

# UC San Diego

## UC San Diego Electronic Theses and Dissertations

### Title

Spectral Inference for High-Dimensional Time Series

### Permalink

<https://escholarship.org/uc/item/0q13m5hk>

### ISBN

9798270239718

### Author

Zhang, Chi

### Publication Date

2025-12-20

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Spectral Inference for High-Dimensional Time Series

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy

in

Mathematics

by

Chi Zhang

Committee in charge:

Professor Danna Zhang, Chair  
Professor Ery Arias-Castro  
Professor Dimitris N. Politis  
Professor Jingjing Zou

2025

Copyright

Chi Zhang, 2025

All rights reserved.

The Dissertation of Chi Zhang is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2025

## DEDICATION

For my family, and for my mother, whose love lives on.

## EPIGRAPH

Keep your eyes on the stars, and your feet on the ground.

*Anonymous*

Until the day when God will deign to reveal the future to man,  
all human wisdom is contained in these two words,  
— “Wait and hope.”

*Alexandre Dumas*

## TABLE OF CONTENTS

|                                                                             |      |
|-----------------------------------------------------------------------------|------|
| Dissertation Approval Page .....                                            | iii  |
| Dedication .....                                                            | iv   |
| Epigraph.....                                                               | v    |
| Table of Contents .....                                                     | vi   |
| List of Figures .....                                                       | viii |
| List of Tables.....                                                         | ix   |
| Acknowledgements .....                                                      | x    |
| Vita.....                                                                   | xi   |
| Abstract of the Dissertation .....                                          | xii  |
| Chapter 1    Spectral Inference for High Dimensional Time Series .....      | 1    |
| 1.1    Introduction .....                                                   | 1    |
| 1.2    Consistency and Asymptotic Normality of Spectral Estimate .....      | 5    |
| 1.3    Simultaneous Inference of Spectral Density .....                     | 15   |
| 1.3.1    Gaussian Approximation for High Dimensional Spectral Estimates     | 15   |
| 1.3.2    Applications .....                                                 | 19   |
| 1.3.3    Resampling Methods .....                                           | 21   |
| 1.4    Extension .....                                                      | 25   |
| 1.4.1    Relaxing the Moment and Dependence Conditions .....                | 25   |
| 1.4.2    Lag-window Estimator of Spectral Density .....                     | 27   |
| 1.5    Generalization .....                                                 | 27   |
| 1.5.1    Consistency and Asymptotic Normality of Spectral Estimate .....    | 28   |
| 1.5.2    Gaussian Approximation for High Dimensional Spectral Estimates     | 32   |
| 1.5.3    Resampling Methods .....                                           | 34   |
| 1.6    Numerical Results .....                                              | 35   |
| 1.6.1    Simulation Study.....                                              | 35   |
| 1.6.2    Data Analysis .....                                                | 38   |
| 1.7    Proofs .....                                                         | 40   |
| 1.7.1    Proofs of Results in Section 1.2 .....                             | 41   |
| 1.7.2    Proofs of Results in Section 1.3 .....                             | 46   |
| 1.7.3    Proofs of Results in Section 1.5 .....                             | 56   |
| 1.8    Conclusion .....                                                     | 62   |
| 1.9    Appendix.....                                                        | 62   |
| 1.9.1    Useful Lemmas for Expectation and Covariance of Spectral Estimates | 62   |
| 1.9.2    Probability Inequalities for High Dimensional Time Series .....    | 72   |

|                                                                                                                           |     |
|---------------------------------------------------------------------------------------------------------------------------|-----|
| 1.10 Acknowledgements .....                                                                                               | 97  |
| Chapter 2 Probability and moment inequalities for quadratic forms in independent<br>random variables with fat tails ..... | 98  |
| 2.1 Introduction .....                                                                                                    | 98  |
| 2.2 Main Result .....                                                                                                     | 100 |
| 2.3 Proofs .....                                                                                                          | 106 |
| 2.4 Acknowledgements .....                                                                                                | 114 |
| Bibliography .....                                                                                                        | 115 |

## LIST OF FIGURES

|             |                                                                                                                                                        |    |
|-------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1.1. | Q-Q plot comparing maximum deviation $F^*$ and its Gaussian approximation $ Z _\infty$ when $p = 5$ . . . . .                                          | 37 |
| Figure 1.2. | Q-Q plot comparing maximum deviation $F^*$ and its Gaussian approximation $ Z _\infty$ when $p = 15$ . . . . .                                         | 37 |
| Figure 1.3. | Standardized first-order differences of hourly wind speed series between 02/01/2021 to 04/30/2021 from selected stations in Arizona and Texas. . . . . | 39 |

## LIST OF TABLES

|            |                                                            |    |
|------------|------------------------------------------------------------|----|
| Table 1.1. | Empirical coverages of simultaneous confidence bands. .... | 38 |
| Table 1.2. | Pairwise Cross-Spectral Test Results .....                 | 40 |

## ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to Professor Danna Zhang, the chair of my committee, for her constant support throughout my studies. Her guidance has been invaluable not only in my academic journey but also in my personal life. Her insights in statistics have continuously inspired and shaped my research, while her kindness and encouragement as a friend have helped me through difficult times. I am truly grateful for her mentorship, patience, and kind support, which have had a lasting impact on both my work and my life.

I am also grateful to Professor Dimitris Politis, Professor Ery Arias-Castro, and Professor Jingjing Zou for serving on my Ph.D. committee. Their valuable advice and insightful comments during my advancement talk provided me with great inspiration for my future research.

In addition, I would like to thank my friends for their help and companionship over the years. In particular, I am grateful to Jason Lin, Linbo Liu, Jinjie Zhang, Shitao Li, Chang Liu, and Kun Zhu for their support and friendship. I really appreciate the valuable conversations and memories we have shared together.

Finally, I would like to express my deepest gratitude to my parents, Lidong Zhang and Yan Liu, and to all my family. Without their love, support, and encouragement, completing my Ph.D. studies would not have been possible.

Chapter 1, in part, has been submitted for publication of the material “Spectral Inference for High Dimensional Time Series” by Chi Zhang and Danna Zhang, as it appears in IEEE Transactions on Information Theory, 2025. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, has been submitted for publication of the material “Probability and Moment Inequalities for Quadratic Forms in Independent Random Variables with Fat Tails” by Chi Zhang and Danna Zhang, as it appears in Statistics & Probability Letters, 2025. The dissertation author was the primary investigator and author of this paper.

## VITA

- |      |                                                                                             |
|------|---------------------------------------------------------------------------------------------|
| 2017 | Bachelor of Science in Mathematics and Applied Mathematics<br>Shanghai Jiao Tong University |
| 2018 | Master of Arts in Mathematics<br>University of Wisconsin-Madison                            |
| 2025 | Doctor of Philosophy in Mathematics<br>University of California San Diego                   |

## PUBLICATIONS

C. Zhang and D. Zhang, “Spectral Inference for High Dimensional Time Series”, IEEE Transactions on Information Theory, vol. 71, no. 4, pp. 2909-2929, April 2025.

C. Zhang and D. Zhang, “Probability and Moment Inequalities for Quadratic Forms in Independent Random Variables with Fat Tails”, Statistics & Probability Letters, Volume 217, 2025.

# ABSTRACT OF THE DISSERTATION

Spectral Inference for High-Dimensional Time Series

by

Chi Zhang

Doctor of Philosophy in Mathematics

University of California San Diego, 2025

Professor Danna Zhang, Chair

Spectral analysis is a central tool in the study of time series, yet much of the established distributional theory for spectral density estimation is limited to low-dimensional settings. Modern applications frequently involve high-dimensional data, nonlinear dynamics, and non-Gaussian distributions, for which existing asymptotic results are inadequate. This dissertation develops a comprehensive inference framework for spectral density estimation in high-dimensional time series. The theory provides distributional approximations that hold across different dimensions and frequencies, thereby enabling principled inference in complex dependence structures. To support practical implementation, two distinct resampling methods are introduced and rigorously justified, offering broadly applicable

procedures for high-dimensional spectral inference.

Complementing this work, the dissertation also develops new probabilistic tools for analyzing quadratic forms in independent random variables with heavy-tailed distributions. Classical probability and moment inequalities for quadratic forms typically assume sub-exponential or exponential tails, which exclude many realistic data-generating mechanisms. To address this gap, we derive a sharp Nagaev-type tail probability inequality and a Rosenthal-type moment inequality that apply under only polynomial moment assumptions. These results substantially broaden the theoretical foundations available for studying second-order statistics in the presence of fat tails.

Taken together, the contributions of this dissertation advance both the methodology and theory of modern statistical inference. On the one hand, they establish a distributional theory and resampling framework for spectral analysis in high-dimensional time series. On the other, they provide new probability and moment inequalities for quadratic forms under weak moment conditions. The combined results offer versatile tools for analyzing high-dimensional and heavy-tailed data, thereby addressing fundamental challenges in contemporary statistics.

# Chapter 1

## Spectral Inference for High Dimensional Time Series

**Abstract:** Spectral analysis plays a fundamental role in the study of time series. While there is a well-developed asymptotic theory for spectral density estimate in low-dimensional cases, a corresponding distributional theory for high-dimensional time series is still lacking. This paper aims to fill this gap by introducing a comprehensive inference theory for the spectral density estimate of high-dimensional time series with possibly nonlinear generating systems and non-Gaussian distributions. Our result is built across different dimensions and frequencies and can serve as a versatile tool for addressing various time series inference challenges. Additionally, we present two distinct resampling methods aimed at practical implementation of high-dimensional spectral inference, each accompanied by a theoretical justification of its validity.

**Key words:** Asymptotic distribution, high dimensional time series, resampling, simultaneous inference, spectral density.

### 1.1 Introduction

The information era has witnessed an explosion in the collection of high-dimensional time series data across a wide range of areas. Spectral density, as an adequate description of serial auto- and cross-correlation in the frequency domain, plays an essential role in

helping understanding the dynamics in serially correlated data without necessarily needing to develop complex parametric models in high dimensions. It has been widely applied in measuring dynamic co-movement of multiple economic time series [44], detecting functional brain network connectivity in neuroscience [41], [7], and investigating multichannel event-related potentials in cognitive science [29]. Besides, it is a powerful tool for waveform analysis in geophysics [8], wave analysis in oceanography [47], [39], signal analysis in weather radar [101], flow analysis in transportation engineering [92], climate oscillation analysis [38] and studies in many other disciplines. Accurate estimation of the spectral density matrix is also a crucial problem in methodological studies. For example, spectral and cross-spectral density estimates can be used in linear prediction of vector-valued stationary processes [78], parameter estimation of generalized least squares type for dynamic models [43], [76], instrumental variable estimation of general linear systems [77], and estimation of generalized dynamic factor models [35], [36]. Also, it permits optimal linear interpolation of multiple missing values in multivariate time series [63].

To fix the idea, for  $p \in \mathbb{N}^+$  and stationary process  $X_t \in \mathbb{R}^p$ , let

$$\Gamma(k) = (\Gamma_{ij}(k))_{i,j=1}^p = \text{Cov}(X_0, X_k)$$

be the autocovariance matrix function. For all  $1 \leq i, j \leq p$ , assume  $\sum_{k \in \mathbb{Z}} |\Gamma_{ij}(k)| < \infty$ . The spectral density matrix function is defined by

$$F(\theta) = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} \Gamma(k) e^{-\iota k \theta}, \text{ where } \iota = \sqrt{-1}.$$

Properties of spectral estimates have been extensively explored for one dimensional stationary processes in many classical time series textbooks; see, for example [2], [71], [81], [15], [14] among others. See also [87] and [59] for more references. Without loss of generality, we assume  $\mathbb{E}X_t = 0$  throughout the paper. It is well known that the

matrix-valued periodogram

$$\widehat{G}_T(\theta) = \frac{1}{2\pi T} \left( \sum_{t=1}^T X_t e^{-it\theta} \right) \left( \sum_{t=1}^T X_t^\top e^{it\theta} \right)$$

is an asymptotically unbiased but inconsistent estimator of  $F(\theta)$ . Consistent estimators of the spectral density function can be obtained by tapering the autocovariance estimator or equivalently by smoothing the periodogram. One natural choice is Welch's [102] non-overlapped segment averaging spectral density estimator given by

$$\widehat{F}(\theta) = \frac{1}{2\pi T} \sum_{r=1}^{T/B} \left( \sum_{t=(r-1)B+1}^{rB} X_t e^{-it\theta} \right) \left( \sum_{t=(r-1)B+1}^{rB} X_t^\top e^{it\theta} \right) \quad (1.1.1)$$

where  $B = o(T)$  is the window size and for notational convenience it is assumed that  $T$  is divisible by  $B$ ; see [44], [71], [66] for more lag-window type variants using different smoothing functions.

In this work, we shall consider a general framework of high-dimensional stationary processes of a causal form

$$X_t = G(\mathcal{F}_t) = (X_{t1}, \dots, X_{tp})^\top, \quad (1.1.2)$$

where  $\mathcal{F}_t = (\varepsilon_t, \varepsilon_{t-1}, \dots)$ ,  $\varepsilon_t$ ,  $t \in \mathbb{Z}$ , are i.i.d. random elements,  $G(\cdot) = (g_1(\cdot), \dots, g_p(\cdot))^\top$  is an  $\mathbb{R}^p$ -valued measurable function such that  $X_t = G(\mathcal{F}_t)$  is a well-defined random vector. In the univariate case where  $p = 1$ , framework (1.1.2) provides a natural paradigm for both linear and nonlinear time series models and represents a huge class of stationary processes which appear frequently in practice; see [103], [79], [72], [97], [106] among many others. Estimation and inference theory on spectral density has been well developed by [87] and [59] in the one dimensional case and [109] in the multivariate case. By allowing the data-generating function to be  $\mathbb{R}^p$ -valued where  $p$  may diverge to infinity,

it extends a large number of existing low dimensional stationary processes into their high-dimensional counterparts in a natural way; see [20], [108] and [113] for examples. In the high-dimensional case, [95], [34] and [114] concerned the convergence of spectral density matrix estimates under different moment and dependence assumptions on the underlying process. [83] established the convergence in distribution of the maximum sample coherence for Gaussian time series to the Gumbel distribution under the high-dimensional regime when  $B = O(T^\rho)$  for some  $0 < \rho < 1$  and  $p/B \rightarrow c$  for some constant  $c \in (0, \infty)$ . It is still an open problem on how to perform simultaneous inference of spectral density matrix for high-dimensional non-Gaussian time series. The main goal of this paper is to develop a comprehensive distributional theory for the high-dimensional spectral density estimate. In particular, we shall establish the asymptotic normality for point-wise inference and a Gaussian approximation result for simultaneous inference. We take into account all of the following features that may arise in real-world applications involving time series data: (i) possibly nonlinear generating system, (ii) temporal dependence and cross-sectional dependence, (iii) high-dimensional data and (iv) non-Gaussian distribution. As another important contribution, we introduce two practical implementation approaches for simultaneous inference of high-dimensional spectral density matrix: one is Gaussian multiplier resampling method and another is subsampling. There has been an extensive study on resampling methods for low dimensional time series including the widely used moving block bootstrap [54], [58] and subsampling [68], [42]. The validity of these methods in high dimensions is still unknown. In the recent decade, Gaussian multiplier bootstrap was introduced in [22] for high-dimensional mean vector inference. As an extension, [115] proposed a blockwise multiplier bootstrap. We shall provide rigorous theoretical justification on resampling approaches to infer spectral density for high-dimensional time series.

