^\perp\right). \tag{D.22}$$

In fact, one can verify that

$$T \cap \Omega = \left\{ \boldsymbol{u}\boldsymbol{z}^\top + \boldsymbol{z}\boldsymbol{u}^\top \mid \boldsymbol{z}_{\Omega^c} = \boldsymbol{0} \right\},$$

and that both the column and row spaces of $T^\perp \cap \Omega$ can be spanned by a set of $k-1$ orthonormal vectors that are supported on Ω and orthogonal to $\boldsymbol{u}$. As pointed out by [59], T and Ω are compatible in the sense that

$$\mathcal{P}_T \mathcal{P}_\Omega = \mathcal{P}_\Omega \mathcal{P}_T = \mathcal{P}_{T\cap\Omega}. \tag{D.23}$$

Additionally, for simplicity of notation, we use $\delta_{r,l}^{\mathrm{lb}}$ and $\delta_{r,l}^{\mathrm{ub}}$ to represent $\delta_{r,l}^{\mathrm{lb},k}$ and $\delta_{r,l}^{\mathrm{ub},k}$ in short whenever there is no ambiguity.

Suppose that $\hat{\boldsymbol{X}} = \boldsymbol{x}\boldsymbol{x}^\top + \boldsymbol{H}$ is the solution to (5.11). Then for any $\boldsymbol{W} \in T^\perp$ and $\boldsymbol{Y} \in \Omega^\perp$ satisfying $\|\boldsymbol{W}\| \leq 1$ and $\|\boldsymbol{Y}\|_\infty \leq \infty$, the matrix $\boldsymbol{u}\boldsymbol{u}^\top + \boldsymbol{W} + \lambda \mathrm{sign}\left(\boldsymbol{u}\right) \mathrm{sign}\left(\boldsymbol{u}\right)^\top + \lambda \boldsymbol{Y}$ forms a subgradient of the function $\|\cdot\|_* + \lambda \|\cdot\|_1$ at point $\boldsymbol{X}_\Omega$. If we pick $\boldsymbol{W}$ and $\boldsymbol{Y}$ such that $\boldsymbol{Y} = \mathrm{sign}\left(\boldsymbol{H}_{\Omega^\perp}\right)$ and $\langle \boldsymbol{W}, \boldsymbol{H}\rangle = \|\boldsymbol{H}_{T^\perp \cap \Omega}\|_*$, then

$$
\begin{aligned}
0 &\geq \|\boldsymbol{X}+\boldsymbol{H}\|_* + \lambda \|\boldsymbol{X}+\boldsymbol{H}\|_1 - \|\boldsymbol{X}\|_* - \lambda \|\boldsymbol{X}\|_1 && \text{(D.24)}\\
&\geq \|\boldsymbol{X}_\Omega+\boldsymbol{H}\|_* - \|\boldsymbol{X}_\mathrm{c}\|_* + \lambda \|\boldsymbol{X}_\Omega+\boldsymbol{H}\|_1 - \lambda \|\boldsymbol{X}_\mathrm{c}\|_1 - \|\boldsymbol{X}_\Omega\|_* - \|\boldsymbol{X}_\mathrm{c}\|_* - \lambda \|\boldsymbol{X}_\Omega\|_1 - \lambda \|\boldsymbol{X}_\mathrm{c}\|_1 \\
& && \text{(D.25)}\\
&\geq \left\langle \boldsymbol{u}\boldsymbol{u}^\top + \boldsymbol{W}, \boldsymbol{H}\right\rangle + \left\langle \lambda \mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top + \lambda\boldsymbol{Y}, \boldsymbol{H}\right\rangle - 2\|\boldsymbol{X}_\mathrm{c}\|_* - 2\lambda\|\boldsymbol{X}_\mathrm{c}\|_1 && \text{(D.26)}\\
&= \left\langle \boldsymbol{u}\boldsymbol{u}^\top, \boldsymbol{H}_T\right\rangle + \lambda\left\langle \mathcal{P}_T\left(\mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top\right), \boldsymbol{H}_T\right\rangle + \lambda\left\langle \mathcal{P}_{T^\perp}\left(\mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top\right), \boldsymbol{H}_{T^\perp}\right\rangle \\
&\quad + \|\boldsymbol{H}_{T^\perp\cap\Omega}\|_* + \lambda\|\boldsymbol{H}_{\Omega^\perp}\|_1 - 2\|\boldsymbol{X}_\mathrm{c}\|_* - 2\lambda\|\boldsymbol{X}_\mathrm{c}\|_1 \\
&\geq \left\langle \boldsymbol{u}\boldsymbol{u}^\top + \lambda\mathcal{P}_T\left(\mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top\right), \boldsymbol{H}_{T\cap\Omega}\right\rangle + \|\boldsymbol{H}_{T^\perp\cap\Omega}\|_* && \text{(D.27)}\\
&\qquad\qquad + \lambda\|\boldsymbol{H}_{\Omega^\perp}\|_1 - 2\|\boldsymbol{X}_\mathrm{c}\|_* - 2\lambda\|\boldsymbol{X}_\mathrm{c}\|_1,
\end{aligned}
$$

where (D.24) follows from the optimality of $\hat{\boldsymbol{X}}$, (D.25) follows from the definitions of $\boldsymbol{X}_\Omega$ and $\boldsymbol{X}_\mathrm{c}$ and the triangle inequality, (D.26) follows from the definition of subgradient. Finally, (D.27) follows from the following two facts:

(i) $\boldsymbol{H}_{T^\perp} \succeq 0$, a consequence of the feasibility constraint of (5.11). This further gives

$$\left\langle \mathcal{P}_{T^\perp}\left(\mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top\right), \boldsymbol{H}_{T^\perp}\right\rangle = \mathrm{sign}\left(\boldsymbol{u}\right)^\top \boldsymbol{H}_{T^\perp}\mathrm{sign}\left(\boldsymbol{u}\right) \geq 0.$$

(ii) It follows from (D.23) and the fact $\mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top \in \Omega$ that

$$\left\langle \mathcal{P}_T\left(\mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top\right), \boldsymbol{H}_T\right\rangle = \left\langle \mathcal{P}_{T\cap\Omega}\left(\mathrm{sign}\left(\boldsymbol{u}\right)\mathrm{sign}\left(\boldsymbol{u}\right)^\top\right), \boldsymbol{H}_{T\cap\Omega}\right\rangle. \qquad \text{(D.28)}$$

Since any matrix in T has rank at most 2, one can bound

$$\begin{aligned}
\left\|\mathcal{P}_T\left(\operatorname{sign}(\boldsymbol{u})\operatorname{sign}(\boldsymbol{u})^\top\right)\right\|_*^2 &\leq 2\left\|\mathcal{P}_T\left(\operatorname{sign}(\boldsymbol{u})\operatorname{sign}(\boldsymbol{u})^\top\right)\right\|_{\mathrm{F}}^2 \leq 4\left\|\boldsymbol{u}\boldsymbol{u}^\top\operatorname{sign}(\boldsymbol{u})\operatorname{sign}(\boldsymbol{u})^\top\right\|_{\mathrm{F}}^2 \\
&= 4\left|\langle\boldsymbol{u},\operatorname{sign}(\boldsymbol{u})\rangle\right|^2\left\|\operatorname{sign}(\boldsymbol{u})\right\|_{\mathrm{F}}^2 \\
&\leq 4k\left\|\boldsymbol{u}\right\|_1^2\left\|\operatorname{sign}(\boldsymbol{u})\right\|_\infty^2 \leq 4k\left\|\boldsymbol{u}\right\|_1^2 \leq \frac{4}{\lambda^2}, \qquad \text{(D.29)}
\end{aligned}$$

where (D.29) follows from the assumption on λ. Combining (D.29) with (D.27) yields

$$\begin{aligned}
&\left\|\boldsymbol{H}_{T^\perp\cap\Omega}\right\|_* + \lambda\left\|\boldsymbol{H}_{\Omega^\perp}\right\|_1 \\
&\quad\leq -\left\langle\boldsymbol{u}\boldsymbol{u}^\top,\boldsymbol{H}_{T\cap\Omega}\right\rangle - \lambda\left\langle\mathcal{P}_T\left(\operatorname{sign}(\boldsymbol{u})\operatorname{sign}(\boldsymbol{u})^\top\right),\boldsymbol{H}_{T\cap\Omega}\right\rangle + 2\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_* + 2\lambda\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_1 \\
&\quad\leq \left|\boldsymbol{u}^\top\boldsymbol{H}_{T\cap\Omega}\boldsymbol{u}\right| + \lambda\left\|\mathcal{P}_T\left(\operatorname{sign}(\boldsymbol{u})\operatorname{sign}(\boldsymbol{u})^\top\right)\right\|_*\cdot\left\|\boldsymbol{H}_{T\cap\Omega}\right\| + 2\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_* + 2\lambda\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_1 \\
&\quad\leq 3\left\|\boldsymbol{H}_{T\cap\Omega}\right\| + 2\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_* + 2\lambda\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_1, \qquad \text{(D.30)}
\end{aligned}$$

where (D.30) results from $\left\|\boldsymbol{u}\right\|_2 = 1$ and (D.29).

Divide $\boldsymbol{H}_{T^\perp\cap\Omega}$ into $M_1 := \left\lceil\frac{k-2}{K_1}\right\rceil$ orthogonal matrices $\boldsymbol{H}_{T^\perp\cap\Omega}^{(1)}$, $\boldsymbol{H}_{T^\perp\cap\Omega}^{(2)}$, $\cdots$, $\boldsymbol{H}_{T^\perp\cap\Omega}^{(M_1)} \in T^\perp\cap\Omega$ satisfying the following: (i) the largest singular value of $\boldsymbol{H}_{T^\perp\cap\Omega}^{(i+1)}$ does not exceed the smallest non-zero singular value of $\boldsymbol{H}_{T^\perp\cap\Omega}^{(i)}$, and (ii)

$$\left\|\boldsymbol{H}_{T^\perp\cap\Omega}\right\|_* = \sum_{i=1}^{M}\left\|\boldsymbol{H}_{T^\perp\cap\Omega}^{(i)}\right\|_* \quad\text{and}\quad \operatorname{rank}\left(\boldsymbol{H}_{T^\perp\cap\Omega}^{(i)}\right) = K_1 \ (1\leq i\leq M_1-1).$$

In the meantime, divide $\boldsymbol{H}_{\Omega^\perp}$ into $M_2 = \left\lceil\frac{n^2-k^2}{K_2}\right\rceil$ orthogonal matrices $\boldsymbol{H}_{\Omega^\perp}^{(1)}$, $\boldsymbol{H}_{\Omega^\perp}^{(2)}$, $\cdots$, $\boldsymbol{H}_{\Omega^\perp}^{(M_2)} \in \Omega^\perp$ of *non-overlapping* support such that (i) the largest entry magnitude of $\boldsymbol{H}_{\Omega^\perp}^{(i+1)}$ does not exceed that of the smallest non-zero entry of $\boldsymbol{H}_{\Omega^\perp}^{(i)}$, and (ii)

$$\left\|\boldsymbol{H}_{\Omega^\perp}^{(i)}\right\|_0 = K_2 \ (1\leq i\leq M_2-1).$$

This decomposition gives the following bound

$$\sum_{i=2}^{M_1}\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(i)}\right\|_{\mathrm{F}}\leq\sum_{i=2}^{M_1}\frac{1}{\sqrt{K_1}}\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(i-1)}\right\|_{*}\leq\frac{1}{\sqrt{K_1}}\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}\right\|_{*},$$

which combined with the RIP-ℓ_2/ℓ_1 property of $\mathcal{B}$ yields

$$\sum_{i=2}^{M_1}\frac{1}{m}\left\|\mathcal{B}\left(\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(i)}\right)\right\|_{1}\leq\frac{\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_1}}\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}\right\|_{*},\tag{D.31}$$

and, similarly,

$$\sum_{i=2}^{M_2}\frac{1}{m}\left\|\mathcal{B}\left(\boldsymbol{H}_{\Omega^{\perp}}^{(i)}\right)\right\|_{1}\leq\frac{\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_2}}\left\|\boldsymbol{H}_{\Omega^{\perp}}\right\|_{1}.\tag{D.32}$$

Set $K_2:=\left\lceil\frac{K_1}{\lambda^2}\right\rceil$, and hence $\sqrt{\frac{K_1}{K_2}}\leq\lambda$. Recalling $\boldsymbol{H}=\boldsymbol{H}_{T\cap\Omega}+\boldsymbol{H}_{T^{\perp}\cap\Omega}+\boldsymbol{H}_{\Omega^{\perp}}$, one can proceed as follows

$$\begin{aligned}
\frac{2\epsilon_1}{m}&\geq\frac{1}{m}\left\|\mathcal{B}\left(\boldsymbol{H}\right)\right\|_{1}\\
&\geq\frac{1}{m}\left\|\mathcal{B}\left(\boldsymbol{H}_{T\cap\Omega}+\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(1)}+\boldsymbol{H}_{\Omega^{\perp}}^{(1)}\right)\right\|_{1}-\sum_{i=2}^{M_1}\frac{1}{m}\left\|\mathcal{B}\left(\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(i)}\right)\right\|_{1}-\sum_{i=2}^{M_2}\frac{1}{m}\left\|\mathcal{B}\left(\boldsymbol{H}_{\Omega^{\perp}}^{(i)}\right)\right\|_{1}\\
&\geq\left(1-\delta_{2K_1,2K_2}^{\mathrm{lb}}\right)\left\|\boldsymbol{H}_{T\cap\Omega}+\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(1)}+\boldsymbol{H}_{\Omega^{\perp}}^{(1)}\right\|_{\mathrm{F}}-\frac{\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_1}}\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}\right\|_{*}-\frac{\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_2}}\left\|\boldsymbol{H}_{\Omega^{\perp}}\right\|_{1}\\
&\geq\left(1-\delta_{2K_1,2K_2}^{\mathrm{lb}}\right)\left\|\boldsymbol{H}_{T\cap\Omega}+\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(1)}+\boldsymbol{H}_{\Omega^{\perp}}^{(1)}\right\|_{\mathrm{F}}-\frac{\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_1}}\left(\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}\right\|_{*}+\lambda\left\|\boldsymbol{H}_{\Omega^{\perp}}\right\|_{1}\right)\\
&\geq\frac{\left(1-\delta_{2K_1,2K_2}^{\mathrm{lb}}\right)}{\sqrt{3}}\left(\left\|\boldsymbol{H}_{T\cap\Omega}\right\|_{\mathrm{F}}+\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(1)}\right\|_{\mathrm{F}}+\left\|\boldsymbol{H}_{\Omega^{\perp}}^{(1)}\right\|_{\mathrm{F}}\right)-\frac{\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_1}}\left(\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}\right\|_{*}+\lambda\left\|\boldsymbol{H}_{\Omega^{\perp}}\right\|_{1}\right).
\end{aligned}$$

This taken collectively with (D.30) gives

$$\begin{aligned}
&\frac{2\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_1}}\left(\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_{*}+\lambda\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_{1}\right)+\frac{2\epsilon}{m}\\
&\quad\geq\left(\frac{1-\delta_{2K_1,2K_2}^{\mathrm{lb}}}{\sqrt{3}}-\frac{3\left(1+\delta_{K_1,K_2}^{\mathrm{ub}}\right)}{\sqrt{K_1}}\right)\left(\left\|\boldsymbol{H}_{T\cap\Omega}\right\|_{\mathrm{F}}+\left\|\boldsymbol{H}_{T^{\perp}\cap\Omega}^{(1)}\right\|_{\mathrm{F}}+\left\|\boldsymbol{H}_{\Omega^{\perp}}^{(1)}\right\|_{\mathrm{F}}\right).
\end{aligned}$$

Therefore, if we know that

$$\frac{\frac{1-\delta^{\mathrm{lb}}_{2K_1,2K_2}}{\sqrt{3}} - \frac{3\left(1+\delta^{\mathrm{ub}}_{K_1,K_2}\right)}{\sqrt{K_1}}}{2\max\left\{\frac{1+\delta^{\mathrm{ub}}_{K_1,K_2}}{\sqrt{K_1}},1\right\}} \geq \beta_3 > 0$$

for some absolute constant β_3, then

$$\left\|\boldsymbol{H}_{T\cap\Omega}\right\|_{\mathrm{F}} + \left\|\boldsymbol{H}^{(1)}_{T^\perp\cap\Omega}\right\|_{\mathrm{F}} + \left\|\boldsymbol{H}^{(1)}_{\Omega^\perp}\right\|_{\mathrm{F}} \leq \frac{1}{\beta_3}\left(\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_* + \lambda\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_1 + \frac{\epsilon_1}{m}\right). \tag{D.33}$$