The paper is organized as follows. In Section 1.2, we introduce the functional dependence measures for high-dimensional time series and establish the convergence

and point-wise asymptotic normality of spectral estimates. In Section 1.3, we study simultaneous inference of spectral density by establishing a Gaussian approximation result for the maximum deviation of the spectral density estimate over dimensions and frequencies, which could be used to tackle a variety of time series inference problems. We also introduce resampling methods to perform statistical inference in practice and provide theoretical justification on the validity. In Section 1.4, we discuss potential extensions of our results. Section 1.6 presents a simulation study and real data analysis. In Section 1.5, we discuss in more detail the generalization of the results in Section 1.4.1. The proofs of the main results are provided in Section 1.7 and part of the proofs are relegated to the Appendix.

We first introduce some notation. For a random variable  $X$  and  $q \geq 1$ , we define  $\|X\|_q = (\mathbb{E}|X|^q)^{1/q}$ . Denote the operator  $\mathbb{E}_0$  with  $\mathbb{E}_0(X) := X - \mathbb{E}X$ . Define the projection operators  $\mathcal{P}_t \cdot = \mathbb{E}(\cdot | \mathcal{F}_t) - \mathbb{E}(\cdot | \mathcal{F}_{t-1})$  and  $\mathcal{P}^t \cdot = \mathbb{E}(\cdot | \mathcal{F}^t) - \mathbb{E}(\cdot | \mathcal{F}^{t+1})$ , where  $\mathcal{F}_t = (\varepsilon_t, \varepsilon_{t-1}, \dots)$  and  $\mathcal{F}^t = (\varepsilon_t, \varepsilon_{t+1}, \dots)$ . For a complex number  $a \in \mathbb{C}$ , denote the real and imaginary part of  $a$  as  $\text{Re}(a)$  and  $\text{Im}(a)$  respectively. For a matrix  $A = (a_{ij})_{i,j} \in \mathbb{C}^{m \times n}$ , define the element-wise  $\ell_\infty$  norm  $|A|_\infty = \max_{i,j} |a_{ij}|$ . Denote the transpose of  $A$  as  $A^\top$  and the conjugate transpose of  $A$  as  $A^H$ . For two real numbers, set  $x \vee y = \max(x, y)$  and  $x \wedge y = \min(x, y)$ . We use  $C, C_1, C_2, \dots$  to denote positive constants whose values may differ from place to place. A constant with a symbolic subscript is used to emphasize the dependence of the value on the subscript. Throughout the paper, we use  $r, s, t$  to denote time indexes and use  $i, j, h, l$  to denote dimension indexes.

## 1.2 Consistency and Asymptotic Normality of Spectral Estimate

To develop an asymptotic theory for spectral estimates, we need to introduce appropriate dependence measures. Assume that  $\max_{1 \leq j \leq p} \|X_{tj}\|_q < \infty$  for some  $q \geq 2$ . Following the idea in [106], for  $t \geq 0$  and  $1 \leq j \leq p$ , we define functional dependence

measure

$$\delta_{t,q,j} = \|g_j(\varepsilon_t, \varepsilon_{t-1}, \dots) - g_j(\varepsilon_t, \dots, \varepsilon_1, \varepsilon_0^*, \varepsilon_{-1}, \dots)\|_q, \quad (1.2.1)$$

where  $\varepsilon_0^*$  is an i.i.d. copy of  $\varepsilon_0$ . We can observe that  $\delta_{t,q,j}$  measures the distance between the original process  $X_{tj} = g_j(\varepsilon_t, \varepsilon_{t-1}, \dots)$  and the decoupled process  $g_j(\varepsilon_t, \dots, \varepsilon_1, \varepsilon_0^*, \varepsilon_{-1}, \dots)$  with  $\varepsilon_0$  replaced by  $\varepsilon_0^*$  and other innovations kept the same. Thus, it quantifies the effect of the innovation vector  $\varepsilon_0$  on the process  $(X_{tj})$ , which can be interpreted as a possibly nonlinear impulse response function. In this sense, we call  $\delta_{t,q,j}$  the dependence measure for the  $j$ -th component process  $(X_{\cdot j})$  at lag  $t$  and moment order  $q$ . By convention,  $\delta_{t,q,j} = 0$  for  $t < 0$ . For high-dimensional time series, it is challenging to depict the dependence structure, because both the temporal and the cross-sectional dependence need to be concerned. A main advantage of the representation (1.2.1) is that it lets us define physically meaningful and easily workable dependence measures, even for high-dimensional cases. Apparently,  $\delta_{t,q,j}$  measures the temporal dependence at lag  $t$ . As each component  $X_{tj}$  is dependent on the  $p$ -variate vectors  $X_{t-1}, X_{t-2}, \dots$ ,  $\delta_{t,q,j}$  incorporates the cross-sectional dependence as well.

Equipped with the dependence measures (1.2.1), we define the  $q$ -th dependence adjusted moment: assume that there exists some constant  $\rho \in (0, 1)$  such that

$$\|X_{\cdot j}\|_q = \sup_{m \geq 0} \rho^{-m} \Delta_{m,q,j} < \infty, \text{ where } \Delta_{m,q,j} = \sum_{t=m}^{\infty} \delta_{t,q,j}. \quad (1.2.2)$$

Here a geometric moment contracting (GMC) condition is imposed on the cumulative (tail) dependence measure  $\Delta_{m,q,j} = \sum_{t=m}^{\infty} \delta_{t,q,j}$  by noting that  $\Delta_{m,q,j} \leq \|X_{\cdot j}\|_q \cdot \rho^m$  for all  $m \geq 0$ . We can interpret the quantity  $\|X_{\cdot j}\|_q$  as the  $q$ -th moment of the component process  $(X_{\cdot j})$  by taking dependence into account. Elementary calculations show that, if  $X_{tj}$ ,  $t \in \mathbb{Z}$ , are i.i.d., then  $\|X_{tj}\|_q \leq \|X_{\cdot j}\|_q \leq 2\|X_{tj}\|_q$ , suggesting that the dependence adjusted moment is equivalent to the classical  $\mathcal{L}^q$  moment for the independent case. For

high-dimensional time series, it could happen that the quantity  $\|X_{\cdot j}\|_q$  may be expressed as a function of  $p$  and may increase with  $p$ . In the following examples, we work out a bound for the dependence measure of high-dimensional time series. This is a key step in applying our theorems.

**Example 1** (High Dimensional Linear Processes). *Let  $\varepsilon_t = (\varepsilon_{t1}, \dots, \varepsilon_{tp})$ ,  $t \in \mathbb{Z}$ , be i.i.d innovation vectors, where  $\varepsilon_{tj}$  are i.i.d. random variables with mean 0 and  $\mu_q = \|\varepsilon_{tj}\|_q < \infty$  for some  $q \geq 2$ . Let  $A_k$ ,  $k \geq 0$ , be  $p \times p$  real coefficient matrices such that  $\sum_{k=0}^{\infty} \text{tr}(A_k A_k^\top) < \infty$ . By Kolmogorov's three-series theorem, the  $p$ -dimensional linear process*

$$X_t = \sum_{k=0}^{\infty} A_k \varepsilon_{t-k} \quad (1.2.3)$$

*is well-defined. Many researchers have worked on this model recently; see for example, [11], [12], [52], [57], [21] among others. It includes the vector AR processes [90], [94], [98], [33] and MA processes [73], [61], [15]. Denote the  $(j, j')$ -th entry of  $A_t$  by  $A_{t,jj'}$ . We assume that there exist some  $0 < \rho < 1$  and  $0 < K_p < \infty$  such that  $\max_{1 \leq j \leq p} (\sum_{j'=1}^p A_{t,jj'}^2)^{1/2} \leq K_p \rho^t$  for all  $t \geq 0$ . By Theorem 2.1 of [74], we can obtain*

$$\delta_{t,q,j} = \left\| \sum_{j'=1}^p A_{t,jj'} (\varepsilon_{0j'} - \varepsilon_{0j'}^*) \right\|_q \leq \sqrt{q} \mu_q K_p \rho^t.$$

**Example 2** (High dimensional autoregressive conditional heteroskedasticity models). *We consider the vector ARCH process  $X_t = (X_{t1}, \dots, X_{tp})^\top$  with*

$$X_{tj} = \varepsilon_{tj} (a_j + X_{t-1}^\top A_j X_{t-1})^{1/2}, \quad j = 1, \dots, p, \quad (1.2.4)$$

*where  $\varepsilon_{tj}$  are i.i.d. random variables with mean 0 and  $\mu_q = \|\varepsilon_{tj}\|_q < \infty$  for some  $q \geq 2$ ,  $a_j$  are positive constants, and  $A_j$  are  $p \times p$  nonnegative-definite matrices with the spectral norms denoted by  $\lambda_j = \|A_j\|_2$ . Denote  $\varepsilon_t = (\varepsilon_{t1}, \dots, \varepsilon_{tp})$  and  $X_t = \mathbf{f}(X_{t-1}, \varepsilon_t)$ . Since for*

any  $X, Y \in \mathbb{R}^p$ ,

$$|(a_j + X^\top A_j X)^{1/2} - (a_j + Y^\top A_j Y)^{1/2}|^2 \leq \lambda_j |X - Y|_2^2, \quad (1.2.5)$$

it follows that

$$|\mathbf{f}(X, \varepsilon_t) - \mathbf{f}(Y, \varepsilon_t)|_2 \leq \left( \sum_{j=1}^p \lambda_j \varepsilon_{tj}^2 \right)^{1/2} |X - Y|_2.$$

Assume that  $\rho := \|(\sum_{j=1}^p \lambda_j \varepsilon_{tj}^2)^{1/2}\|_q < 1$ . Denote

$$K_{p,q} = \|(\sum_{j=1}^p a_j \varepsilon_{tj}^2)^{1/2}\|_q = \|\mathbf{f}(\mathbf{0}, \varepsilon_t)|_2\|_q.$$

By the arguments for Theorem 2 in [107], the recursion (1.2.4) admits a stationary solution  $X_t = \mathbf{g}(\varepsilon_t, \varepsilon_{t-1}, \dots)$ . By (1.2.4) and (1.2.5), we can compute the dependence measure

$$\delta_{t,q,j} \leq \mu_q K_{p,q} \sqrt{\lambda_j} \rho^t / (1 - \rho).$$

Before proceeding, we state the main assumption required in studying the asymptotic properties of the spectral density estimator (1.1.1).

**Assumption 1.** *There exists some constant  $\alpha \geq 0$  such that*

$$\|X_{\cdot j}\|_{\psi_\alpha} := \sup_{q \geq 2} q^{-\alpha} \|X_{\cdot j}\|_q < \infty.$$

Denote  $\|X\|_{\psi_\alpha} = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_{\psi_\alpha}$ .

In Assumption 1, it is assumed that  $X_{tj}$  has finite moment with any order. For  $\alpha \geq 0$ ,  $\|X_{\cdot j}\|_{\psi_\alpha}$  is interpreted as the dependence adjusted exponential type norm for the  $j$ -th component process and  $\|X\|_{\psi_\alpha}$  is interpreted as the uniform dependence adjusted norm by taking the maximum over all the component processes, which accounts for high dimensionality. By this definition, if  $X_{tj}$ ,  $t \in \mathbb{Z}$ , are i.i.d.,  $\|X_{\cdot j}\|_{\psi_\alpha}$  is equivalent to the

classical sub-Gaussian norm when  $\alpha = 1/2$  or sub-exponential norm when  $\alpha = 1$ , due to the equivalence of  $\|X_{\cdot j}\|_q$  and  $\|X_{tj}\|_q$ . The parameter  $\alpha$  measures how fast  $\|X_{\cdot j}\|_q$  increases with  $q$ . With this assumption, we are able to deal with tails which are heavier than Gaussian tails: the larger  $\alpha$  is, the heavier the tail is. It is worth noting that Assumption 1 imposes some additional restrictions on the underlying process, even when the innovation has finite moments of all orders. These restrictions are characterized by the parameter  $\alpha$ , which quantifies the rate at which the  $q$ -th dependence-adjusted norm  $\|X_{\cdot j}\|_q$  increases with  $q$ . Specifically, under the setting of Example 1, it requires  $\mu_q \leq q^{\alpha-1/2}(1-\rho)K_p^{-1}\|X_{\cdot j}\|_{\psi_\alpha}$  for all  $1 \leq j \leq p$ . In the case of Example 2, the condition becomes  $\mu_q K_{p,q} \leq q^\alpha(1-\rho)^2\lambda_j^{-1/2}\|X_{\cdot j}\|_{\psi_\alpha}$  for all  $1 \leq j \leq p$ .

For notational convenience, we assume  $B$  is an even integer and  $B \geq 4$ . Define the set of discrete Fourier frequencies in  $[0, \pi]$

$$\Theta = \{2k\pi/B : k = 0, 1, 2, \dots, (B/2)\}.$$

Theorem 1 below provides a uniform non-asymptotic tail probability bound for

$$\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty,$$

the maximum deviation of the spectral estimate from its expectation across different dimensions and frequencies. Since the process can be nonlinear, non-Gaussian and high-dimensional, it can be quite involved to derive a tail probability inequality for  $\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty$ . In our framework, equipped with functional dependence measures, we can have a closed form of the upper bound depending on the sample size  $T$ , window size  $B$ , dimension  $p$ , parameter  $\alpha$  and dependence adjusted norm  $\|X_{\cdot}\|_{\psi_\alpha}$ .

**Theorem 1.** *Let Assumption 1 be satisfied for some  $\alpha \geq 0$ . Then there exists some*

constant  $C_\alpha$  such that for any  $x > 0$ ,

$$\mathbb{P}\left(\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty \geq x\right) \lesssim Bp^2 \cdot \exp\left[-C_\alpha \left(\frac{\sqrt{T}x}{\sqrt{B}\|X_\cdot\|_{\psi_\alpha}^2}\right)^{\frac{1}{2\alpha+1}}\right]. \quad (1.2.6)$$

**Remark 1.** *It is worth mentioning that the result in Theorem 1 also applies to the uniform deviation for  $\theta$  over the real line. In fact, for a trigonometric polynomial of order  $B$ , we can bound the uniform deviation by taking the maximum over a fine grid, i.e., by Theorem 7.28 in Ch. X of [117], for  $B' = 4B$  and  $\theta'_k = 2\pi k/B'$ ,*

$$\max_{\theta \in \mathbb{R}} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty \leq 2 \max_{0 \leq k \leq B'} |\widehat{F}(\theta'_k) - \mathbb{E}\widehat{F}(\theta'_k)|_\infty.$$

**Remark 2.** *A closely related result to our Theorem 1 is Theorem 3.1 in [34], which presented a tail probability inequality for smoothed periodogram under finite sub-exponential moment. Our requirement and availability are significantly different from theirs. Firstly, we are able to deal with a wide class of exponential type tails characterized by  $\alpha$ . Moreover, based on (1.2.6), we can obtain*

$$\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty = O_{\mathbb{P}}\left(\|X_\cdot\|_{\psi_\alpha}^2 \sqrt{\frac{B}{T}} \log^{2\alpha+1}(Bp)\right). \quad (1.2.7)$$

*If  $\|X_\cdot\|_{\psi_\alpha} \asymp 1$  and  $B \asymp T^c$  for some  $0 < c < 1$ , as a natural requirement of consistency, we can allow the dimension  $p$  to increase exponentially with  $T$ , i.e.,*

$$\log p = o\left(T^{\frac{1-c}{4\alpha+2}}\right),$$

*while [34] allows a narrower range  $p = O(n^C)$  for some  $C > 0$ . Proposition 4.3 of [114] delivered a probability inequality for lag-window spectral estimates of Gaussian processes:*

$$\mathbb{P}\left(\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty \geq x\right) \lesssim Bp^2 \cdot \exp\left[-C \min\left(\frac{Tx^2}{B\Phi_2^4}, \frac{Tx}{B\Phi_2^2}\right)\right], \quad (1.2.8)$$

where  $\Phi_2 = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_2$ . (1.2.8) suggests two types of bounds for the tail probability: a sub-Gaussian type tail  $\exp(-CTx^2/B\Phi_2^4)$  and a sub-exponential type tail  $\exp(-CTx/B\Phi_2^2)$ . Our result concerns the tails heavier than Gaussian tails and the probability bound is thus less sharper than the Gaussian case. In view of (1.2.7), the error rate is non-decreasing with  $\alpha$ ,  $p$  and the dependence adjusted norm  $\|X\|_{\psi_\alpha}$ . In other words, the rate is less sharper when the process has a heavier tail, higher dimension and stronger dependence.