On the other hand, we know from (D.31) and (D.32) that

$$\begin{aligned}
\sum_{i=2}^{M_1}\left\|\boldsymbol{H}^{(i)}_{T^\perp\cap\Omega}\right\|_{\mathrm{F}} + \sum_{i=2}^{M_2}\left\|\boldsymbol{H}^{(i)}_{\Omega^\perp}\right\|_{\mathrm{F}} &\leq \frac{1}{1-\delta^{\mathrm{lb}}_{K_1,K_2}}\sum_{i=2}^{M_1}\left\|\mathcal{B}\left(\boldsymbol{H}^{(i)}_{T^\perp\cap\Omega}\right)\right\|_1 + \sum_{i=2}^{M_2}\left\|\mathcal{B}\left(\boldsymbol{H}^{(i)}_{\Omega^\perp}\right)\right\|_1 \\
&\leq \frac{1+\delta^{\mathrm{ub}}_{K_1,K_2}}{\left(1-\delta^{\mathrm{lb}}_{K_1,K_2}\right)\sqrt{K_1}}\left\|\boldsymbol{H}_{T^\perp\cap\Omega}\right\|_* + \frac{\left(1+\delta^{\mathrm{ub}}_{K_1,K_2}\right)}{\left(1-\delta^{\mathrm{lb}}_{K_1,K_2}\right)\sqrt{K_2}}\left\|\boldsymbol{H}_{\Omega^\perp}\right\|_1 \\
&= \frac{1+\delta^{\mathrm{ub}}_{K_1,K_2}}{\left(1-\delta^{\mathrm{lb}}_{K_1,K_2}\right)\sqrt{K_1}}\left(\left\|\boldsymbol{H}_{T^\perp\cap\Omega}\right\|_* + \lambda\left\|\boldsymbol{H}_{\Omega^\perp}\right\|_1\right) \\
&\leq \frac{1+\delta^{\mathrm{ub}}_{K_1,K_2}}{\left(1-\delta^{\mathrm{lb}}_{K_1,K_2}\right)\sqrt{K_1}}\left(3\left\|\boldsymbol{H}_{T\cap\Omega}\right\| + 2\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_* + 2\lambda\left\|\boldsymbol{X}_{\mathrm{c}}\right\|_1\right),
\end{aligned}$$

where the last inequality arises from (D.30). This together with (D.33) completes the proof.

D.5 Proof of Lemma 5.3

Simple calculation yields that

$$\mathbb{E}\left[\boldsymbol{A}_i\left\langle\boldsymbol{A}_i,\boldsymbol{X}\right\rangle\right] = 2\boldsymbol{X} + \left(1+\frac{\mu_4-3}{n}\right)\mathrm{tr}\left(\boldsymbol{X}\right)\cdot\boldsymbol{I}. \tag{D.34}$$

When $\mu_4 = 3$, one can see that

$$\mathbb{E}\left[\boldsymbol{B}_i \left\langle \boldsymbol{B}_i, \boldsymbol{X} \right\rangle\right] = \frac{1}{4}\mathbb{E}\left(\boldsymbol{A}_{2i} - \boldsymbol{A}_{2i+1}\right)\left\langle \boldsymbol{A}_{2i} - \boldsymbol{A}_{2i+1}, \boldsymbol{X} \right\rangle = \boldsymbol{X}. \tag{D.35}$$

When $\mu_4 \neq 3$, consider the linear combination

$$\boldsymbol{B} = a\boldsymbol{A}_1 + b\boldsymbol{A}_2 + c\boldsymbol{A}_3,$$

where we aim to find the coefficients a, b and c that makes $\boldsymbol{B}$ isotropic. If we further require

$$\mathbb{E}\left[\boldsymbol{B}\right] = a + b + c = \frac{\epsilon}{\sqrt{n}}, \tag{D.36}$$

then one can compute

$$\mathbb{E}\left[\boldsymbol{B}\left\langle \boldsymbol{B}, \boldsymbol{X} \right\rangle\right] = 2\left(a^2 + b^2 + c^2\right)\boldsymbol{X} + \left[\left(1 + \frac{\mu_4 - 3}{n}\right)\left(a^2 + b^2 + c^2\right) + 2\left(ab + bc + ac\right)\right]\mathrm{tr}\left(\boldsymbol{X}\right)\cdot\boldsymbol{I}.$$

Our goal is thus to determine a, b and c that satisfy

$$\left(1 + \frac{\mu_4 - 3}{n}\right)\left(a^2 + b^2 + c^2\right) + 2\left(ab + bc + ac\right) = 0,$$

which combined with (D.36) gives

$$\frac{\mu_4 - 3}{n}\left(a^2 + b^2 + c^2\right) + \frac{\epsilon^2}{n} = 0. \tag{D.37}$$

If we set $a = 1$, then (D.37) reduces to

$$\frac{\mu_4 - 3}{n}\left(1 + b^2 + \left(\frac{\epsilon}{\sqrt{n}} - 1 - b\right)^2\right) + \frac{\epsilon^2}{n} = 0$$

$$\Rightarrow \quad b^2 + b\left(1 - \frac{\epsilon}{\sqrt{n}}\right) + \frac{1}{2}\left(1 - \frac{\epsilon}{\sqrt{n}}\right)^2 + \frac{1}{2} + \frac{\epsilon^2}{2\left(\mu_4 - 3\right)} = 0.$$

Solving this quadratic equation yields

$$b = \frac{1}{2}\left(-\left(1 - \frac{\epsilon}{\sqrt{n}}\right) + \sqrt{\Delta}\right); \quad c = \frac{1}{2}\left(-\left(1 - \frac{\epsilon}{\sqrt{n}}\right) - \sqrt{\Delta}\right), \tag{D.38}$$

where

$$\Delta := \left(1-\frac{\epsilon}{\sqrt{n}}\right)^2 - 4\left(\frac{1}{2}\left(1-\frac{\epsilon}{\sqrt{n}}\right)^2+\frac{1}{2}+\frac{\epsilon^2}{2\left(\mu_4-3\right)}\right) = -\left(1-\frac{\epsilon}{n}\right)^2 - 2 - \frac{2\epsilon^2}{\mu_4-3}.$$

Note that $\Delta > 0$ when $\epsilon^2 > 1.5 \cdot |3-\mu_4|$. Also, b and c satisfy

$$1+b^2+c^2 = \epsilon^2/\left(3-\mu_4\right). \tag{D.39}$$

By choosing $\alpha = \sqrt{\frac{3-\mu_4}{2\epsilon^2}}$, $\beta = b\alpha$, and $\gamma = c\alpha$, we derive the form of $\boldsymbol{B}_i$ as introduced in (5.31), which satisfies

$$\mathbb{E}\left[\boldsymbol{B}_i \left\langle \boldsymbol{B}_i, \boldsymbol{X}\right\rangle\right] = \boldsymbol{X}.$$

Finally, we remark that for *any* norm $\|\cdot\|_{\mathrm{n}}$. This can be easily bounded as follows

$$\begin{aligned}\|\boldsymbol{B}_i\|_{\mathrm{n}} &\le \sqrt{\frac{|3-\mu_4|}{2\epsilon^2}}\left(1+|b|+|c|\right)\max_{i:1\le i\le m}\|\boldsymbol{A}_i\|_{\mathrm{n}} \le \sqrt{3}\sqrt{\frac{|3-\mu_4|}{2\epsilon^2}\left(1+b^2+c^2\right)}\max_{i:1\le i\le m}\|\boldsymbol{A}_i\|_{\mathrm{n}} \\ &= \sqrt{3}\max_{i:1\le i\le m}\|\boldsymbol{A}_i\|_{\mathrm{n}}, \end{aligned} \tag{D.40}$$

concluding the proof.

D.6 Proof of Lemma 5.4

Let $\boldsymbol{M}$ represent the symmetric Toeplitz matrix as follows

$$\boldsymbol{M} = \left[\boldsymbol{M}_{|i-l|}\right]_{1\le i,l\le n} := \mathcal{T}\left(\boldsymbol{z}\boldsymbol{z}^*\right),$$

and hence

$$\boldsymbol{M}_k := \frac{1}{n-k}\sum_{l=k+1}^{n}\boldsymbol{z}_l\boldsymbol{z}_{l-k}, \quad 0\le k<n.$$

Apparently, one has $\mathbb{E}[\boldsymbol{M}_0] = 1$ and $\mathbb{E}[\boldsymbol{M}_k] = 0$ $(1\le k<n)$.

The harmonic structure of the Toeplitz matrix $\boldsymbol{M}$ motivates us to embed it into a circulant matrix $\boldsymbol{C}_{\boldsymbol{M}}$. Specifically, a $(2n-1)\times(2n-1)$ circulant matrix

$$\boldsymbol{C}_{\boldsymbol{M}} := \begin{bmatrix} c_0 & c_1 & \cdots & c_{2n-2} \\ c_{2n-2} & c_0 & c_1 & c_2 \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{bmatrix}$$

is constructed such that

$$c_i := \begin{cases} \boldsymbol{M}_i, & \text{if } 0 \le i < n; \\ \boldsymbol{M}_{2n-i-1}, & \text{if } n \le i \le 2n-2. \end{cases}$$

Since $\boldsymbol{M}$ is a submatrix of $\boldsymbol{C}_{\boldsymbol{M}}$, it suffices to bound the spectral norm of $\boldsymbol{C}_{\boldsymbol{M}}$. Define $\omega_i := \exp\left(\frac{2\pi j}{2n-1}\cdot i\right)$, then the corresponding eigenvalues of $\boldsymbol{C}_{\boldsymbol{M}}$ are given by

$$\lambda_i := \sum_l c_l \omega_i^l = \boldsymbol{M}_0 + \sum_{l=1}^{n-1} \boldsymbol{M}_l \omega_i^l + \sum_{l=n}^{2n-2} \boldsymbol{M}_{2n-l-1}\omega_i^l = \boldsymbol{M}_0 + 2\sum_{l=1}^{n-1} \boldsymbol{M}_l \cos\left(\frac{2\pi i l}{2n-1}\right),\ 0 \le i \le 2n-2,$$

which satisfies $\mathbb{E}[\lambda_i] = \mathbb{E}[\boldsymbol{M}_0] = 1$. This leads to an upper bound as follows

$$\|\boldsymbol{M}\| \le \|\boldsymbol{C}_{\boldsymbol{M}}\| \le \max_{0\le i\le 2n-2} |\lambda_i|\,. \tag{D.41}$$

Note that λ_i is a quadratic form in $\{\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_n\}$. Define the symmetric coefficient matrix $\boldsymbol{G}^{(i)}$ such that

$$\forall \alpha, \beta\ (1 \le \alpha, \beta \le n): \quad \boldsymbol{G}^{(i)}_{\alpha,\beta} = \frac{1}{n-|l|}\cos\left(\frac{2\pi i\,|l|}{2n-1}\right), \quad \text{if } \alpha - \beta = l,$$

satisfying

$$\lambda_i = \mathbb{E}[\boldsymbol{M}_0] + \sum_{1\le\alpha,\beta\le n} \boldsymbol{G}^{(i)}_{\alpha,\beta}\left(\boldsymbol{z}_\alpha \boldsymbol{z}_\beta - \mathbb{E}[\boldsymbol{z}_\alpha \boldsymbol{z}_\beta]\right) = 1 + \sum_{1\le\alpha,\beta\le n} \boldsymbol{G}^{(i)}_{\alpha,\beta}\left(\boldsymbol{z}_\alpha \boldsymbol{z}_\beta - \mathbb{E}[\boldsymbol{z}_\alpha \boldsymbol{z}_\beta]\right).$$

When $\boldsymbol{z}$ are drawn from a sub-Gaussian measure, Lemma D.1 asserts that there exists an absolute constant $c_{10} > 0$ such that

$$\mathbb{P}\left(|\lambda_i - 1| \geq t\right) \leq \exp\left(-c_{10}\min\left\{\frac{t}{\|\boldsymbol{G}^{(i)}\|}, \frac{t^2}{\|\boldsymbol{G}^{(i)}\|_{\mathrm{F}}^2}\right\}\right) \tag{D.42}$$

holds for any $t > 0$.

It remains to compute $\|\boldsymbol{G}^{(i)}\|_{\mathrm{F}}$ and $\|\boldsymbol{G}^{(i)}\|$. Since $\boldsymbol{G}^{(i)}$ is a symmetric Toeplitz matrix, we have

$$\|\boldsymbol{G}^{(i)}\|_{\mathrm{F}}^2 = \sum_{\alpha,\beta=1}^{n} |\boldsymbol{G}_{\alpha,\beta}|^2 \leq 2\sum_{l=0}^{n-1}\frac{1}{n-l} \leq 2\log n. \tag{D.43}$$

It then follows that

$$\|\boldsymbol{G}^{(i)}\| \leq \|\boldsymbol{G}^{(i)}\|_{\mathrm{F}} \leq \sqrt{2\log n}. \tag{D.44}$$

Substituting these two bounds into (D.42) immediately yields that there exists a constant $c_{12} > 0$ such that

$$\lambda_i \leq c_{12}\log^{\frac{3}{2}} n, \quad 1 \leq i \leq 2n-2 \tag{D.45}$$

with probability at least $1 - \frac{1}{n^{10}}$. This together with (D.41) concludes the proof.

D.7 Proof of Lemma 5.5

Out of technical convenience, we introduce another collection of events

$$\forall 1 \leq i \leq m: \quad F_i := \left\{\|\boldsymbol{B}_i\|_{\mathrm{F}} \leq 20n\log n\right\}.$$

Since the restriction of $\mathcal{B}_i$ to Toeplitz matrices is isotropic and $\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T} \succeq 0$, we have $\mathcal{T} = \mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\right] \succeq \mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_E\right] \succeq \mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E\cap F_i}\right]$, which yields

$$\left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_E\right] - \mathcal{T}\right\| \leq \left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E\cap F_i}\right] - \mathcal{T}\right\|. \tag{D.46}$$

Thus, it is sufficient to evaluate $\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathbf{1}_{E\cap F_i}\right]-\mathcal{T}\|$. To this end, we adopt an argument of similar spirit as [91, Appendix B]. Write

$$\mathcal{T}=\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\right]=\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E\cap F_i}\right]+\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E^c\cup F_i^c}\right],$$

and, consequently,

$$\begin{aligned}\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E\cap F_i}\right]-\mathcal{T}\| &= \left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E^c\cup F_i^c}\right]\right\| \\ &\le \|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{F_i\cap E^c}\right]\|+\left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{F_i^c}\right]\right\|, \end{aligned} \tag{D.47}$$

which allows us to bound $\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{F_i\cap E^c}\right]\|$ and $\left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{F_i^c}\right]\right\|$ separately.