We also note that Theorem 4.1 in [114] established a concentration bound on spectral density estimators for high-dimensional and locally stationary process, assuming finite polynomial moment. Specifically, we revise the component-wise dependence measure in (1.2.1) and define the  $L_\infty$  dependence measure as

$$w_{t,q} = \| |G(\varepsilon_t, \varepsilon_{t-1}, \dots) - G(\varepsilon_t, \dots, \varepsilon_1, \varepsilon_0^*, \varepsilon_{-1}, \dots)|_\infty \|_q. \quad (1.2.9)$$

We further assume that  $\| |X|_\infty \|_q = \sup_{m \geq 0} \rho^{-m} \sum_{t=m}^\infty w_{t,q} < \infty$  for some  $0 < \rho < 1$ . Under the stationarity assumption and with  $q > 4$ , Theorem 6.1 in [114] yields

$$\begin{aligned} & \mathbb{P}\left(\max_{\theta \in \Theta} |\hat{F}(\theta) - \mathbb{E}\hat{F}(\theta)|_\infty \geq x\right) \\ & \lesssim \frac{T^{1-q/2} B^{q/2} (1 \vee \log p)^{5q/4} \| |X|_\infty \|_q^q}{x^{q/2}} + Bp^2 \cdot \exp\left(-\frac{CTx^2}{B\Phi_4^4}\right). \end{aligned} \quad (1.2.10)$$

This result, a Nagaev-type inequality, provides two types of bounds on the tail probability: a polynomial tail bound and a sub-Gaussian type tail bound. By comparison, our result involves only one term when assuming a general exponential-type tail quantified by  $\|X\|_{\psi_\alpha}$ . In Section 1.4.1, we discuss an extension to the case with polynomial decayed dependence measures, as explored in [114].

**Remark 3.** An alternative estimator of  $F(\theta)$ , which is applicable when the mean  $\mathbb{E}(X_t)$  is

unknown, is the sample-mean-adjusted version given by

$$\tilde{F}(\theta) = \frac{1}{2\pi T} \sum_{r=1}^{T/B} \left( \sum_{t=(r-1)B+1}^{rB} (X_t - \bar{X}) e^{-it\theta} \right) \times \left( \sum_{t=(r-1)B+1}^{rB} (X_t - \bar{X})^\top e^{it\theta} \right), \quad (1.2.11)$$

where  $\bar{X} = T^{-1} \sum_{t=1}^T X_t$ .

Obviously  $\tilde{F}(\theta) \equiv \hat{F}(\theta)$  for  $\theta \in \Theta$  but  $\theta \neq 0$ , since  $\sum_{t=(r-1)B+1}^{rB} e^{-it\theta} = 0$ . If  $\theta = 0$ , we can obtain  $|\tilde{F}(\theta) - \hat{F}(\theta)|_\infty \lesssim B|\bar{X}|_\infty^2$ , while for  $\theta \notin \Theta$ ,  $|\tilde{F}(\theta) - \hat{F}(\theta)|_\infty \lesssim B|\bar{X}|_\infty |\bar{X}(\theta)|_\infty$ , where  $\bar{X}(\theta) = T^{-1} \sum_{t=1}^T X_t e^{-it\theta}$ . Applying Theorem 3 in [108] to  $\bar{X}$  and  $\bar{X}(\theta)$ , and using the Bonferroni technique, we can conclude that the difference  $|\tilde{F}(\theta) - \hat{F}(\theta)|_\infty$  is negligible and Theorem 1 also holds for  $\tilde{F}(\theta)$ .

In Proposition 2, we concern the bias of the spectral estimate.

**Proposition 2.** Denote  $\Phi_2 = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_2$ . Then we have

$$\max_{\theta \in \Theta} |\mathbb{E} \hat{F}(\theta) - F(\theta)|_\infty \lesssim B^{-1} \Phi_2^2. \quad (1.2.12)$$

We can select the window size  $B$  by considering a trade-off between the stochastic deviation bound in (1.2.7) and the bias order provided in Proposition 2. For example, if  $\Phi_2$  and  $\|X_{\cdot}\|_{\psi_\alpha}$  are both of a constant order, to minimize the rate  $\sqrt{B/T} \log^{2\alpha+1}(Bp) + B^{-1}$ , it is suggested that

$$B \asymp \left( \frac{T}{\log^{4\alpha+2}(p \vee T)} \right)^{1/3}.$$

The order of the best  $B$  is analytically expressed in terms of the sample size  $T$ , the dimension  $p$  and the parameter  $\alpha$ .

To perform statistical inference of spectral density, one shall establish the distributional theory of the spectral estimate. Theorem 3 below focuses on the joint convergence of two arbitrary entries in  $\hat{F}(\theta)$ .

**Theorem 3.** Denote  $\Phi_4 = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_4$ . Assume  $B/T \rightarrow 0$  and  $\Phi_4^8(\log B)/B \rightarrow 0$ . Take any  $\theta, \theta' \in \Theta$  and any  $1 \leq i, j, h, l \leq p$ . Then there exists some bivariate (complex) normal random variable  $\boldsymbol{\xi} = (\xi_{ij}(\theta), \xi_{hl}(\theta'))^\top$  such that

$$\sqrt{\frac{T}{B}} \begin{pmatrix} \widehat{F}_{ij}(\theta) - \mathbb{E}\widehat{F}_{ij}(\theta) \\ \widehat{F}_{hl}(\theta') - \mathbb{E}\widehat{F}_{hl}(\theta') \end{pmatrix} - \boldsymbol{\xi} \xrightarrow{d} \mathbf{0}.$$

(i) If  $\theta = \theta' \in (0, \pi)$ ,  $\boldsymbol{\xi} = (\xi_{ij}(\theta), \xi_{hl}(\theta))^\top$  is a bivariate complex normal distribution with mean 0, covariance matrix

$$\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^H) = \begin{pmatrix} F_{ii}(\theta)F_{jj}(\theta) & F_{ih}(\theta)F_{lj}(\theta) \\ F_{hi}(\theta)F_{jl}(\theta) & F_{hh}(\theta)F_{ll}(\theta) \end{pmatrix} \triangleq \Sigma_1$$

and relation matrix (a.k.a. pseudo-covariance matrix)

$$\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^\top) = \begin{pmatrix} F_{ij}^2(\theta) & F_{il}(\theta)F_{hj}(\theta) \\ F_{il}(\theta)F_{hj}(\theta) & F_{hl}^2(\theta) \end{pmatrix} \triangleq \Sigma_2.$$

(ii) If  $\theta = \theta' = 0$  or  $\pi$ ,  $\boldsymbol{\xi}$  is a bivariate normal distribution with mean 0 and covariance matrix

$$E(\boldsymbol{\xi}\boldsymbol{\xi}^\top) = \Sigma_1 + \Sigma_2.$$

(iii) If  $\theta \neq \theta'$ ,  $\xi_{ij}(\theta)$  and  $\xi_{hl}(\theta')$  are independent (complex) normal distributions. In particular, the off-diagonals in matrices  $\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^H)$  and  $\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^\top)$  are all 0.

Theorem 3 covers all possible cases for the asymptotic joint behavior of two (cross-)spectral density estimates  $\widehat{F}_{ij}(\theta)$  and  $\widehat{F}_{hl}(\theta')$ . By (i) and (ii), the covariance matrix (as well as the pseudo-covariance matrix for complex-valued spectral estimates) of the two at the same frequency is fully characterized by the entries of  $F(\theta)$ , which will further be

used in performing simultaneous inference of spectral density in Section 1.3. Result (iii) suggests the asymptotic independence of spectral density estimates at different frequencies. By Proposition 2, with the bias concerned, we can replace  $\mathbb{E}\widehat{F}_{ij}(\theta)$  with  $F_{ij}(\theta)$  in the statement of Theorem 3 under the additional condition  $T\Phi_2^4 = o(B^3)$ ,

By considering the marginal case of the joint convergence, we can study the limiting distribution of  $\widehat{F}_{ij}(\theta)$ , the estimator of each cross spectral density. We represent the limiting distribution of  $\sqrt{T/B}(\widehat{F}_{ij}(\theta) - \mathbb{E}\widehat{F}_{ij}(\theta))$  as  $\xi_{ij}(\theta) = \text{Re}(\xi_{ij}(\theta)) + \iota \text{Im}(\xi_{ij}(\theta))$ . By Theorem 3 (i), for  $\theta \in (0, \pi)$ , we can conclude  $(\text{Re}(\xi_{ij}(\theta)), \text{Im}(\xi_{ij}(\theta)))^\top$  is a bivariate normal distribution with mean 0. Looking into the diagonals of the covariance matrices, we can get the variance of the real part and imaginary part as follows

$$\text{Var}(\text{Re}(\xi_{ij}(\theta))) = \frac{1}{2}F_{ii}(\theta)F_{jj}(\theta) + \frac{1}{2}\text{Re}(F_{ij}^2(\theta)), \quad (1.2.13)$$

$$\text{Var}(\text{Im}(\xi_{ij}(\theta))) = \frac{1}{2}F_{ii}(\theta)F_{jj}(\theta) - \frac{1}{2}\text{Re}(F_{ij}^2(\theta)), \quad (1.2.14)$$

$$\text{Cov}(\text{Re}(\xi_{ij}(\theta)), \text{Im}(\xi_{ij}(\theta))) = \frac{1}{2}\text{Im}(F_{ij}^2(\theta)), \quad (1.2.15)$$

while Theorem 3 (ii) implies  $\xi_{ij}(\theta) = N(0, F_{ij}^2(\theta) + F_{ii}(\theta)F_{jj}(\theta))$  if  $\theta = 0$  or  $\pi$ . This result agrees with Theorem 3 in [109] for the single cross spectral estimate and agrees with Theorem 2 in [59] for the real-valued spectrum estimate by letting  $i = j$ . The central limit theorem of the spectral estimate is highly non-trivial even for the special case of one dimensional time series. Many previous results require restrictive conditions on the underlying process such as linear processes [2], strong mixing conditions [80] and Gaussian processes [10]. Our framework covers a wide class of possibly nonlinear stationary processes and our conditions involving functional dependence measures can be easily verified. In addition, we impose a mild moment condition to deal with non-Gaussian data.

## 1.3 Simultaneous Inference of Spectral Density

In this section, we shall consider the simultaneous inference of the spectral density matrix for high-dimensional time series. In the one dimensional case, the distribution of  $\max_{\theta \in [0, \pi]} |\hat{F}(\theta) - F(\theta)|$  or the discretized version  $\max_{\theta \in \Theta} |\hat{F}(\theta) - F(\theta)|$  allows one to construct the confidence band of  $F(\theta)$  or conduct a parametric specification test for  $F(\theta)$ . Gumbel convergence of this maximum deviation has been established by earlier work. For example, in the one dimensional case, [104] dealt with linear processes and [85] worked on Gaussian processes. Under the framework of (1.1.2) with  $p = 1$ , [59] derived the asymptotic distribution of the maximum deviation for the spectral density estimate and it was generalized to the multivariate case by [109] concerning the cross-spectral density. [83] considered the asymptotic distribution of the maximum sample coherence for Gaussian time series under the high-dimensional regime when  $B = O(T^\rho)$  for some  $0 < \rho < 1$  and  $p/B \rightarrow c$  for some constant  $c \in (0, \infty)$ . There is little investigation on the simultaneous inference of spectral density for non-Gaussian processes with a possibly ultra-high dimension. In this section, we shall develop an inference theory on high-dimensional spectral density.

### 1.3.1 Gaussian Approximation for High Dimensional Spectral Estimates

Performing simultaneous inference of the high-dimensional spectral density matrix is a challenging task, especially when the dimension  $p$  is comparable to or even greater than the sample size  $T$ . If  $p$  can also diverge to infinity, [70] showed that the central limit theorem will fail for i.i.d. random vectors if  $\sqrt{T} = o(p)$ . Classical probability tools are inadequate for formulating a distribution theory for high-dimensional cases. We shall consider an alternative form: Gaussian approximation for the largest magnitude of the entries of the spectral density matrix. The idea of Gaussian approximation for

high-dimensional random vectors was previously adopted in mean vector inference; see, for example [113], [22] and [116]. Addressing the spectral density matrix still poses a new challenge. Firstly, the spectral density matrix exhibits complex-valued off-diagonals (cross-spectral densities) when  $p > 1$ . Secondly, the spectral estimate manifests as a quadratic form in the time-dependent samples, further complicating the analytical landscape. In this section, we will address both aspects and formulate a new distribution theory to facilitate the inference of the spectral density matrix. In particular, we consider two types of deviation:

$$\begin{aligned}\mathbf{F} &= \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\hat{F}(\theta) - \mathbb{E}\hat{F}(\theta))|_\infty, |\operatorname{Im}(\hat{F}(\theta) - \mathbb{E}\hat{F}(\theta))|_\infty \}; \\ \tilde{\mathbf{F}} &= \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\hat{F}(\theta) - \mathbb{E}\hat{F}(\theta)|_\infty.\end{aligned}$$

The former separates the real and imaginary parts of each element, while the latter concerns the maximum modulus of these entries. While it is quite natural to consider the largest modulus of those complex-valued elements, it is worth mentioning that investigation on the real and imaginary parts separately of cross-spectral density can be of important use in many applications. For example, the imaginary part has gained increasing popularity to identify brain interaction in EEG and MEG connectivity studies [7], [67], [62], [37], [48], etc.

For the convenience of simultaneous inference, we exclude the two end frequencies  $0, \pi$  and consider  $\Theta = \{\theta_k : \theta_k = 2k\pi/B, k = 1, 2, \dots, (B/2 - 1)\}$ . Let  $\mathbf{vec}_0(\hat{F}) \in \mathbb{R}^{(B-2)p^2}$  be the vectorization of  $(\operatorname{Re}(\hat{F}(\theta)), \operatorname{Im}(\hat{F}(\theta)))_{\theta \in \Theta}$ . We aim to establish a distributional theory for  $\mathbf{F}$  by the idea of Gaussian approximation. In particular, denote

$$\boldsymbol{\Sigma} = \lim_{B, T \rightarrow \infty} \frac{T}{B} \operatorname{Cov}(\mathbf{vec}_0(\hat{F})) \text{ and } \mathbf{Z} \sim N(0, \boldsymbol{\Sigma}). \quad (1.3.1)$$

By Theorem 3, we can work out the expression of each entry in  $\boldsymbol{\Sigma}$ . For  $\theta \in (0, \pi)$  and

arbitrary combination of  $i, j, h, l \in \{1, 2, \dots, p\}$ , it can be computed that

$$\begin{aligned}
\lim_{B, T \rightarrow \infty} \frac{T}{B} \text{Cov}(\text{Re}(\widehat{F}_{ij}(\theta)), \text{Re}(\widehat{F}_{hl}(\theta))) &= \frac{1}{2} \text{Re}(F_{ih}(\theta)F_{lj}(\theta)) + \frac{1}{2} \text{Re}(F_{il}(\theta)F_{hj}(\theta)), \\
\lim_{B, T \rightarrow \infty} \frac{T}{B} \text{Cov}(\text{Im}(\widehat{F}_{ij}(\theta)), \text{Im}(\widehat{F}_{hl}(\theta))) &= \frac{1}{2} \text{Re}(F_{ih}(\theta)F_{lj}(\theta)) - \frac{1}{2} \text{Re}(F_{il}(\theta)F_{hj}(\theta)), \\
\lim_{B, T \rightarrow \infty} \frac{T}{B} \text{Cov}(\text{Im}(\widehat{F}_{ij}(\theta)), \text{Re}(\widehat{F}_{hl}(\theta))) &= \frac{1}{2} \text{Im}(F_{il}(\theta)F_{hj}(\theta)) + \frac{1}{2} \text{Im}(F_{ih}(\theta)F_{lj}(\theta)), \\
\lim_{B, T \rightarrow \infty} \frac{T}{B} \text{Cov}(\text{Re}(\widehat{F}_{ij}(\theta)), \text{Im}(\widehat{F}_{hl}(\theta))) &= \frac{1}{2} \text{Im}(F_{il}(\theta)F_{hj}(\theta)) - \frac{1}{2} \text{Im}(F_{ih}(\theta)F_{lj}(\theta)).
\end{aligned}$$

All the other entries of  $\Sigma$  are set to be 0 by the asymptotic independence of two spectral density estimates at different frequencies.

Before proceeding, we state the main assumptions required in studying the distribution of  $\mathbf{F}$ .

**Assumption 2.** *There exists some constant  $c_1 > 0$  such that*

$$\min_{\theta \in \Theta} F_{ii}(\theta) \geq c_1 \text{ for all } 1 \leq i \leq p.$$

**Assumption 3.** *There exists some constant  $c_2 > 0$  such that*

$$\min_{\theta \in \Theta} \{F_{ii}(\theta)F_{jj}(\theta) - |\text{Re}(F_{ij}^2(\theta))|\} \geq c_2 \text{ for all } 1 \leq i \neq j \leq p.$$

In view of (1.2.13) and (1.2.14), Assumption 2 and Assumption 3 ensure all the variances of non-constantly-zero entries in  $\mathbf{Z}$  are bounded away from 0, where the former requires non-vanishing spectrum for each component process and the latter accounts for a lower bound for the asymptotic variances of  $\text{Re}(\widehat{F}_{ij}(\theta))$  and  $\text{Im}(\widehat{F}_{ij}(\theta))$  when  $i \neq j$ .

In Theorem 4, we shall present the main result for the approximate distribution on spectral estimate. That is, the distribution of  $\mathbf{F}$  or  $\tilde{\mathbf{F}}$  can be approximated by its Gaussian analogue.

**Theorem 4.** *Let Assumption 1, Assumption 2 and Assumption 3 be satisfied.*

*For  $\kappa_1 = \max\{7, 2\alpha + 6\}$  and  $\kappa_2 = 4\alpha + 4$ , let*

$$\mathcal{E}_{T,B,p} = \left( \frac{B \log^{\kappa_1}(Tp)}{T} \right)^{\frac{1}{6}} + \left( \frac{\log^{\kappa_2}(Bp)}{B} \right)^{\frac{1}{2}} + \left( \frac{\log^5(Bp)}{B} \right)^{\frac{1}{6}}. \quad (1.3.2)$$

*Assume that  $\|X.\|_{\psi_\alpha}^2 \mathcal{E}_{T,B,p} \rightarrow 0$ . Then it holds that*

$$\begin{aligned} d_{T,B,p} &:= \sup_{x \in \mathbb{R}} |\mathbb{P}(\mathbf{F} \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \lesssim \|X.\|_{\psi_\alpha}^2 \mathcal{E}_{T,B,p}, \\ \tilde{d}_{T,B,p} &:= \sup_{x \in \mathbb{R}} |\mathbb{P}(\tilde{\mathbf{F}} \leq x) - \mathbb{P}(|\tilde{\mathbf{Z}}|_\infty \leq x)| \lesssim \|X.\|_{\psi_\alpha}^2 \mathcal{E}_{T,B,p}, \end{aligned}$$

*where*

$$\mathbf{Z} = (Z_1, Z_2, \dots, Z_{(B-2)p^2})^\top$$

*is defined in (1.3.1) and*

$$\tilde{\mathbf{Z}} = (\sqrt{Z_{2j-1}^2 + Z_{2j}^2})_{j=1, \dots, (B/2-1)p^2}.$$

In the low dimensional case when  $p$  is a fixed number, the conditions of Theorem 4 are simplified to the very mild requirement  $B = o(T/\log^{\kappa_1}(T))$ . We can also investigate how large the dimension  $p$  can be allowed for the validity of the Gaussian approximation result in high dimensions. For example, if  $\|X.\|_{\psi_\alpha} \asymp 1$ , we shall require

$$\log p = o\left(\min\left\{\left(\frac{T}{B}\right)^{1/\kappa_1}, B^{1/\kappa_2}, B^{1/5}\right\}\right)$$

If we choose  $B \asymp T^c$  for some  $0 < c < 1$ ,  $p$  can be allowed to increase exponentially with the sample size  $T$ , i.e.,  $\log p = o(T^C)$  where  $C = \min\{(1-c)/\kappa_1, c/\kappa_2, c/5\}$ .

We now consider the version centered by  $F(\theta)$ . Define

$$\begin{aligned}\mathbf{F}^\star &= \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\widehat{F}(\theta) - F(\theta))|_\infty, |\operatorname{Im}(\widehat{F}(\theta) - F(\theta))|_\infty \}; \\ \tilde{\mathbf{F}}^\star &= \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\widehat{F}(\theta) - F(\theta)|_\infty.\end{aligned}\tag{1.3.3}$$

By accounting for the bias of  $\widehat{F}(\theta)$ , the Gaussian approximation result also holds for  $\mathbf{F}^\star$  and  $\tilde{\mathbf{F}}^\star$ ; see Corollary 5 below.

**Corollary 5.** *Let the assumptions of Theorem 4 be satisfied. Further assume that*

$$\frac{T \log(Bp) \|X.\|_{\psi_\alpha}^4}{B^3} \rightarrow 0.\tag{1.3.4}$$

*Then we have*

$$\begin{aligned}d_{T,B,p}^\star &:= \sup_{x \in \mathbb{R}} |\mathbb{P}(\mathbf{F}^\star \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \rightarrow 0, \\ \tilde{d}_{T,B,p}^\star &:= \sup_{x \in \mathbb{R}} |\mathbb{P}(\tilde{\mathbf{F}}^\star \leq x) - \mathbb{P}(|\tilde{\mathbf{Z}}|_\infty \leq x)| \rightarrow 0.\end{aligned}$$

Our results in Theorem 4 and Corollary 5 concern the maximum deviation over all the entries in  $\widehat{F}(\theta)$  and over all the frequencies in  $\Theta$ . In various inference scenarios, we only need to select a subset of these elements and the asymptotic covariance matrix  $\Sigma$  should also have an adaptively reduced size to account for the variances and pairwise covariances of elements in the chosen subset. In Section 1.3.2 we provide some examples as such.

### 1.3.2 Applications

Our result can be applied to address a variety of time series inference problems. In hypothesis testing, we can compute the test statistics defined by (1.3.3), where  $\widehat{F}(\theta)$  is calculated from the sample data, and the plugged-in  $F(\theta)$  is from the null hypothesis.

Corollary 5 provides the theoretical guarantee of validity.

An immediate application is to test the spectrum for any component process marginally, i.e.,  $H : F_{ii}(\theta) = f_i(\theta)$ , for  $1 \leq i \leq p$ , where  $f_i(\theta)$  is the hypothesized spectral density for the component process  $(X_{ti})_t$ . In this case,  $\hat{F}_{ii}(\theta)$  is real-valued and only Assumption 2 is required to ensure a non-vanishing variance of  $\hat{F}_{ii}(\theta)$ . The two test statistics are equal

$$\mathbf{F}^* = \tilde{\mathbf{F}}^* = \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max_{1 \leq i \leq p} |\hat{F}_{ii}(\theta) - f_i(\theta)|$$

and  $\Sigma$  consists of covariances of  $\hat{F}_{ii}(\theta)$  and  $\hat{F}_{jj}(\theta)$  for  $1 \leq i, j \leq p$  and  $\theta \in \Theta$ .

Moreover, we can provide a new solution to high-dimensional white noise test, which is used for serial correlation detection of time series. Classical portmanteau-type tests including [13] and [60] are popular choices in the one dimensional case, and are well generalized to the multivariate case (e.g., [25], [50]). high-dimensional white noise tests are developed for linear processes by [55] concerning the sum of squared singular values of the first few lagged sample autocovariance matrices and for  $\beta$ -mixing processes by [18] using the maximum absolute auto-correlations and cross-correlations (see [99] for a similar test statistic formulation). The forementioned tests in the time domain concern the null hypothesis  $H : \Gamma(k) = 0$ , for  $1 \leq k \leq L$ , to test for serial correlation and present the additional problem of selecting the number of lags to be considered, which does not appear in the frequency domain by stating the null hypothesis in terms of the spectral density matrix function, namely,  $H : F(\theta) = F_0$ , where  $F_0$  consists of constant-valued entries free of  $\theta$ . The most common statistics in the one dimensional case are based on standardized cumulative periodogram; see [6], [40], [31], [32], just to name a few. With the given  $F_0$  inserted in our statistic (1.3.3), our result can be used to design a new high-dimensional white noise test for a large class of stationary processes under very mild regularity conditions. To practically select an initial  $F_0$ , we observe that  $F(\theta) = (2\pi)^{-1}\Gamma(0)$  if  $X_t$  is white noise, where  $\Gamma(0) = \text{Cov}(X_t, X_t)$  denotes

the covariance matrix of the stationary process. We can refer to existing methods of consistently estimating the covariance matrix for high-dimensional time series, including the sample covariance matrix studied in [114], the regularized covariance estimator under sparse settings in [20], and the robust covariance estimator for fat-tailed processes proposed in [112].

We can also test for pairwise independence between two component time series. Unlike the tests based on the cross-correlation function in the time domain (e.g., [46], [53], [49]), we test  $H : F_{ij}(\theta) \equiv 0$  for fixed  $i \neq j$ . Suppose Assumption 3 holds for the selected  $i, j$  and let  $p = 2$  in the theory. We shall use the test statistic

$$\mathbf{F}_{ij}^* = \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max\{|\operatorname{Re}(\widehat{F}_{ij}(\theta))|, |\operatorname{Im}(\widehat{F}_{ij}(\theta))|\}$$

or  $\tilde{\mathbf{F}}_{ij}^* = \sqrt{T/B} \max_{\theta \in \Theta} |\widehat{F}_{ij}(\theta)|$ . Our Corollary 5 provides a new criteria for the cross-spectral density testing.

Spectral analysis has been widely used to detect functional network connectivity; see [7], [30], [89], [56] and [75] for applications involving neuroimaging data. Denote the partition of the index set  $\{1, 2, \dots, p\}$  as  $K$  disjoint subsets  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_K$  with  $\cup_{k=1}^K \mathcal{A}_k = \{1, 2, \dots, p\}$ . We can identify disconnected brain regions by testing the joint independence of the  $K$  subset processes  $(X_{ti})_{i \in \mathcal{A}_k}$ ,  $k = 1, \dots, K$ . That is, the problem is to consider  $H : F_{ij}(\theta) \equiv 0$  for any combinations of  $i \in \mathcal{A}_{k_1}$ ,  $j \in \mathcal{A}_{k_2}$  and  $1 \leq k_1 \neq k_2 \leq K$  with the test statistic

$$\tilde{\mathbf{F}}^* = \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max_{1 \leq k_1 \neq k_2 \leq K} \max_{i \in \mathcal{A}_{k_1}, j \in \mathcal{A}_{k_2}} |\widehat{F}(\theta)|.$$

### 1.3.3 Resampling Methods

To perform statistical inference on high-dimensional spectral density in practice, given the level  $\theta_0 \in (0, 1)$ , one needs to work out the cut-off value  $c_{1-\theta_0}$ , the  $(1-\theta_0)$ -quantile

of  $|\mathbf{Z}|_\infty$  or  $|\tilde{\mathbf{Z}}|_\infty$  for  $\mathbf{Z} \sim N(0, \Sigma)$ . Despite the theoretical advancements on the distribution approximation of the spectral estimate, practical implementation remains challenging due to the involvement of unknown covariance matrix  $\Sigma$ . In spectral inference, we leverage the unique feature that the asymptotic covariances of the spectral estimates (1.1.1) can be entirely expressed through the true spectral densities. Hence, a good estimator of  $\hat{F}(\theta)$  with Theorem 1 concerning the convergence of  $\hat{F}(\theta)$  ensures a good plug-in estimator of  $\Sigma$ .

In this context, we carry out the implementation of estimating the distribution of  $\mathbf{F}^*$ . The case of  $\tilde{\mathbf{F}}^*$  can be dealt with in a similar way. By Theorem 3 and looking into their real and imaginary parts pairwise, we have the nice property of expressing the entries in  $\Sigma$  in terms of the entries in  $F(\theta)$ , as justified in Section 1.3.1. We can implement a Gaussian multiplier resampling method for spectral inference as follows:

1. Given the observations  $X_1, \dots, X_t$ , estimate  $F(\theta)$  by  $\hat{F}(\theta)$ .
2. Plug the entries of  $\hat{F}(\theta)$  into the formula of entries of  $\Sigma$  to obtain  $\hat{\Sigma}$ .
3. Obtain  $\hat{c}_{1-\theta_0}$ , the conditional  $(1 - \theta_0)$ -quantile of  $|\hat{\Sigma}^{1/2} \mathbf{Z}_0|_\infty$  given  $(X_t)_{t=1}^T$ , where  $\mathbf{Z}_0$  is a standard normal random vector, independent of  $(X_t)_{t=1}^T$ .

The technique of Gaussian multiplier bootstrap was employed in [22] for high-dimensional mean vector inference in the independent setting. As an extension, blockwise multiplier bootstrap was proposed by [115] for  $M$ -dependent processes. In particular, to apply the blockwise multiplier bootstrap for  $M$ -dependent processes, the method divides the sample into consecutive large and small blocks of size  $N$  and  $M$ , respectively, with  $N \gg M$ . The statistic over these blocks is then computed and further multiplied by i.i.d. Gaussian variables to facilitate statistical inference. However, for the general weakly dependent processes addressed in our paper, selecting appropriate values for  $N$  and  $M$  is challenging and introduces additional uncertainty into the inference. Based on our Gaussian approximation theory for spectral estimates, in which  $\Sigma$  is specified by entries

from the true  $F(\theta)$ , directly substituting  $\hat{F}(\theta)$  to estimate  $\Sigma$  offers a straightforward approach by avoiding the need for additional block-size selection.

Theorem 6 is introduced to concern the validity of our approach.