First, it follows from the identity $\|\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\|=\|\mathcal{T}\left(\boldsymbol{B}_i\right)\|_{\mathrm{F}}^2$ and the definition of the event F_i that

$$\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{F_i\cap E^c}\right]\|\le\left(20n\log n\right)^2\mathbb{P}\left(E^c\right)<\frac{1}{n^2}. \tag{D.48}$$

Second, applying the tail inequality on the quadratic form (e.g. [143, Proposition 1.1]) yields

$$\mathbb{P}\left(\|\boldsymbol{A}_i\|_{\mathrm{F}}\ge c_{20}\left(n+2\sqrt{nt}+2t\right)\right)\le e^{-t}. \tag{D.49}$$

Thus, for any $t>\left(20n\log n\right)^2$, one has

$$\mathbb{P}\left(\|\boldsymbol{A}_i\|_{\mathrm{F}}\ge\sqrt{t/3}\right)\le e^{-c_{21}\sqrt{t}} \tag{D.50}$$

for some constant $c_{21}>0$. Recall that $\|\boldsymbol{B}_i\|_{\mathrm{F}}\le\sqrt{3}\max\left\{\|\boldsymbol{A}_{3i-2}\|_{\mathrm{F}},\|\boldsymbol{A}_{3i-1}\|_{\mathrm{F}},\|\boldsymbol{A}_{3i}\|_{\mathrm{F}}\right\}$, which indicates

$$\begin{aligned}\mathbb{P}\left(\|\boldsymbol{B}_i\|_{\mathrm{F}}^2\ge t\right) &\le \mathbb{P}\left(\|\boldsymbol{A}_{3i-1}\|_{\mathrm{F}}^2\ge\frac{t}{3}\right)+\mathbb{P}\left(\|\boldsymbol{A}_{3i-2}\|_{\mathrm{F}}^2\ge\frac{t}{3}\right)+\mathbb{P}\left(\|\boldsymbol{A}_{3i}\|_{\mathrm{F}}^2\ge\frac{t}{3}\right) \\ &\le 3\mathbb{P}\left(\|\boldsymbol{A}_i\|_{\mathrm{F}}\ge\sqrt{\frac{t}{3}}\right)\le 3e^{-c_{21}\sqrt{t}}:=g(t).\end{aligned}$$

A similar approach as introduced in [91, Appendix B] gives

$$\begin{aligned}\left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{F_i^c}\right]\right\| &\leq \mathbb{E}\left[\|\boldsymbol{B}_i\|_{\mathrm{F}}^2\,\mathbf{1}_{F_i^c}\right] \leq (20n\log n)^2\, g\left((20n\log n)^2\right) + \int_{(20n\log n)^2}^{\infty} g(t)\,\mathrm{d}t \\ &< (20n\log n)^2\, g\left((20n\log n)^2\right) + \int_{(20n\log n)^2}^{\infty} \frac{1}{t^5}\mathrm{d}t < \frac{c_{15}}{n^2} \qquad \text{(D.51)}\end{aligned}$$

for some absolute constant $c_{15} > 0$. This taken collectively with (D.46), (D.47) and (D.48) yields

$$\left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E}\right] - \mathcal{T}\right\| \leq \left\|\mathbb{E}\left[\mathcal{T}\mathcal{B}_i^*\mathcal{B}_i\mathcal{T}\mathbf{1}_{E\cap F_i}\right] - \mathcal{T}\right\| \leq \frac{\tilde{c}_{15}}{n^2}$$

for some absolute constant $\tilde{c}_{15} > 0$.

D.8 Proof of Lemma D.3

Dudley's inequality [144, Theorem 11.17] allows us to bound the supremum of the Gaussian process as follows

$$\mathbb{E}\left[\sup_{T\in\mathcal{M}_r^2,\boldsymbol{X}\in T,\|\boldsymbol{X}\|_{\mathrm{F}}=1}\left|g_i\sum_{i=1}^{m}\left|\mathcal{B}_i(\boldsymbol{X})\right|^2\right|\right] \leq 24\int_0^{\infty}\log^{\frac{1}{2}} N\left(\mathcal{D}_{2r}^2, d(\cdot,\cdot), u\right)\mathrm{d}u, \qquad \text{(D.52)}$$

where $\mathcal{D}_r^2 := \{\boldsymbol{X} \mid \|\boldsymbol{X}\|_{\mathrm{F}} = 1, \mathrm{rank}(\boldsymbol{X}) \leq 2r\}$. Here, $N(\mathcal{Z}, d(\cdot,\cdot), u)$ denotes the smallest number of balls of radius u centered in points of $\mathcal{Z}$ needed to cover the set $\mathcal{Z}$, under the pseudo metric $d(\cdot,\cdot)$ defined as follows

$$d(\boldsymbol{X},\boldsymbol{Y}) := \sqrt{\sum_{i=1}^{m}\left(\left|\mathcal{B}_i(\boldsymbol{X})\right|^2 - \left|\mathcal{B}_i(\boldsymbol{Y})\right|^2\right)^2}.$$

For any $(\boldsymbol{X},\boldsymbol{Y})$ that satisfy $\|\boldsymbol{X}\|_{\mathrm{F}}=\|\boldsymbol{Y}\|_{\mathrm{F}}=1$, $\mathrm{rank}\,(\boldsymbol{X})\leq r$ and $\mathrm{rank}\,(\boldsymbol{Y})\leq r$, the pseudo metric satisfies

$$\begin{aligned}
d\left(\boldsymbol{X},\boldsymbol{Y}\right) &\leq \sqrt{\left(\max_{i:1\leq i\leq m}\left|\mathcal{B}_i\left(\boldsymbol{X}-\boldsymbol{Y}\right)\right|^2\right)\sum_{i=1}^{m}\left|\mathcal{B}_i\left(\boldsymbol{X}+\boldsymbol{Y}\right)\right|^2} \\
&\leq \sqrt{2}\sqrt{\sum_{i=1}^{m}\left|\mathcal{B}_i\left(\boldsymbol{X}\right)\right|^2+\left|\mathcal{B}_i\left(\boldsymbol{Y}\right)\right|^2}\max_{i:1\leq i\leq m}\left|\mathcal{B}_i\left(\boldsymbol{X}-\boldsymbol{Y}\right)\right| \\
&\leq \sqrt{2}\left\{\sqrt{\left\langle \boldsymbol{X},\left(\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i\right)\left(\boldsymbol{X}\right)\right\rangle}+\sqrt{\left\langle \boldsymbol{Y},\left(\sum_{i=1}^{m}\mathcal{B}_i^*\mathcal{B}_i\right)\left(\boldsymbol{Y}\right)\right\rangle}\right\}\max_{i:1\leq i\leq m}\left|\mathcal{B}_i\left(\boldsymbol{X}-\boldsymbol{Y}\right)\right| \\
&\leq 2\sqrt{2}\sup_{T:T\in\mathcal{M}_r^2}\sqrt{\left\|\sum_{i=1}^{m}\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|}\max_{i:1\leq i\leq m}\left|\left\langle \boldsymbol{B}_i,\boldsymbol{X}-\boldsymbol{Y}\right\rangle\right|,
\end{aligned}$$

where the last inequality relies on the observation that $\|\boldsymbol{X}\|_{\mathrm{F}}=\|\boldsymbol{Y}\|_{\mathrm{F}}=1$.

If we introduce the quantity

$$R:=\sup_{T:T\in\mathcal{M}_r^2}\sqrt{\left\|\sum_{i=1}^{m}\mathcal{P}_T\mathcal{B}_i^*\mathcal{B}_i\mathcal{P}_T\right\|} \tag{D.53}$$

and define another pseudo metric $\|\cdot\|_{\mathcal{B}}$ as

$$\|\boldsymbol{X}\|_{\mathcal{B}}:=\max_{i:1\leq i\leq m}\left|\left\langle \boldsymbol{B}_i,\boldsymbol{X}\right\rangle\right|, \tag{D.54}$$

then $d\left(\boldsymbol{X},\boldsymbol{Y}\right)\leq 2\sqrt{2}R\left\|\boldsymbol{X}-\boldsymbol{Y}\right\|_{\mathcal{B}}$, which allows us to bound

$$\begin{aligned}
\int_0^{\infty}\log^{\frac{1}{2}}N\left(\mathcal{D}_{2r}^2,d\left(\cdot,\cdot\right),u\right)\mathrm{d}u &\leq \int_0^{\infty}\log^{\frac{1}{2}}N\left(\mathcal{D}_{2r}^2,2\sqrt{2}R\left\|\cdot\right\|_{\mathcal{B}},u\right)\mathrm{d}u \\
&= \int_0^{\infty}\log^{\frac{1}{2}}N\left(\frac{1}{\sqrt{2r}}\mathcal{D}_{2r}^2,\left\|\cdot\right\|_{\mathcal{B}},\frac{u}{4R\sqrt{r}}\right)\mathrm{d}u \\
&\leq \int_0^{\infty}\log^{\frac{1}{2}}N\left(\mathcal{D}_{2r}^1,\left\|\cdot\right\|_{\mathcal{B}},\frac{u}{4R\sqrt{r}}\right)\mathrm{d}u \qquad \text{(D.55)} \\
&\leq\ 4R\sqrt{r}\int_0^{\infty}\log^{\frac{1}{2}}N\left(\mathcal{D}^1,\left\|\cdot\right\|_{\mathcal{B}},u\right)\mathrm{d}u. \qquad \text{(D.56)}
\end{aligned}$$