**Theorem 6.** *Let the conditions in Corollary 5 be satisfied. Further assume*

$$\sqrt{\frac{B}{T}} \log^{2\alpha+3}(Bp) \|X_{\cdot}\|_{\psi_{\alpha}}^4 \rightarrow 0. \quad (1.3.5)$$

Then

$$\sup_{\theta_0 \in (0,1)} |\mathbb{P}(\mathbf{F}^* \geq \hat{c}_{1-\theta_0}) - \theta_0| \rightarrow 0. \quad (1.3.6)$$

If we use  $\tilde{\mathbf{F}}^*$ , it suffices to slightly modify the third step to obtain  $\tilde{c}_{1-\theta_0}$ , the conditional  $(1 - \theta_0)$ -quantile of  $|\hat{\tilde{\mathbf{Z}}}|_{\infty}$ , where

$$\hat{\mathbf{Z}} = (\hat{Z}_1, \dots, \hat{Z}_{(B-2)p^2})^{\top} = \hat{\Sigma}^{1/2} \mathbf{Z}_0,$$

$\mathbf{Z}_0$  is a standard normal random vector, independent of  $(X_t)_{t=1}^T$  and

$$\hat{\tilde{\mathbf{Z}}} = (\sqrt{\hat{Z}_{2j-1}^2 + \hat{Z}_{2j}^2})_{j=1, \dots, (B/2-1)p^2}.$$

Another widely employed and effective resampling technique is subsampling, which constructs the resampling distribution by taking the subsamples of the observations. It has been commonly used in inferring low dimensional time series. Some pioneering work includes [68], [42], [88]; see [69] for a comprehensive overview. Despite its intuitive appeal, there has been relatively limited exploration into the theoretical properties of the subsampling method in high dimensions. Making an additional noteworthy contribution, we provide a thorough theoretical justification for the validity of high-dimensional spectral inference through subsampling. Consequently, we obtain a consistent estimator for the distribution function of  $\mathbf{F}^*$ , enabling the implementation of tests and confidence intervals

based on  $\mathbf{F}^*$ . To the best of our knowledge, our work is the first to provide theoretical assurance for high-dimensional spectral inference utilizing subsampling. It is anticipated that the novel proofs can be adapted to offer theoretical support in other high-dimensional subsampling problems.

Let  $M = o(T)$  be the subsample size and  $B = o(M)$  be the window size. Over each subsample  $(X_\ell, X_{\ell+1}, \dots, X_{\ell+M-1})$ , we apply (1.1.1) to estimate the spectral density matrix, denoted by  $\widehat{F}_\ell(\theta)$ . Recall  $\widehat{F}(\theta)$  is the estimator over the whole sample  $X_1, \dots, X_t$ . We define

$$\mathbf{F}_\ell = \sqrt{\frac{M}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\widehat{F}_\ell(\theta) - \widehat{F}(\theta))|_\infty, |\operatorname{Im}(\widehat{F}_\ell(\theta) - \widehat{F}(\theta))|_\infty \}.$$

For  $\mathcal{L} = T - M + 1$ , define the empirical distribution function

$$\mathbf{Q}(x) = \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} \mathbf{1}\{\mathbf{F}_\ell \leq x\}.$$

The deviation based on the maximum modulus of spectral densities can be dealt with similarly with  $\mathbf{F}_\ell$  replaced by

$$\tilde{\mathbf{F}}_\ell = \sqrt{\frac{M}{B}} \max_{\theta \in \Theta} |\widehat{F}_\ell(\theta) - \widehat{F}(\theta)|_\infty.$$

**Theorem 7.** *Assume that  $M = o(T)$  and  $B = o(M)$ . Let the conditions of Corollary 5 be satisfied, in which  $T$  is replaced by the subsample size  $M$ . Further assume the following conditions are met as  $B, M, T \rightarrow \infty$ :*

$$T^{-1} M \log^{4\alpha+5}(Bp) \|X_\cdot\|_{\psi_\alpha}^8 \rightarrow 0 \tag{1.3.7}$$

and

$$B^{-3}M \log^3(Bp) \|X.\|_{\psi_\alpha}^8 \rightarrow 0. \quad (1.3.8)$$

Then we have

$$\sup_{x \in \mathbb{R}} |\mathbf{Q}(x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \xrightarrow{\mathbb{P}} 0.$$

**Remark 4.** *The subsampling approach is also valid for consistent spectral estimates by other smoothing methods when the asymptotic covariance matrix  $\Sigma$  may not be fully characterized. For example, with a general lag-window estimator, or estimating the entries of  $\widehat{F}(\theta)$  element-wisely using different smoothing functions, the Gaussian approximation result still holds if the asymptotic variances of the real and imaginary parts of each element are non-vanishing, but the variances may rely on the smoothing functions in a subtle way and working out a closed-form of  $\Sigma$  could be highly non-trivial. In such cases, we can choose to implement subsampling to avoid the estimation of  $\Sigma$ .*

## 1.4 Extension

### 1.4.1 Relaxing the Moment and Dependence Conditions

In this section, we shall discuss the generalization of the results in Section 1.2 and Section 1.3 by relaxing the moment and dependence conditions. Specifically, we assume the existence of the  $q$ -th moment for some  $q > 4$ , as considered by [113] and [114]. This assumption is weaker than Assumption 1, which requires moments to exist at all orders. Using the  $L_\infty$  dependence measure defined in (1.2.9) and the associated  $q$ -th dependence-adjusted norm  $|||X.|_\infty||_q$ , we can follow the proof of Theorem 6.1 in [114] to

obtain the tail probability bound for  $q > 4$ :

$$\begin{aligned} & \mathbb{P}\left(\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty \geq x\right) \\ & \lesssim \frac{T^{1-q/2} B^{q/2} (1 \vee \log p)^{5q/4} \|X_\cdot\|_\infty^q}{x^{q/2}} + Bp^2 \cdot \exp\left(-\frac{CTx^2}{B\Phi_4^4}\right). \end{aligned} \quad (1.4.1)$$

Using the concentration inequality in (1.4.1) as a replacement for Theorem 1 and applying Gaussian approximation with finite polynomial moments (cf. Proposition 2.1, case (E.2) in [24]) as a substitute for Lemma 26, we can employ similar arguments to derive the Gaussian approximation result and demonstrate the validity of the two resampling methods under polynomial moment conditions.

While the geometric moment contracting (GMC) condition considered in (1.2.2) has been widely applied in spectral inference (e.g., [87] and [109]), a more relaxed condition is polynomial decay, which assumes

$$\|X_{\cdot j}\|_{q,\alpha} = \sup_{m \geq 0} (m \vee 1)^{-\alpha} \sum_{t=m}^{\infty} \delta_{t,q,j} < \infty$$

for some  $\alpha > 0$ . In Example 1, for the linear process  $X_t = \sum_{k=0}^{\infty} A_k \varepsilon_{t-k}$ , we assume that there exist some  $\beta > 1$  and  $0 < K_p < \infty$  such that  $\max_{1 \leq j \leq p} (\sum_{j'=1}^p A_{t,jj'}^2)^{1/2} \leq K_p (t \vee 1)^{-\beta}$  for any  $t \geq 0$ , where  $A_{t,jj'}$  denotes the  $(j, j')$ -th entry of  $A_t$ . This allows us to bound  $\delta_{t,q,j} = \|\sum_{j'=1}^p A_{t,jj'}(\varepsilon_{0j'} - \varepsilon_{0j'}^*)\|_q \leq C_q \mu_q K_p (t \vee 1)^{-\beta}$ . Thus, the upper bound of the dependence measure exhibits polynomial decay with  $\alpha = \beta - 1$ . Essentially, replacing the original tail cumulative dependence measure of the order  $\rho^m$  with the order  $m^{-\alpha}$  allows all proofs to extend to the case where the dependence measure decays polynomially. Notably, an additional assumption,  $\max_{1 \leq j \leq p} \sum_{t=0}^{\infty} t^2 \delta_{t,4,j} < \infty$ , is required to ensure the summability of fourth-order joint cumulants, while this condition is automatically satisfied under the geometric decay as stated in Lemma 17.

### 1.4.2 Lag-window Estimator of Spectral Density

Next, we discuss the possibility of extending the theory to the general lag-window estimator, defined as

$$\check{F}(\theta) = \frac{1}{2\pi T} \sum_{k=-B}^B \mathcal{K}(k/B) \hat{\Gamma}(k) e^{-\imath k\theta}, \quad (1.4.2)$$

where  $\mathcal{K}(\cdot)$  is a continuous, even and bounded kernel function on  $[-1, 1]$  satisfying  $\mathcal{K}(0) = 1$ ,  $B = o(T)$  is the window size, and  $\hat{\Gamma}(k)$  are the sample autocovariance matrices with  $\hat{\Gamma}(k) = T^{-1} \sum_{t=k+1}^T X_{t-k} X_t^\top$  for  $k \geq 0$  and  $\hat{\Gamma}(k) = \hat{\Gamma}(-k)^\top$  for  $k < 0$ . By an inspection of the proof for Theorem 1, applying equation (41) of [111] to derive the key step (1.7.4), the same concentration bound in Theorem 1 applies to the lag-window estimator  $\check{F}(\theta)$ .

For the distributional theory, we observe that the separation needs to be larger than the order  $B^{-1}$  to ensure asymptotic uncorrelation of spectral estimates at neighboring frequencies. This behavior for the lag-window estimator is also supported by Section 6.2.4 in [71], Lemma A.5 in [59] and Lemma A.6 in [109]. Therefore, the covariance matrix  $\Sigma$  contains many non-zero entries, reflecting the strong correlations at close frequencies, making it difficult to compute. After verification, our Gaussian approximation results still hold with  $\Sigma = \lim_{B, T \rightarrow \infty} \frac{T}{B} \text{Cov}(\text{vec}_0(\hat{F}))$  and  $\mathbf{Z} \sim N(0, \Sigma)$ ; however, obtaining a closed form for all entries of  $\Sigma$  is highly non-trivial. Consequently, while the Gaussian multiplier resampling method, which relies on explicit computation of  $\Sigma$ , is not applicable, subsampling remains valid.

## 1.5 Generalization

In this section, we shall show in more detail the generalization of the results in Section 1.2 and Section 1.3 as discussed in Section 1.4.1.

### 1.5.1 Consistency and Asymptotic Normality of Spectral Estimate

To develop the main theory, we need to revise the dependence measure (1.2.1) and  $q$ -th dependence adjusted moment (1.2.2), along with  $L_\infty$  dependence measure defined in (1.2.9) and the associated  $q$ -th dependence-adjusted norm  $\|X\|_{q,\alpha}$ . we define the component and  $L_\infty$  dependence adjusted norms

$$\|X_{\cdot j}\|_{q,\alpha} = \sup_{m \geq 0} (1 \vee m^\alpha) \Delta_{m,q,j}, \quad \alpha \geq 0, \quad \text{where } \Delta_{m,q,j} = \sum_{t=m}^{\infty} \delta_{t,q,j}, \quad (1.5.1)$$

$$\|X_{\cdot j}\|_{q,\alpha} = \sup_{m \geq 0} (1 \vee m^\alpha) \Omega_{m,q}, \quad \alpha \geq 0, \quad \text{where } \Omega_{m,q} = \sum_{t=m}^{\infty} \omega_{t,q}. \quad (1.5.2)$$

We use  $\alpha$  to depict the decay rate of the cumulative (tail) dependence measure  $\Delta_{m,q,j} = \sum_{t=m}^{\infty} \delta_{t,q,j}$  by noting that  $\Delta_{m,q,j} \leq \|X_{\cdot j}\|_{q,\alpha} \cdot m^{-\alpha}$  for all  $m \geq 1$ . In this sense, it quantifies the strength of temporal dependence: larger  $\alpha$  implies faster decay of tail dependence measures and thus weaker temporal dependence. We can interpret the quantity  $\|X_{\cdot j}\|_{q,\alpha}$  as the  $q$ -th moment of the component process  $(X_{tj})_{t \in \mathbb{Z}}$  by taking dependence into account. We then define the uniform dependence adjusted norms by taking the maximum over all the component processes

$$\Phi_{q,\alpha} = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_{q,\alpha}.$$

Both quantities  $\|X_{\cdot j}\|_{q,\alpha}$  and  $\Phi_{q,\alpha}$  are used to account for high dimensionality. They may be unbounded functions in terms of the dimension  $p$ . For  $d \in \mathbb{N}$ , we also define

$$\Psi_{q,d} = \max_{1 \leq j \leq p} \sum_{t=0}^{\infty} t^d \delta_{t,q,j}.$$

This quantity is quite useful for verifying cumulative summability in our framework.

Theorem 8 below provides a non-asymptotic bound of the stochastic part

$$\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty$$

with the existence of finite  $q$ -th moment of the underlying process and Proposition 9 concerns the bias of  $\widehat{F}(\theta)$ .

**Theorem 8.** *Assume  $\|X_\cdot\|_\infty\|_{q,\alpha} < \infty$ , for some  $q > 4$  and  $\alpha > 0$ . Define*

$$\Upsilon_{q,\alpha} = \log(Bp) \|X_\cdot\|_\infty\|_{q,\alpha} \wedge \log B \left( \sum_{j=1}^p \|X_{\cdot j}\|_{q,\alpha}^{q/2} \right)^{2/q}.$$

*Then there exist constants  $C, C_\alpha, C_{q,\alpha} > 0$  such that for any  $x > 0$ ,*

$$\begin{aligned} & \mathbb{P}\left(\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty \geq x\right) \\ & \leq C_{q,\alpha} x^{-q/2} \Upsilon_{q,\alpha}^q T^{-q/2} V_{T,B} + Cp^2 B \exp\left(-\frac{Tx^2}{C_\alpha \Phi_{4,\alpha}^4 B}\right), \end{aligned} \quad (1.5.3)$$

*where  $V_{T,B} = TB^{q/2-1}$  if  $\alpha > 1/2 - 2/q$  and  $V_{T,B} = TB^{q/2-1} + T^{q/4-\alpha q/2} B^{q/4}$  if  $\alpha < 1/2 - 2/q$ .*

Based on (1.5.3), it can be obtained that

$$\max_{\theta \in \Theta} |\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta)|_\infty = O_{\mathbb{P}}(R_{T,B}),$$

where

$$R_{T,B} = \left(\frac{B}{T}\right)^{1-2/q} \Upsilon_{q,\alpha}^2 + \frac{\sqrt{B}}{T^{\alpha+1/2}} \Upsilon_{q,\alpha}^2 \mathbf{1}\{\alpha < 1/2 - 2/q\} + \sqrt{\frac{B \log(Bp)}{T}} \Phi_{4,\alpha}^2.$$

We shall briefly discuss how these different parameters affect the error rate. Specifically, we consider the cases where each component process is balanced with a similar order of dependence adjusted norm, i.e.,  $\|X_{\cdot j}\|_{q,\alpha} \asymp 1$  for all  $1 \leq j \leq p$ . Assume  $\|X_\cdot\|_\infty\|_{q,\alpha} \asymp p^\tau$

for some  $0 \leq \tau \leq 1/q$  in view of  $\|X\|_{q,\alpha} \leq (\sum_{j=1}^p \|X_{\cdot j}\|_{q,\alpha}^q)^{1/q}$ . Then  $\Upsilon_{q,\alpha} \asymp p^\tau \log(Bp)$ . We can observe that  $R_{T,B}$  is non-increasing with  $q$  and  $\alpha$ , and non-decreasing with  $p$  and the dependence norms. As a result, the error rate is less sharper when the data has a heavier tail, higher dimension and stronger dependence. In particular, two types of dependency affect the result: temporal dependence characterized by  $\alpha$  and cross-sectional dependence quantified by  $\|X\|_{q,\alpha}$ . Stronger temporal dependence can increase the error rate by noting that  $\sqrt{B}\Upsilon_{q,\alpha}^2/T^{\alpha+1/2}$  is the additional cost for  $\alpha < 1/2 - 2/q$  and it decreases with  $\alpha$ , while stronger cross-sectional dependence can sharpen the error rate as  $\Upsilon_{q,\alpha}$  is increasing with  $\tau$ . It is noted that that  $\tau = 0$  corresponds to the strongest cross-sectional dependence, which could happen for strongly linearly correlated component processes, for example,  $X_{tj} = a_j X_{t1}$  with  $C_1 \leq |a_j| \leq C_2$  for some constants  $C_1, C_2 > 0$ . In this case, it delivers a rate increasing with the dimension  $p$  logarithmically. When  $\tau > 0$ , it shows a polynomial increase with  $p$ .

We can select the window size  $B$  by considering a trade-off between the deviation bound and the bias order provided in Proposition 9 below. The order of optimal  $B$  can be analytically expressed in terms of various parameters without much difficulty, and the discussion is thus omitted here.

**Proposition 9.** *Assume  $\Phi_{2,\alpha} < \infty$  for some  $\alpha > 0$  and  $\Psi_{2,1} < \infty$ . Then we have*

$$\max_{\theta \in \Theta} |\mathbb{E}\hat{F}(\theta) - F(\theta)|_\infty \lesssim B^{-\alpha}\Phi_{2,0}\Phi_{2,\alpha} + B^{-1}\Phi_{2,0}\Psi_{2,1}. \quad (1.5.4)$$

Consistent with Theorem 3, we establish the distribution theory of the spectral estimate in Theorem 10.

**Theorem 10.** *Assume  $B/T \rightarrow 0$  and*

$$\mathcal{E}_B = B^{-\alpha/(2\alpha+1)}\Phi_{4,0}^3\Phi_{4,\alpha} + B^{-1}\Phi_{4,0}^2\Psi_{4,2}^2 \rightarrow 0. \quad (1.5.5)$$

Take any  $\theta, \theta' \in \Theta$  and any  $1 \leq i, j, h, l \leq p$ . Then there exists some bivariate (complex) normal random variable  $\boldsymbol{\xi} = (\xi_{ij}(\theta), \xi_{hl}(\theta'))^\top$  such that

$$\sqrt{\frac{T}{B}} \begin{pmatrix} \widehat{F}_{ij}(\theta) - \mathbb{E}\widehat{F}_{ij}(\theta) \\ \widehat{F}_{hl}(\theta') - \mathbb{E}\widehat{F}_{hl}(\theta') \end{pmatrix} - \boldsymbol{\xi} \xrightarrow{d} \mathbf{0}.$$

(i) If  $\theta = \theta' \in (0, \pi)$ ,  $\boldsymbol{\xi} = (\xi_{ij}(\theta), \xi_{hl}(\theta))^\top$  is a bivariate complex normal distribution with mean 0, covariance matrix

$$\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^H) = \begin{pmatrix} F_{ii}(\theta)F_{jj}(\theta) & F_{ih}(\theta)F_{lj}(\theta) \\ F_{hi}(\theta)F_{jl}(\theta) & F_{hh}(\theta)F_{ll}(\theta) \end{pmatrix}$$

and relation matrix (or pseudo-covariance matrix)

$$\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^\top) = \begin{pmatrix} F_{ij}^2(\theta) & F_{il}(\theta)F_{hj}(\theta) \\ F_{il}(\theta)F_{hj}(\theta) & F_{hl}^2(\theta) \end{pmatrix}.$$

(ii) If  $\theta = \theta' = 0$  or  $\pi$ ,  $\boldsymbol{\xi}$  is a bivariate normal distribution with mean 0 and covariance matrix

$$E(\boldsymbol{\xi}\boldsymbol{\xi}^\top) = \begin{pmatrix} F_{ii}(\theta)F_{jj}(\theta) + F_{ij}^2(\theta) & F_{ih}(\theta)F_{lj}(\theta) + F_{il}(\theta)F_{hj}(\theta) \\ F_{ih}(\theta)F_{lj}(\theta) + F_{il}(\theta)F_{hj}(\theta) & F_{hh}(\theta)F_{ll}(\theta) + F_{hl}^2(\theta) \end{pmatrix}.$$

(iii) If  $\theta \neq \theta'$ ,  $\xi_{ij}(\theta)$  and  $\xi_{hl}(\theta')$  are independent (complex) normal distributions. In particular, the off-diagonals in matrices  $\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^H)$  and  $\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^\top)$  are all 0.

Similarly, with the bias concerned, we can replace  $\mathbb{E}\widehat{F}_{ij}(\theta)$  with  $F_{ij}(\theta)$  in the statement of Theorem 10 under the additional conditions  $T\Phi_{2,0}\Phi_{2,\alpha} = o(B^{2\alpha+1})$  and  $T\Phi_{2,0}\Psi_{2,1} = o(B^3)$ .

### 1.5.2 Gaussian Approximation for High Dimensional Spectral Estimates

In this section, we shall discuss the results for Gaussian approximation with finite polynomial moments. Recall the maximum deviation defined in Section 1.3.1

$$\mathbf{F} = \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta))|_\infty, |\operatorname{Im}(\widehat{F}(\theta) - \mathbb{E}\widehat{F}(\theta))|_\infty \}.$$

Theorem 11 below provides the analogue of Theorem 4, established under finite polynomial moment conditions. To state Theorem 11, we introduce the following quantities:

$$\begin{aligned} L_1 &= (\Phi_{4,0}\Phi_{4,\alpha}\log(Bp))^{2/\alpha}, \\ L_2 &= ((B/T)^{1-4/q}(\log(Bp))^5 \| |X|_\infty \|_{q,\alpha}^4)^{1/\alpha}, \\ R_1 &= B(\Phi_{4,0}\Phi_{4,\alpha}\log(Bp))^{-4}, \\ R_2 &= T^{1-4/q}B^{4/q}(\log(Bp))^{-5} \| |X|_\infty \|_{q,\alpha}^{-4}. \end{aligned}$$

**Theorem 11.** *Let Assumption 2 and Assumption 3 be satisfied. Assume  $\| |X|_\infty \|_{q,\alpha} < \infty$ , for some  $q > 4$  and  $\alpha > 0$ . Further assume that*

$$B^{-1}(\log(Bp))^2 \Phi_{4,0}^2 \Psi_{4,2}^2 \rightarrow 0, \tag{1.5.6}$$

$$\| |X|_\infty \|_{q,\alpha}^4 \Phi_{8,0}^{12} (B/T)^{1-4/q} (\log(Bp))^2 (\log(Tp))^7 \rightarrow 0, \tag{1.5.7}$$

and

$$\max(L_1, L_2) = o(1) \min(R_1, R_2). \tag{1.5.8}$$

(i) If  $\alpha > 1/2 - 2/q$ , then

$$d_{T,B} := \sup_{x \in \mathbb{R}} |\mathbb{P}(\mathbf{F} \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \rightarrow 0. \quad (1.5.9)$$

(ii) If  $\alpha < 1/2 - 2/q$ , (1.5.9) holds under the additional condition

$$T^{-2\alpha}(\log(Bp))^5 \| |X|_\infty \|_{q,\alpha}^4 \rightarrow 0. \quad (1.5.10)$$

In the low-dimensional case when  $p$  is a fixed number, with all the dependence norms of a constant order, the conditions of Theorem 11 are simplified to the very mild requirement  $B = o(T/(\log T)^C)$  for some  $C > 0$ . We can also investigate how large the dimension  $p$  is allowed for the validity of the Gaussian approximation result in high dimensions. Similarly as discussed in Section 1.3.1, we consider the balanced case where  $\|X_{\cdot j}\|_{q,\alpha} \asymp 1$  and  $\| |X|_\infty \|_{q,\alpha} \asymp p^\tau$  for some  $\tau \geq 0$ . In addition, assume that  $\Psi_{4,2} \asymp 1$ . Elementary but tedious calculations show that the conditions of Theorem 11 first require  $\log p = o(B^{\frac{\alpha}{2+4\alpha}})$ . Moreover, in the weaker temporal dependence case with  $\alpha > 1/2 - 2/q$ , if  $\tau = 0$ , it should further satisfy  $\log p = o(\mathcal{H}^{C_1})$  for some  $0 < C_1 < 1$ , where

$$\mathcal{H} = \frac{(T/B)^{1/4-1/q}}{(\log T)^{C_2}}$$

for some  $C_2 > 0$ , and it requires  $p = o(\mathcal{H}^{1/\tau})$  when  $\tau > 0$ . As a result, if we choose  $B = O(T^{\beta_1})$  and  $T/B = O(T^{\beta_2})$  for some  $0 < \beta_1, \beta_2 < 1$ ,  $p$  is allowed to grow at either an exponential rate ( $\tau = 0$ ) or polynomial rate ( $\tau > 0$ ) as a function of  $T$  depending on the moment order  $q$ , the temporal dependence strength  $\alpha$  and the cross-sectional dependence strength  $\tau$ . A larger dimension is allowed when the distribution of the process has a lighter tail (larger  $q$ ), the temporal dependence is weaker (larger  $\alpha$ ) and the cross-sectional dependence is stronger (smaller  $\tau$ ). For the case  $\alpha < 1/2 - 2/q$  that allows stronger temporal dependence, it is additionally required that  $\log p = o(T^{2\alpha/5})$  when  $\tau = 0$  and

$p = o((T^{\alpha/2}/(\log T)^{5/4})^{1/\tau})$  when  $\tau > 0$ .

Next, we also consider the version centered by  $F(\theta)$ . Define

$$\mathbf{F}^\star = \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\widehat{F}(\theta) - F(\theta))|_\infty, |\operatorname{Im}(\widehat{F}(\theta) - F(\theta))|_\infty \}. \quad (1.5.11)$$

By accounting for the bias of  $\widehat{F}(\theta)$ , the Gaussian approximation result also holds for  $\mathbf{F}^\star$ ; see Corollary 12 below.

**Corollary 12.** *Let the assumptions of Theorem 11 be satisfied. Further assume that*

$$\frac{\Phi_{2,0}\Phi_{2,\alpha}T^{1/2}\sqrt{\log(Bp)}}{B^{\alpha+1/2}} + \frac{\Phi_{2,0}\Psi_{2,1}T^{1/2}\sqrt{\log(Bp)}}{B^{3/2}} \rightarrow 0. \quad (1.5.12)$$

Then we have

$$d_{T,B}^\star := \sup_{x \in \mathbb{R}} |\mathbb{P}(\mathbf{F}^\star \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \rightarrow 0.$$

### 1.5.3 Resampling Methods

In this section, we establish the validity of the two resampling methods discussed in Section 1.3.3 under finite polynomial moment conditions.

Theorem 13 below presents the corresponding result to Theorem 6, establishing the validity of the resampling method under polynomial moment conditions.

**Theorem 13.** *Let the conditions in Corollary 12 be satisfied. Further assume*

$$R_{T,B}^\star \Phi_{2,0}^2 (\log(Bp))^2 \rightarrow 0$$

with

$$\begin{aligned} R_{T,B}^\star &= T^{-1} V_{T,B}^{2/q} \| |X|_\infty \|_{q,\alpha}^2 (\log(Bp))^2 + (B/T)^{1/2} \Phi_{4,\alpha}^2 (\log(Bp))^{1/2} \\ &\quad + B^{-\alpha} \Phi_{2,0} \Phi_{2,\alpha} + B^{-1} \Phi_{2,0} \Psi_{2,1}, \end{aligned}$$

where  $V_{T,B}$  is the same as defined in Theorem 8. Then

$$\sup_{\alpha_0 \in (0,1)} |\mathbb{P}(\mathbf{F}^\star \geq \widehat{c}_{1-\alpha_0}) - \alpha_0| \rightarrow 0. \quad (1.5.13)$$

Theorem 14 establishes the validity of the subsampling approach, serving as the analogue of Theorem 7 under finite polynomial moment conditions.

**Theorem 14.** *Assume that  $M = o(T)$  and  $B = o(M)$ . Let the conditions of Corollary 12 be satisfied, in which  $T$  is replaced by the subsample size  $M$ . Further assume the following conditions are met as  $B, M, T \rightarrow \infty$ ,*

$$\begin{aligned} (i) \quad & T^{-1} M (\log(Bp))^5 \Phi_{2,0}^4 \|X.\|_\infty^4 \rightarrow 0. \\ (ii) \quad & B^{-1} M (\log(Bp))^3 (B^{-2\alpha} \Phi_{2,0}^6 \Phi_{2,\alpha}^2 + B^{-2} \Phi_{2,0}^6 \Psi_{2,1}^2) \rightarrow 0. \\ (iii) \quad & T^{-1} R(\mathcal{V}) (\log(Bp))^{3/2} \Phi_{2,0}^2 \rightarrow 0. \end{aligned}$$

where  $\mathcal{V} = \log(Bp) \|X.\|_\infty \|X.\|_{4,\alpha}$ ,  $R(\mathcal{V}) = \mathcal{V}^{1/\alpha}$  if  $\alpha > 1$ ,  $R(\mathcal{V}) = \mathcal{V} \log T$  if  $\alpha = 1$  and  $R(\mathcal{V}) = \mathcal{V} T^{1-\alpha}$  if  $\alpha < 1$ . Then we have

$$\sup_{x \in \mathbb{R}} |\mathbf{Q}(x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \xrightarrow{\mathbb{P}} 0.$$

## 1.6 Numerical Results

### 1.6.1 Simulation Study

In this section, we conduct a simulation study to evaluate the accuracy of the Gaussian approximation results presented in Section 1.3.1, followed by an assessment of the performance of the Gaussian multiplier resampling method and the subsampling

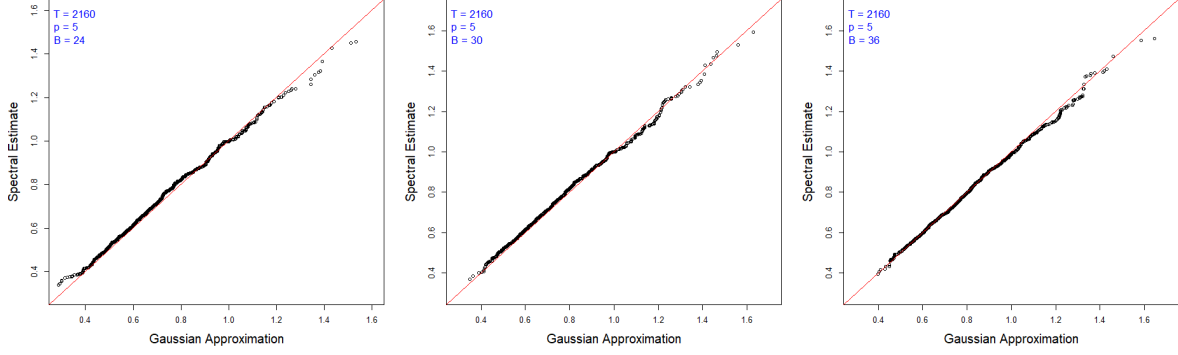
method proposed in Section 1.3.2. We consider the linear process

$$X_t = \sum_{k=0}^{\infty} A_k \varepsilon_{t-k}, \quad (1.6.1)$$

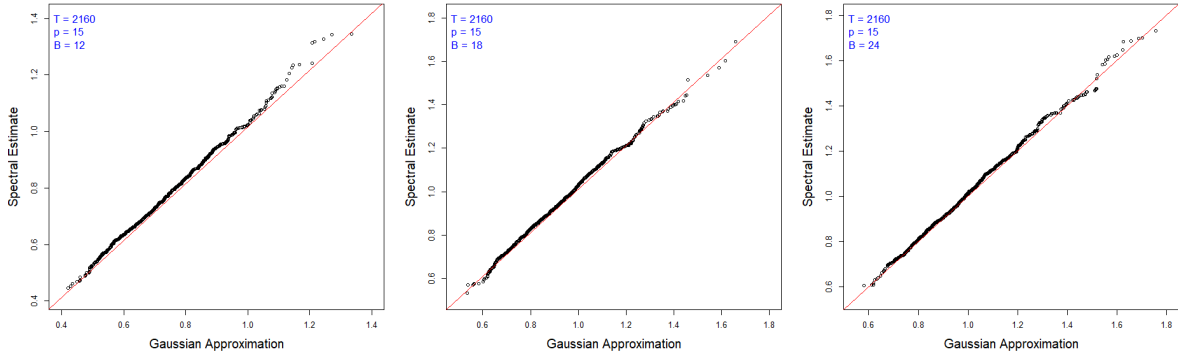
where the coefficient matrices  $A_k$  are constructed with a Toeplitz structure, parameterized by a correlation factor  $\rho$ . Specifically,  $A_k = \rho^k M$  where  $M = (m_{ij})_{i,j=1}^p$  and  $m_{ij} = \rho^{|i-j|+1}$ , with  $\rho \in (0, 1)$ . To ensure the process exhibits sub-exponential decay, we consider errors  $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{ip})$ , where  $\varepsilon_{ij}$  are i.i.d. random variables following a Gamma distribution  $\Gamma(2, 1)$ . We adopt the following numerical configurations for our analysis: decay rate  $\rho = 0.5$ , sample size  $T = 2160$ , dimension  $p = 5, 15$ , with the sum in (1.6.1) truncated to  $X_t = \sum_{k=0}^{2000} A_k \varepsilon_{t-k}$ .