Here, $\mathcal{D}_r^1 := \{\boldsymbol{X} \mid \|\boldsymbol{X}\|_* \leq 1, \mathrm{rank}\,(\boldsymbol{X}) \leq r\}$, $\mathcal{D}^1 := \{\boldsymbol{X} \mid \|\boldsymbol{X}\|_* \leq 1\}$, and we have exploited the containment $\frac{1}{\sqrt{2r}}\mathcal{D}_{2r}^2 \subseteq \mathcal{D}_{2r}^1 \subseteq \mathcal{D}^1$. Hence it suffices to bound

$$E_2 := 4R\sqrt{r}\int_0^\infty \log^{\frac{1}{2}} N\left(\mathcal{D}^1, \|\cdot\|_{\mathcal{B}}, u\right) \mathrm{d}u.$$

It remains to bound the covering number (or metric entropy) of the nuclear-norm ball $\mathcal{D}^1$. Repeating the well-known procedure as in [145, Page 1113] yields

$$\int_0^\infty \sqrt{\log N\left(\mathcal{D}^1, \|\cdot\|_{\mathcal{B}}, u\right)}\mathrm{d}u \leq C_{10} K \left(\log n\right)^{5/2} \sqrt{\log m} \leq C_{11} K \log^3 n$$

for some constants $C_{10}, C_{11} > 0$. This taken collectively with (D.52) and (D.56) gives that conditioning on $\mathcal{B}_i$'s, one has

$$\mathbb{E}\left[\sup_{T \in \mathcal{M}_r^2} \left\|\mathcal{P}_T\left(\sum_{i=1}^m g_i \mathcal{B}_i^* \mathcal{B}_i\right)\mathcal{P}_T\right\|\right] \leq C_{14}\sqrt{r} K \log^3 n \sqrt{\sup_{T: T \in \mathcal{M}_r^2} \left\|\sum_{i=1}^m \mathcal{P}_T \mathcal{B}_i^* \mathcal{B}_i \mathcal{P}_T\right\|}. \tag{D.57}$$

for some absolute constant $C_{14} > 0$.

Bibliography

[1] D.L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289 –1306, April 2006.

[2] C. E. Shannon. *A mathematical theory of communication.* University of Illinois Press, Urbana, 1949.

[3] R. G. Gallager. *Information theory and reliable communication.* John Wiley & Sons, Inc, New York, 1968.

[4] P. Bello. Characterization of randomly time-variant linear channels. *IEEE Trans. on Communications Systems*, 11(4):360 –393, Dec. 1963.

[5] M. Medard. The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel. *IEEE Trans. on Info Theory*, 46(3):933 –946, May. 2000.

[6] A.J. Jerri. The Shannon sampling theorem: Its various extensions and applications: A tutorial review. *Proceedings of the IEEE*, 65(11):1565 – 1596, Nov. 1977.

[7] A. Aldroubi and K. Gröchenig. Nonuniform sampling and reconstruction in shift-invariant spaces. *SIAM Review*, 43(4):585–620, 2001.

[8] R. M. Yeung. *An Introduction to Nonharmonic Fourier Series.* Academic Press, 2001.

[9] J. Yen. On nonuniform sampling of bandwidth-limited signals. *IRE Transactions on Circuit Theory*, 3(4):251 – 257, December 1956.

[10] A. Papoulis. Generalized sampling expansion. *IEEE Transactions on Circuits and Systems*, 24(11):652 – 654, Nov 1977.

[11] P. L. Butzer and R. L. Stens. Sampling theory for not necessarily band-limited functions: A historical overview. *SIAM Review*, 34(1):40–53, 1992.

[12] M. Mishali and Y. C. Eldar. Sub-Nyquist sampling: Bridging theory and practice. *IEEE Signal Processing Magazine*, 28(6):98–124, Nov. 2011.

[13] H. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Mathematica*, 117:37–52, 1967.

[14] Yuan-Pei Lin and P.P. Vaidyanathan. Periodically nonuniform sampling of bandpass signals. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(3):340 –351, Mar. 1998.

[15] M. Unser and J. Zerubia. A generalized sampling theory without band-limiting constraints. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(8):959 –969, Aug. 1998.

[16] Toby Berger. *Nyquist's Problem in Data Transmission Theory*. PhD dissertation, Harvard University, Cambridge, Dec. 1965.

[17] D. Chan and R. Donaldson. Optimum pre- and postfiltering of sampled signals with application to pulse modulation and data compression systems. *IEEE Trans. on Communication Technology*, 19(2):141 –157, Apr. 1971.

[18] T. Ericson. Optimum PAM filters are always band limited (corresp.). *IEEE Transactions on Information Theory*, 19(4):570 – 573, Jul 1973.

[19] E. Biglieri, J. Proakis, and S. Shamai. Fading channels: information-theoretic and communications aspects. *IEEE Transactions on Information Theory*, 44(6):2619 –2692, Oct 1998.

[20] Abbas El Gamal and Y. H. Kim. *Network Information theory*. Cambridge University Press, 2011.

[21] David Blackwell, Leo Breiman, and A. J. Thomasian. The capacity of a class of channels. *The Annals of Mathematical Statistics*, 30(4):1229–1241, 1959.

[22] M. Effros, A. Goldsmith, and Yifan Liang. Generalizing capacity: New definitions and capacity theorems for composite channels. *IEEE Transactions on Information Theory*, 56(7):3069 –3087, July 2010.

[23] E. J. Candes, J. K. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, Feb. 2006.

[24] E. J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE Transactions on Information Theory*, 52(12):5406–5425, Dec. 2006.

[25] David L Donoho, Arian Maleki, and Andrea Montanari. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.

[26] D.L. Donoho, A. Javanmard, and A. Montanari. Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing. Dec. 2011.

[27] Yihong Wu and S. Verdu. Optimal phase transitions in compressed sensing. *IEEE Transactions on Information Theory*, 58(10):6241 –6263, Oct. 2012.

[28] A. Guionnet and O. Zeitouni. Concentration of the spectral measure for large matrices. *Electronic Communications in Probability*, (5):119 –136, 2000.

[29] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[30] T. Zhang, S. Chowdhury, M. Lustig, R. A. Barth, M. T. Alley, T. Grafendorfer, P. D. Calderon, F. J.L. Robb, J. M. Pauly, and S. S. Vasanawala. Clinical

performance of contrast enhanced abdominal pediatric MRI with fast combined parallel imaging compressed sensing reconstruction. *Journal of Magnetic Resonance Imaging*, 40(1):13–25, July 2014.

[31] L.C. Potter, E. Ertin, J.T. Parker, and M. Cetin. Sparsity and compressed sensing in radar imaging. *Proceedings of the IEEE*, 98(6):1006–1020, 2010.

[32] Yuejie Chi, Yao Xie, and Robert Calderbank. Compressive demodulation of mutually interfering signals. *submitted to IEEE Transactions on Information Theory*, 2013.

[33] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.

[34] R. Roy and T. Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(7):984 –995, Jul 1989.

[35] Y. Hua and T. K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(5):814 –824, may 1990.

[36] Martin Vetterli, Pina Marziliano, and Thierry Blu. Sampling signals with finite rate of innovation. *IEEE Trans on Signal Proc.*, 50(6):1417–1428, 2002.

[37] Kfir Gedalyahu, Ronen Tur, and Yonina. C. Eldar. Multichannel sampling of pulse streams at the rate of innovation. *IEEE Transactions on Signal Processing*, 59(4):1491–1504, 2011.

[38] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets strang-fix. *IEEE Transactions on Signal Processing*, 55(5):1741 –1757, May 2007.

[39] Y. C. Eldar and Gitta Kutyniok. *Compressed Sensing: Theory and Applications*, volume 95. Cambridge University Press Cambridge, 2012.

[40] E. J. Candes, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, 2006.

[41] Xiaodong Li. Compressed sensing and matrix completion with constant proportion of corruptions. *Constructive Approximation*, 37:73–99, 2013.

[42] Y. Chi, L.L. Scharf, A. Pezeshki, and A.R. Calderbank. Sensitivity to basis mismatch in compressed sensing. *IEEE Transactions on Signal Processing*, 59(5):2182–2195, May 2011.