It is worth noting that the window size  $B$  and subsample size  $M$  play important roles in the inference procedure. In high-dimensional settings, where both the sample size  $T$  and the vector dimension  $(B-2)p^2$  in Gaussian approximation can be large, selecting these tuning parameters is challenging and sensitive to constant scaling. In practice, we recommend  $B = T^{1/2}/\log p$ , which minimizes the error bound in Theorem 4 sharpest in terms of the polynomial orders of  $T$  and  $\log p$ . Given the complicated error bounds in the subsampling procedure, determining an ideal order for the subsample size  $M$  is highly non-trivial. We suggest using  $M = T^{1/3}B$ , which meets the additional conditions (1.3.7) and (1.3.8). With the suggested orders  $B = T^{1/2}/\log p$  and  $M = T^{1/3}B$ , we adjust  $B$  to the nearest even divisor of  $T$  and  $M$  to the closest integer divisible by  $B$  for computational convenience. For example, for  $p = 5$ , it can be computed that  $T^{1/2}/\log p \approx 28.877$ , so we set  $B = 30$ , an even divisor of  $T$  and choose  $M = 360$ . For comparison, we also examine a smaller bandwidth  $B = 24$  with  $M = 288$  and a larger bandwidth  $B = 36$  with  $M = 432$ . Similarly, for  $p = 15$ , we use  $B = 18$  and  $M = 216$  as the best choices, and include comparisons with  $B = 12$ ,  $M = 144$ , and  $B = 24$ ,  $M = 288$ .

Figure 1.1 and Figure 1.2 display Q-Q plots comparing  $\mathbf{F}^*$ , as defined in (1.3.3),



**Figure 1.1.** Q-Q plot comparing maximum deviation  $F^*$  and its Gaussian approximation  $|Z|_\infty$  when  $p = 5$ .



**Figure 1.2.** Q-Q plot comparing maximum deviation  $F^*$  and its Gaussian approximation  $|Z|_\infty$  when  $p = 15$ .

and its Gaussian approximation  $|Z|_\infty$ , as defined in (1.3.1), based on 1000 repetitions. In each figure, a Q-Q plot is generated for each of the 3 values of  $B$  considered. As shown, the fit aligns closely with the reference line when  $p = 5$  and show slightly less accuracy at the tail but remains satisfactory when  $p = 15$ . These results validate our Gaussian approximation through simulations. We also observe that the suggested  $B$  provides the best approximation across each set of three plots.

We further employ the Gaussian multiplier resampling method and subsampling to evaluate the validity of these two approaches. Specifically, we compare the values of  $F^*$  with the cut-off values  $\hat{c}_{1-\theta_0}$  obtained from the resampling and subsampling methods discussed in Section 1.3.2 using nominal levels  $\theta_0 = 0.1$  and  $0.05$ , corresponding to the

90% and 95% empirical quantiles respectively. Table 1.1 shows the coverage probability for each case, representing the proportion of instances where the cut-off value  $\widehat{c}_{1-\theta_0}$  exceeds  $F^*$  over 1000 repetitions. The results indicate that our simultaneous confidence bands achieve the desired coverage levels. Additionally, the procedure becomes more sensitive to the choice of  $B$  as  $p$  increases, particularly for the Gaussian multiplier resampling method.

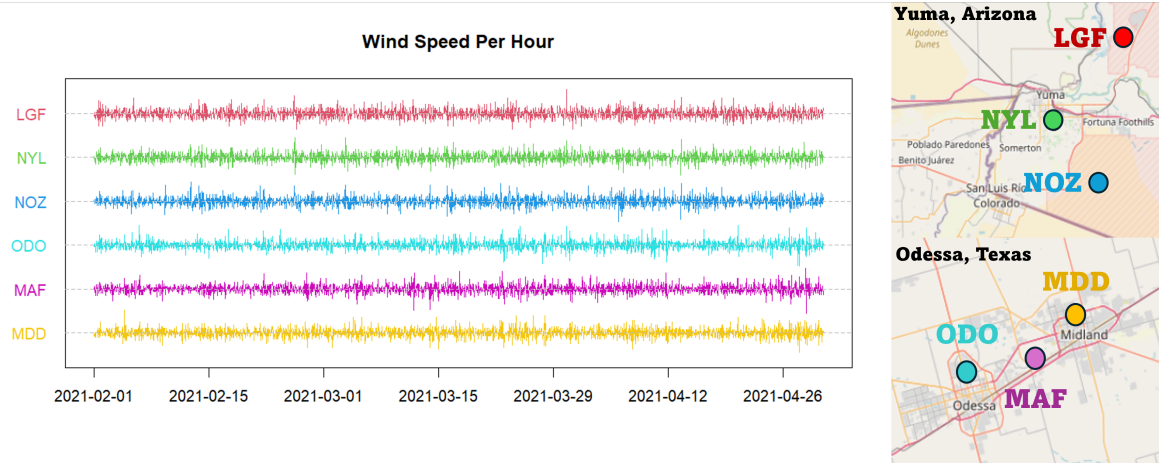
**Table 1.1.** Empirical coverages of simultaneous confidence bands.

| Dimension | Bandwidth | Resampling |      | Subsampling |      |
|-----------|-----------|------------|------|-------------|------|
|           |           | 90%        | 95%  | 90%         | 95%  |
| $p = 5$   | $B = 24$  | 87.8       | 93.4 | 87.6        | 92.9 |
|           | $B = 30$  | 89.8       | 94.3 | 89.4        | 95.1 |
|           | $B = 36$  | 90.6       | 94.7 | 88.1        | 93.0 |
| $p = 15$  | $B = 12$  | 84.7       | 91.2 | 89.3        | 94.7 |
|           | $B = 18$  | 87.5       | 93.7 | 90.3        | 95.2 |
|           | $B = 24$  | 86.4       | 92.7 | 87.4        | 94.6 |

## 1.6.2 Data Analysis

In this section, we test out method on wind data obtained from the Iowa State University Environmental Mesonet (IEM) Automated Surface Observing System (ASOS) Network, a publicly accessible database ([3], [96]). ASOS stations, primarily located at airports, provide weather observations and general reports for the National Weather Service (NWS), the Federal Aviation Administration (FAA), and the Department of Defense (DOD). For this study, we analyze data collected from six airports, three located in Arizona: LGF (Laguna Army Airfield), NYL (Yuma International Airport), and NOZ (Rolle Airfield), and three in Texas: ODO (Odessa-Schlemeyer Field), MAF (Midland International Air and Space Port), and MDD (Midland Airpark). By examining the cross-spectra across these 15 pairs, we aim to test for pairwise independence between each pair of wind speed time series, based on the theory presented in Section 1.3.2.

We consider the hourly wind speed from 02/01/2021 to 04/30/2021 for each location. Since our method relies on the assumption of stationarity, we select this time period to



**Figure 1.3.** Standardized first-order differences of hourly wind speed series between 02/01/2021 to 04/30/2021 from selected stations in Arizona and Texas.

avoid extreme values and non-stationarity associated with rainfall and storms, which are prevalent in other months. We note that there is virtually no rainfall during this period at these locations. Previous studies on wind behavior have shown that certain locations exhibit quasi-periodic patterns influenced by geographic and climatic factors, such as jet streams in summer and winter ([93]), which we aim to avoid in our analysis. To eliminate any local trends, we conduct a joint analysis of the 6-dimensional data by taking the first-order differences of each time series. Each differenced series is further standardized by subtracting its mean and dividing by its standard deviation to improve computational stability. Figure 1.3 presents the resulting 6-dimensional time series data, consisting of observations at 2136 time points, along with the 6 airport locations.

We shall consider the hypothesis test:  $H : F_{ij}(\theta) = 0$  for each pairs among the six airports, where  $1 \leq i < j \leq 6$  with significance level  $\theta_0 = 0.05$ . The test statistic  $\tilde{\mathbf{F}}_{ij}^*$ , defined in Section 1.3.2, is employed. Unlike in the simulation study, the true  $\Sigma$  is not available in real data analysis. To validate our test, we apply resampling and subsampling methods as discussed in Section 1.3.3, obtaining (i) the conditional  $(1 - \alpha)$ -quantile  $\hat{c}_{1-\alpha}$  via resampling, and (ii) the empirical quantile from the empirical distribution  $Q(x)$  by subsampling. Following the recommendations discussed in Section 1.6.1, we set  $B = 24$

and  $M = 288$ .

Table 1.2 presents the test statistics and cut-off values obtained through resampling and subsampling. Results indicate that for pairs within the groups (LGF, NYL, NOZ) and (ODO, MAF, MDD), the test statistics significantly exceed the corresponding cut-off values, leading us to reject the null hypothesis of independence within these pairs. Additionally, for tests between pairs from different groups, the test statistics are smaller than the cut-off values, so we fail to reject the null hypothesis, supporting independence between these pairs. This finding aligns with geographic proximity, as locations close to each other tend to share similar wind profile patterns, whereas locations farther apart do not exhibit this similarity.

**Table 1.2.** Pairwise Cross-Spectral Test Results

| Location | Station pairs | Test statistic | Rsampling cut-off value | Subsampling cut-off value |
|----------|---------------|----------------|-------------------------|---------------------------|
| Yuma     | LGF & NYL     | 0.648          | 0.418                   | 0.391                     |
|          | LGF & NOZ     | 0.593          | 0.420                   | 0.507                     |
|          | NYL & NOZ     | 0.545          | 0.431                   | 0.493                     |
| Odessa   | ODO & MAF     | 0.905          | 0.399                   | 0.496                     |
|          | ODO & MDD     | 0.805          | 0.408                   | 0.431                     |
|          | MAF & MDD     | 0.889          | 0.417                   | 0.433                     |
| Yuma     | LGF & ODO     | 0.321          | 0.400                   | 0.445                     |
|          | LGF & MAF     | 0.311          | 0.416                   | 0.410                     |
|          | LGF & MDD     | 0.333          | 0.412                   | 0.496                     |
| vs.      | NYL & ODO     | 0.299          | 0.412                   | 0.446                     |
|          | NYL & MAF     | 0.225          | 0.413                   | 0.402                     |
|          | NYL & MDD     | 0.264          | 0.421                   | 0.392                     |
| Odessa   | NOZ & ODO     | 0.263          | 0.410                   | 0.478                     |
|          | NOZ & MAF     | 0.286          | 0.422                   | 0.400                     |
|          | NOZ & MDD     | 0.372          | 0.418                   | 0.409                     |

## 1.7 Proofs

In this section, we provide the proofs of the results introduced in the main body.

### 1.7.1 Proofs of Results in Section 1.2

*Proof of Theorem 1.* Denote  $n = T/B$ . For  $1 \leq r \leq n$  and  $1 \leq i \leq p$ , define

$$Y_{ri}(\theta) = \sum_{t=(r-1)B+1}^{rB} X_{ti} e^{-it\theta}.$$

To prove the result (1.2.6), we shall first consider the quadratic form

$$R_{ij}(\theta) = 2\pi T \widehat{F}_{ij}(\theta) = \sum_{r=1}^n Y_{ri}(\theta) Y_{rj}(\theta)$$

and apply the Markov inequality: for  $C_0 > 0$  and  $\eta > 0$  which are to be determined later,

$$\mathbb{P}(|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)| \geq x) \leq e^{-\eta x^{C_0}} \mathbb{E}[\exp(\eta |R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)|^{C_0})]. \quad (1.7.1)$$

Take the martingale decomposition  $R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta) = \sum_{l=-\infty}^T \mathcal{P}_l R_{ij}(\theta)$ . By Theorem 2.1 of [74], we have for  $q \geq 4$ ,

$$\begin{aligned} \|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)\|_{q/2}^2 &\leq (q/2 - 1) \sum_{l=-\infty}^T \|\mathcal{P}_l \sum_{r=1}^n Y_{ri}(\theta) Y_{rj}(\theta)\|_{q/2}^2 \\ &\leq (q/2 - 1) \sum_{l=-\infty}^T \left( \sum_{r=1}^n \|\mathcal{P}_l Y_{ri}(\theta) Y_{rj}(\theta)\|_{q/2} \right)^2. \end{aligned} \quad (1.7.2)$$

By Equation (33) in Proposition 8 of [110], we can get

$$\|Y_{ri}(\theta)\|_q \leq (q-1)^{1/2} \sqrt{B} \Delta_{0,q,i} \leq (q-1)^{1/2} \sqrt{B} \|X_{\cdot i}\|_q.$$

By Equation (31) in Proposition 8 of [110], we have

$$\|Y_{ri}(\theta) - Y_{ri,\{l\}}(\theta)\|_q \leq \sum_{t=(r-1)B+1}^{rB} \delta_{t-l,q,i}.$$

By Jensen's inequality, the triangle inequality and Hölder's inequality, it then follows that

$$\begin{aligned}
& \sum_{r=1}^n \|\mathcal{P}_l Y_{ri}(\theta) Y_{rj}(\theta)\|_{q/2} \\
& \leq \sum_{r=1}^n \|Y_{ri}(\theta) Y_{rj}(\theta) - Y_{rj, \{l\}}(\theta) Y_{rj, \{l\}}(\theta)\|_{q/2} \\
& \leq \sum_{r=1}^n (\|Y_{ri}(\theta)\|_q \|Y_{rj}(\theta) - Y_{rj, \{l\}}(\theta)\|_q + \|Y_{ri} - Y_{ri, \{l\}}(\theta)\|_q \|Y_{rj, \{l\}}(\theta)\|_q) \\
& \leq (q-1)^{1/2} \sqrt{B} \left( \|X_{\cdot i}\|_q \sum_{t=1}^T \delta_{t-l, q, j} + \|X_{\cdot j}\|_q \sum_{t=1}^T \delta_{t-l, q, i} \right).
\end{aligned} \tag{1.7.3}$$