[43] Y. Hua. Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. *IEEE Transactions on Signal Processing*, 40(9):2267 –2280, Sep 1992.

[44] R. C. Daniels and R. W. Heath. 60 GHz wireless communications: emerging requirements and design recommendations. *IEEE Vehicular Technology Magazine*, 2(3):41–50, 2007.

[45] A. Jung, G. Taubock, and F/ Hlawatsch. Compressive spectral estimation for nonstationary random processes. *IEEE Transactions on Information Theory*, 59(5), May 2013.

[46] L. L. Scharf and B. Friedlander. Matched subspace detectors. *IEEE Transactions on Signal Processing*, 42(8):2146–2157, 1994.

[47] M.G. Raymer, M. Beck, and D. McAlister. Complex wave-field reconstruction using phase-space tomography. *Physical review letters*, 72:1137, 1994.

[48] L. Tian, J. Lee, S. B. Oh, and G. Barbastathis. Experimental compressive phase space tomography. *Optics Express*, 20(8):8296, 2012.

[49] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 21(15):2758–2769, 1982.

[50] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225, 2013.

[51] A. Beck and Y. C. Eldar. Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *to appear in SIAM Optimization*, 2013.

[52] Philip Schniter and Sundeep Rangan. Compressive phase retrieval via generalized approximate message passing. In *Annual Allerton Conference on Communication, Control, and Computing*, pages 815–822, 2012.

[53] K. Jaganathan, S. Oymak, and B. Hassibi. Recovery of sparse 1-D signals from the magnitudes of their Fourier transform. In *IEEE ISIT*, pages 1473–1477, 2012.

[54] E. J. Candes, T. Strohmer, and V. Voroninski. PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 2012.

[55] Irène Waldspurger, Alexandre d'Aspremont, and Stéphane Mallat. Phase recovery, maxcut and complex semidefinite programming. *arXiv preprint arXiv:1206.0102*, 2012.

[56] P. Netrapalli, P. Jain, and S. Sanghavi. Phase retrieval using alternating minimization. *Advances in Neural Information Processing Systems (NIPS)*, 2013.

[57] S. Muthukrishnan. *Data streams: Algorithms and applications*. Now Publishers Inc, 2005.

[58] E. J. Candes and X. Li. Solving quadratic equations via PhaseLift when there are about as many equations as unknowns. *Foundations of Computational Mathematics*, 2013.

[59] X. Li and V. Voroninski. Sparse signal recovery from quadratic measurements via convex programming. *SIAM Journal on Mathematical Analysis*, 2013.

[60] M. Rudelson and R. Vershynin. On sparse reconstruction from Fourier and Gaussian measurements. *Communications on Pure and Applied Mathematics*, 61(8):1025–1045, 2008.

[61] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Trans. on Inf. Theory*, 57(3):1548–1566, March 2011.

[62] Yi-Kai Liu. Universal low-rank matrix recovery from pauli measurements. In *Advances in Neural Information Processing Systems*, pages 1638–1646, 2011.

[63] Iain M Johnstone. High dimensional statistical inference and random matrices. In *Proceedings of the International Congress of Mathematicians: Madrid, August 22-30, 2006: invited lectures*, pages 307–333, 2006.

[64] N. Karoui. Operator norm consistent estimation of large-dimensional sparse covariance matrices. *The Annals of Statistics*, pages 2717–2756, 2008.

[65] G. Dasarathy, P. Shah, B. N. Bhaskar, and R. Nowak. Sketching sparse matrices. *arXiv preprint arXiv:1303.6544*, 2013.

[66] Erkki Oja. *Subspace methods of pattern recognition*, volume 4. Research Studies Press England, 1983.

[67] L. Balzano, R. Nowak, and B. Recht. Online identification and tracking of subspaces from highly incomplete information. *Proc. Allerton 2010*, 2010.

[68] Yuejie Chi, Y.C. Eldar, and R. Calderbank. PETRELS: Parallel subspace estimation and tracking by recursive least squares from partial observations. *IEEE Transactions on Signal Processing*, 61(23):5947–5959, 2013.

[69] Yuxin Chen, Y. C. Eldar, and A. J. Goldsmith. Shannon meets Nyquist: Capacity of sampled Gaussian channels. *IEEE Transactions on Information Theory*, 59(8):4889–4914, Aug 2013.

[70] Yuxin Chen, A. J. Goldsmith, and Y. C. Eldar. Channel capacity under sub-Nyquist nonuniform sampling. *IEEE Transactions on Information Theory*, 60(8):4739–4756, Aug 2014.

[71] Yuxin Chen, Andrea J. Goldsmith, and Yonina C. Eldar. Minimax capacity loss under sub-Nyquist universal sampling. *under revision, IEEE Transactions on Information Theory*, 2013.

[72] Yuxin Chen, Andrea J. Goldsmith, and Yonina C. Eldar. Backing off from infinity: Performance bounds via concentration of spectral measure for random MIMO channels. *IEEE Transactions on Information Theory*, 61(1), January 2015.

[73] Y. Chen and Y. Chi. Robust spectral compressed sensing via structured matrix completion. *IEEE Transactions on Information Theory*, 60(10):6576 – 6601, Oct. 2014.

[74] Yuxin Chen and Yuejie Chi. Spectral compressed sensing via structured matrix completion. *International Conference on Machine Learning (ICML)*, June 2013.

[75] Yuxin Chen, Yuejie Chi, and Andrea J. Goldsmith. Exact and stable covariance estimation from quadratic sampling via convex programming. *under revision, IEEE Transactions on Information Theory, arXiv preprint arXiv:1310.0807*, 2013.

[76] A. Beurling. *The Collected Works of Arne Beurling. Vol. 1, Complex Analysis*. Birkhauser, Boston, 1989.

[77] Y. C. Eldar and T. Michaeli. Beyond bandlimited sampling. *IEEE Signal Processing Magazine*, 26(3):48 –68, May. 2009.

[78] M. Mishali and Y. C. Eldar. From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals. *IEEE Journal of Selected Topics in Signal Processing*, 4(2):375 –391, Apr. 2010.

[79] E. Telatar. Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, 10(6):585 – 595, Nov. - Dec. 1999.

[80] Paolo Tilli. Singular values and eigenvalues of non-Hermitian block Toeplitz matrices. *Linear Algebra and its Applications*, 272(1-3):59 – 89, 1998.

[81] M. I. Kadec. The exact value of the Paley-Wiever constant. *Soviet Math. Dokl.*, pages 559 – 561, 1964.

[82] LTE in a nutshell: White paper on physical layer. 2010.

[83] E. L. Lehmann and George Casella. *Theory of Point Estimation*. Springer, New York, 1998.

[84] Maxim Raginsky and Igal Sason. Concentration of measure inequalities in information theory, communications and coding. *Foundations and Trends in Communications and Information Theory*, 10, 2013.

[85] Pascal Massart and Jean Picard. *Concentration Inequalities and Model Selection*. Lecture Notes in Mathematics. Springer, 2007.

[86] Terrance Tao and Van Vu. Random matrices: Universality of local spectral statistics of non-Hermitian matrices. June 2012.

[87] T. Tao and Van Vu. Random covariance matrices: Universality of local statistics of eigenvalues. *The Annals of Probability*, 40(3):1285 – 1315, 2012.

[88] E. J. Candes and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, April 2009.

[89] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010.

[90] James A Cadzow. Spectral estimation: An overdetermined rational model equation approach. *Proceedings of the IEEE*, 70(9):907–939, 1982.

[91] E.J. Candes and Y. Plan. A probabilistic and RIPless theory of compressed sensing. *IEEE Transactions on Information Theory*, 57(11):7235–7254, 2011.

[92] M.F. Duarte and R.G. Baraniuk. Spectral compressive sensing. *Applied and Computational Harmonic Analysis*, 2012.

[93] Albert Fannjiang and Wenjing Liao. Coherence pattern-guided compressive sensing with unresolved grids. *SIAM Journal on Imaging Sciences*, 5(1):179–202, 2012.

[94] Emmanuel J Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.

[95] Emmanuel J Candès and Carlos Fernandez-Granda. Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254, 2013.

[96] Gongguo Tang, B.N. Bhaskar, P. Shah, and B. Recht. Compressed sensing off the grid. *IEEE Trans on Information Theory*, 59(11):7465–7490, Nov 2013.

[97] Gongguo Tang, Badri Narayan Bhaskar, and Benjamin Recht. Near minimax line spectral estimation. 2013.

[98] Y. Chi and Y. Chen. Compressive recovery of 2-D off-grid frequencies. In *Asilomar Conference on Signals, Systems and Computers*, pages 687–691, 2013.

[99] Weiyu Xu, Jian-Feng Cai, Kumar Vijay Mishra, Myung Cho, and Anton Kruger. Precise semidefinite programming formulation of atomic norm minimization for recovering d-dimensional ($d \geq 2$) off-the-grid frequencies. In *ITA Workshop*, pages 1–4, 2014.

[100] Yuanxin Li and Yuejie Chi. Off-the-grid line spectrum denoising and estimation with multiple measurement vectors. *arXiv preprint arXiv:1408.2242*, 2014.

[101] E. J. Candes and T. Tao. The power of convex relaxation: Near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053 –2080, May 2010.

[102] Raghunandan H Keshavan, Andrea Montanari, and Sewoong Oh. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56(6):2980–2998, 2010.

[103] Raghunandan H. Keshavan, Andrea Montanari, and Sewoong Oh. Matrix completion from noisy entries. *Journal of Machine Learning Research*, 11:2057–2078, August 2010.

[104] Yudong Chen. Incoherence-optimal matrix completion. *arXiv preprint arXiv:1310.0154*, 2013.

[105] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of ACM*, 58(3):11:1–11:37, Jun 2011.

[106] S. Negahban and M.J. Wainwright. Restricted strong convexity and weighted matrix completion: Optimal bounds with noise. *The Journal of Machine Learning Research*, 98888:1665–1697, May 2012.

[107] Venkat Chandrasekaran, Sujay Sanghavi, Pablo A Parrilo, and Alan S Willsky. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 21(2):572–596, 2011.

[108] Yudong Chen, Huan Xu, Constantine Caramanis, and Sujay Sanghavi. Robust matrix completion with corrupted columns. *International Conference on Machine Learning (ICML)*, June 2011.

[109] Yudong Chen, Ali Jalali, Sujay Sanghavi, and Constantine Caramanis. Low-rank matrix recovery from errors and erasures. *IEEE Transactions on Information Theory*, 59(7):4324–4337, 2013.

[110] Mingrui Wu. Collaborative filtering via ensembles of matrix factorizations. In *Proceedings of KDD Cup and Workshop*, volume 2007, 2007.

[111] T. Zhang, J. M. Pauly, and I. R. Levesque. Accelerating parameter mapping with a locally low rank constraint. *Magnetic Resonance in Medicine, DOI: 10.1002/mrm.25161*, 2014.

[112] T. Zhang, J. Y. Cheng, A. G. Potnick, R. A. Barth, M. T. Alley, M. Uecker, M. Lustig, J. M. Pauly, and S. S. Vasanawala. Fast pediatric 3D free-breathing

abdominal dynamic contrast enhanced MRI with high spatiotemporal resolution. *Journal of Magnetic Resonance Imaging, DOI: 10.1002/jmri.24551*, 2013.

[113] Wenjing Liao and Albert Fannjiang. MUSIC for single-snapshot spectral estimation: Stability and super-resolution. *arXiv preprint arXiv:1404.1484*, 2014.

[114] Maryam Fazel, Ting Kei Pong, Defeng Sun, and Paul Tseng. Hankel matrix rank minimization with applications to system identification and realization. *SIAM Journal on Matrix Analysis and Applications*, 34(3):946–977, 2013.

[115] A. Sankaranarayanan, P. Turaga, R. Baraniuk, and R. Chellappa. Compressive acquisition of dynamic scenes. *Computer Vision–ECCV 2010*, pages 129–142, 2010.

[116] M Lustig, M Elad, and JM Pauly. Calibrationless parallel imaging reconstruction by structured low-rank matrix completion. In *ISMRM*, page 2870, 2010.

[117] J. F. Cai, E. J. Candes, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

[118] M. Grant, S. Boyd, and Y. Ye. CVX: Matlab software for disciplined convex programming. *Online accessiable: http://stanford. edu/˜ boyd/cvx*, 2008.

[119] E.J. Candes 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.

[120] Henrik Ohlsson, Allen Y Yang, Roy Dong, and S Shankar Sastry. Compressive phase retrieval from squared output measurements via semidefinite programming. *arXiv preprint arXiv:1111.6323*, 2011.

[121] T Tony Cai and Anru Zhang. Rop: Matrix recovery via rank-one projections. *arXiv preprint arXiv:1310.5791*, 2013.

[122] Samet Oymak, Amin Jalali, Maryam Fazel, Yonina C Eldar, and Babak Hassibi. Simultaneously structured models with application to sparse and low-rank matrices. *arXiv preprint arXiv:1212.3753*, 2012.

[123] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346(9):589–592, 2008.

[124] S. Jafarpour, W. Xu, B. Hassibi, and R. Calderbank. Efficient and robust compressed sensing using optimized expander graphs. *IEEE Transactions on Information Theory*, 55(9):4299–4308, 2009.

[125] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, 2008.

[126] Michel Talagrand. Majorizing measures: the generic chaining. *The Annals of Probability*, pages 1049–1103, 1996.

[127] Terence Tao. *Topics in Random Matrix Theory*. Graduate Studies in Mathematics. American Mathematical Society, Providence, Rhode Island, 2012.

[128] Ulf Grenander and Gábor Szegő. *Toeplitz forms and their applications*. Univ of California Press, 1958.

[129] Alon Kipnis, Andrea J. Goldsmith, Tsachy Weissman, and Yonina C. Eldar. Distortion-rate function of sub-Nyquist sampled Gaussian sources. *arXiv preprint arXiv:1405.5329*, 2014.

[130] R. G. Gallager. *Introductory Functional Analysis with Applications*. Wiley, 1989.

[131] R. Gray. *Toeplitz and circulant matrices: A review*. Foundations and Trends in Communications and Information Theory. NOW Publisher, 2006.

[132] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, 1985.

[133] Masanao Aoki. *New Approaches to Macroeconomic Modeling*. Cambridge University Press, 1998.

[134] A.E. Litvak, A. Pajor, M. Rudelson, and N. Tomczak-Jaegermann. Smallest singular value of random matrices and geometry of random polytopes. *Advances in Mathematics*, 195(2):491 – 523, 2005.

[135] R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. *Compressed Sensing, Theory and Applications*, pages 210 – 268, 2012.

[136] KE Wang. Random covariance matrices: Universality of local statistics of eigenvalues up to the edge. *Random Matrices: Theory and Applications*, 1(1):1150005, 2012.

[137] Mark Rudelson and Roman Vershynin. Smallest singular value of a random rectangular matrix. *Communications on Pure and Applied Mathematics*, 62(12):1707–1739, 2009.

[138] Yasunori Fujikoshi, Vladimir V. Ulyanov, and Ryoichi Shimizu. *Multivariate Statistics: High-dimensional and Large-sample Approximations*. Wiley Series in Probability and Statistics. Wiley, Hoboken, New Jersey, 2010.

[139] Zizhong Chen and Jack J. Dongarra. Condition numbers of Gaussian random matrices. *SIAM Journal on Matrix Analysis and Applications*, 27(3):603–620, July 2005.

[140] Joel A Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 12(4):389–434, 2012.

[141] D. L. Hanson and F. T. Wright. A bound on tail probabilities for quadratic forms in independent random variables. *The Annals of Mathematical Statistics*, 42(3):1079–1083, 1971.

[142] M. Rudelson and R. Vershynin. Hanson-Wright inequality and sub-Gaussian concentration. *arXiv preprint arXiv:1306.2872*, June 2013.

[143] D. Hsu, S. M. Kakade, and T. Zhang. A tail inequality for quadratic forms of sub-Gaussian random vectors. *Electronic Communications in Probability*, 17(52), 2012.

[144] Michel Ledoux and Michel Talagrand. *Probability in Banach Spaces: isoperimetry and processes*, volume 23. Springer, 1991.

[145] Guillaume Aubrun. On almost randomizing channels with a short Kraus decomposition. *Communications in Mathematical Physics*, 288(3):1103–1116, 2009.