By (1.7.2) and (1.7.3), elementary manipulation implies

$$\|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)\|_{q/2} \leq \sqrt{(q-2)(q-1)} \sqrt{TB} \|X_{\cdot i}\|_q \|X_{\cdot j}\|_q. \tag{1.7.4}$$

Under Assumption 1 and noting that  $\|X_{\cdot}\|_{\psi_\alpha} = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_{\psi_\alpha}$ , (1.7.4) suggests

$$\|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)\|_{q/2} \leq (q-1) q^{2\alpha} \sqrt{TB} \|X_{\cdot}\|_{\psi_\alpha}^2. \tag{1.7.5}$$

Write the negative binomial expansion  $(1-s)^{-1/2} = 1 + \sum_{h=1}^{\infty} c_h s^h$  with

$$c_h = (2h)! / (2^{2h} (h!)^2)$$

for  $|s| < 1$ . By Stirling's formula, we have  $c_h \sim (h\pi)^{-1/2}$  as  $h \rightarrow \infty$ . Hence, there exists absolute constants  $C_1, C_2 > 0$  such that for all  $h \geq 1$ ,

$$C_1 (h/e)^h c_h^{-1} \leq h! \leq C_2 (h/e)^h c_h^{-1}. \tag{1.7.6}$$

Let  $C_0 = 1/(2\alpha + 1)$ . For any positive integer  $h$  such that  $C_0h \geq 2$ , (1.7.5) leads to

$$\|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)\|_{C_0h} \leq (2C_0h - 1)(2C_0h)^{2\alpha}\sqrt{TB}\|X.\|_{\psi_\alpha}^2.$$

Let  $\eta_0 = (2eC_0T^{C_0/2}B^{C_0/2}\|X.\|_{\psi_\alpha}^{2C_0})^{-1}$ . Notice that  $2\alpha + 1 = 1/C_0$ . Then

$$\begin{aligned} & \frac{\eta^h \mathbb{E}|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)|^{C_0h}}{h!} \\ & \leq \frac{\eta^h (2C_0h - 1)^{C_0h} (2C_0h)^{2\alpha C_0h} (TB)^{C_0h/2} \|X.\|_{\psi_\alpha}^{2C_0h}}{C_1(h/e)^h c_h^{-1}} \\ & \leq \frac{c_h \eta^h (2C_0h - 1)^{C_0h}}{C_1 \eta_0^h (2C_0h)^{C_0h}} \leq \frac{c_h \eta^h}{C_1 \sqrt{e} \eta_0^h}. \end{aligned}$$

If  $C_0h < 2$ , then  $\|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)\|_{C_0h} \leq \|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)\|_2 \leq \sqrt{6} \cdot 4^{2\alpha} \sqrt{TB} \|X.\|_{\psi_\alpha}^2$ . So we have

$$\begin{aligned} & \mathbb{E}[\exp(\eta|R_{jk}(\theta) - \mathbb{E}R_{jk}(\theta)|^{C_0})] \\ & \leq 1 + \sum_{1 \leq h < 2/C_0} \frac{\eta^h (\sqrt{6} \cdot 4^{2\alpha} \sqrt{TB} \|X.\|_{\psi_\alpha}^2)^{C_0h}}{h!} + \sum_{h \geq 2/C_0} \frac{c_h \eta^h}{C_1 \sqrt{e} \eta_0^h} \\ & \leq 1 + C_{C_0} \sum_{h=1}^{\infty} c_h \frac{\eta^h}{\eta_0^h} \leq 1 + C_{C_0} \frac{\eta/\eta_0}{(1 - \eta/\eta_0)^{1/2}}. \end{aligned}$$

By choosing  $\eta = \eta_0/2$  in (1.7.1) and recalling  $C_0 = 1/(2\alpha + 1)$ , we have

$$\mathbb{P}(|R_{ij}(\theta) - \mathbb{E}R_{ij}(\theta)| \geq x) \lesssim \exp \left\{ -C_\alpha \left( \frac{x}{\sqrt{TB} \|X.\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}. \quad (1.7.7)$$

As a result, (1.2.6) follows in view of (1.7.7), the Bonferroni inequality and the simple scaling  $\widehat{F}_{ij}(\theta) = (2\pi T)^{-1} R_{ij}(\theta)$ .  $\square$

*Proof of Proposition 2.* By the stationarity of the process, we have

$$\mathbb{E}\widehat{F}_{ij}(\theta) = \frac{1}{2\pi} \sum_{k=-B}^B \left(1 - \frac{|k|}{B}\right) \Gamma_{ij}(k) e^{-ik\theta}.$$

It follows that

$$2\pi|\mathbb{E}\widehat{F}_{ij}(\theta) - F_{ij}(\theta)| \leq \sum_{k:|k|\geq B} |\Gamma_{ij}(k)| + \sum_{k=-B+1}^{B-1} \frac{|k|}{B} |\Gamma_{ij}(k)| := A_1 + A_2. \quad (1.7.8)$$

Since  $X_{ti} = \sum_{r=0}^{\infty} \mathcal{P}_{t-r}(X_{ti})$ , for  $k \geq 0$ , we have

$$|\Gamma_{ij}(k)| = \left| \sum_{r=0}^{\infty} \mathbb{E}[\mathcal{P}_{t-r}(X_{ti}) \mathcal{P}_{t-r}(X_{(t+k)j})] \right| \leq \sum_{r=0}^{\infty} \delta_{r,2,i} \delta_{r+k,2,j}, \quad (1.7.9)$$

which implies  $A_1 \lesssim \Delta_{0,2,i} \Delta_{B,2,j}$ . In view of  $\Delta_{0,2,i} \leq \|X_{\cdot i}\|_2$  and  $\Delta_{B,2,j} \leq \rho^B \|X_{\cdot j}\|_2$ , we have  $A_1 \lesssim \rho^B \|X_{\cdot i}\|_2 \|X_{\cdot j}\|_2$ . Furthermore, we can get

$$\begin{aligned} A_2 &\leq 2B^{-1} \sum_{k=1}^{B-1} \sum_{r=0}^{\infty} k \delta_{r,2,i} \delta_{r+k,2,j} \\ &\leq 2B^{-1} \|X_{\cdot j}\|_2 \sum_{k=1}^{B-1} k \rho^k \sum_{r=0}^{\infty} \delta_{r,2,i} \\ &\lesssim B^{-1} \|X_{\cdot i}\|_2 \|X_{\cdot j}\|_2. \end{aligned} \quad (1.7.10)$$

By combining the bounds of  $A_1$  and  $A_2$ , we have for any  $1 \leq i, j \leq p$  and  $\theta$ ,

$$|\mathbb{E}\widehat{F}_{ij}(\theta) - F_{ij}(\theta)| \lesssim B^{-1} \|X_{\cdot i}\|_2 \|X_{\cdot j}\|_2. \quad (1.7.11)$$

□

*Proof of Theorem 3.* Let  $n = T/B$ . Define  $\mathcal{B}_r = \{(r-1)B + 1, \dots, rB\}$  and  $\mathcal{D}_r =$

$\{(r-1)B+1, \dots, (r-1)B+B-m\}$  for  $r=1, \dots, n$ . Then we can write

$$\widehat{F}(\theta) = (\widehat{F}_{ij}(\theta))_{i,j=1}^p = \frac{1}{2\pi T} \sum_{r=1}^n \sum_{t,s \in B_r} X_t X_s^\top e^{-\iota(t-s)\theta}.$$

For  $m \geq 0$ , define  $X_{t,m} = \mathbb{E}(X_t | \varepsilon_{t-m}, \varepsilon_{t-m+1}, \dots, \varepsilon_t)$ . Let

$$\widehat{F}_m(\theta) = (\widehat{F}_{ij,m}(\theta))_{i,j=1}^p = \frac{1}{2\pi T} \sum_{r=1}^n \sum_{t,s \in B_r} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}.$$

Further define

$$\begin{aligned} \widehat{H}_{r,m}(\theta) &= (\widehat{H}_{ij,r,m}(\theta))_{i,j=1}^p = \frac{1}{2\pi B} \sum_{t,s \in \mathcal{D}_r} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}, \\ \widehat{S}_m(\theta) &= (\widehat{S}_{ij,m}(\theta))_{i,j=1}^p = \frac{1}{n} \sum_{r=1}^n \widehat{H}_{r,m}(\theta). \end{aligned}$$

By Lemma 9 in [111],

$$\|\mathbb{E}_0(\widehat{F}_{ij}(\theta)) - \mathbb{E}_0(\widehat{F}_{ij,m}(\theta))\|_2 \lesssim \sqrt{\frac{B}{T}} \rho^m \Phi_4^2. \quad (1.7.12)$$

Similarly as (1.9.13), we can obtain

$$\|\mathbb{E}_0(\widehat{F}_{ij,m}(\theta)) - \mathbb{E}_0(\widehat{S}_{ij,m}(\theta))\|_2 \lesssim \sqrt{\frac{m}{T}} \Phi_4^2. \quad (1.7.13)$$

Choosing  $m \asymp (\log B)/\log(\rho^{-1})$ , by (1.7.12) and (1.7.13),

$$\sqrt{\frac{T}{B}} \|\mathbb{E}_0(\widehat{F}_{ij}(\theta)) - \mathbb{E}_0(\widehat{S}_{ij,m}(\theta))\|_2 \lesssim \sqrt{\frac{\log B}{B}} \Phi_4^2 \rightarrow 0. \quad (1.7.14)$$

As a result, the distribution of  $\sqrt{T/B}(\widehat{F}_{ij}(\theta) - \mathbb{E}\widehat{F}_{ij}(\theta))$  is dominated by the distribution of  $\sqrt{T/B}(\widehat{S}_{ij}(\theta) - \mathbb{E}\widehat{S}_{ij}(\theta))$ . Note that  $\widehat{H}_{r,m}(\theta)$ ,  $r=1, \dots, n$ , are independent and identically distributed. It remains to concern the covariance structure of  $\widehat{H}_{1,m}(\theta)$ . Take  $m \asymp$

$(\log B)/\log(\rho^{-1})$  in (1.9.12). Under the condition  $\Phi_4^8(\log B)/B \rightarrow 0$ , we have  $\mathcal{E}_{B,m}^{(2)} \rightarrow 0$ . Note that  $(\widehat{H}_{r,m}(\theta))_{r=1}^n$  is an i.i.d. sequence. Theorem 3 follows in view of Lemma 18, Lemma 19, the classical central limit theorem, the Cramér-Wold device and the construction of bivariate complex random variable.  $\square$

### 1.7.2 Proofs of Results in Section 1.3

*Proof of Theorem 4.* We shall mainly focus on the case of  $\mathbf{F}$ .  $\tilde{\mathbf{F}}$  can be dealt with by similar proofs. Recall the definitions of  $\widehat{F}_m(\theta)$  and  $\widehat{S}_m(\theta)$  in (1.9.18) and (1.9.20), respectively. Denote

$$\mathbf{S}_m = (\text{Re}(\widehat{S}_m(\theta_0)), \text{Im}(\widehat{S}_m(\theta_0)), \dots, \text{Re}(\widehat{S}_m(\theta_{B/2})), \text{Im}(\widehat{S}_m(\theta_{B/2}))).$$

Choose  $m \asymp \log B / \log \rho^{-1}$  and denote  $n = T/B$ . We start with

$$\begin{aligned} d_{T,B,p} &\leq \sup_{x \in \mathbb{R}} \left| \mathbb{P}(\mathbf{F} \leq x) - \mathbb{P}(\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0 \mathbf{S}_m|_\infty \leq x) \right| \\ &\quad + \sup_{x \in \mathbb{R}} \left| \mathbb{P}(\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0 \mathbf{S}_m|_\infty \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x) \right| \\ &=: d_{T,B,p}^\circ + d_{T,B,p}^\dagger. \end{aligned} \tag{1.7.15}$$

By the triangle inequality, for every  $y > 0$ , we have

$$\begin{aligned} d_{T,B,p}^\circ &= \sup_{x \in \mathbb{R}} \left| \mathbb{P}(\mathbf{F} \geq x) - \mathbb{P}(\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0 \mathbf{S}_m|_\infty \geq x) \right| \\ &\leq \mathbb{P}(\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{F}(\theta)) - \mathbb{E}_0(\widehat{S}_m(\theta))|_\infty \geq y) + \sup_{x \in \mathbb{R}} \mathbb{P}(|\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0 \mathbf{S}_m|_\infty - x| \leq y) \\ &=: d_{T,B,p}^{\circ(1)} + d_{T,B,p}^{\circ(2)}. \end{aligned}$$

By Lemma 24 and the Bonferroni inequality, for every  $y > 0$ ,

$$\mathbb{P}(\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{F}(\theta)) - \mathbb{E}_0(\widehat{F}_m(\theta))|_\infty \geq y) \lesssim Bp^2 \exp \left\{ -C_\alpha \left( \frac{y}{\rho^m \|X.\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}. \quad (1.7.16)$$

And by Lemma 25, we have

$$\mathbb{P}(\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{F}_m(\theta)) - \mathbb{E}_0(\widehat{S}_m(\theta))|_\infty \geq y) \lesssim Bp^2 \exp \left\{ -C_\alpha \left( \frac{\sqrt{B}y}{\sqrt{m} \|X.\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}. \quad (1.7.17)$$

With the choice of  $m$ , it follows that

$$d_{T,B,p}^{(1)} \lesssim Bp^2 \exp \left\{ -C_\alpha \left( \frac{\sqrt{B}y}{\sqrt{\log B} \|X.\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}. \quad (1.7.18)$$

We next consider

$$\begin{aligned} d_{T,B,p}^{\dagger(1)} &= |\mathbb{P}(\sqrt{n} \max_{\theta \in \Theta} |\mathbb{E}_0 \mathbf{S}_m|_\infty \leq x) - \mathbb{P}(|\mathbf{Z}_m|_\infty \leq x)|, \\ d_{T,B,p}^{\dagger(2)} &= |\mathbb{P}(|\mathbf{Z}_m|_\infty \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \end{aligned}$$

Then we have

$$\begin{aligned} d_{T,B,p}^\dagger &\leq d_{T,B,p}^{\dagger(1)} + d_{T,B,p}^{\dagger(2)}, \\ d_{T,B,p}^{(2)} &\leq 2d_{T,B,p}^{\dagger(1)} + 2d_{T,B,p}^{\dagger(2)} + \sup_{x \in \mathbb{R}} \mathbb{P}(|\mathbf{Z}|_\infty - x| \leq y). \end{aligned}$$

As  $\mathcal{E}_{B,m}^{(2)} = (\rho^m + \sqrt{m/B})\Phi_4^4 \lesssim \sqrt{m/B}\Phi_4^4 = o(1)$ , Assumption 2, Assumption 3 and Lemma 19 ensure that the variance of each non-zero entry of  $\mathbf{S}_m$  is lower bounded. Applying in Lemma 26, we have

$$d_{T,B,p}^{\dagger(1)} \lesssim \left( \frac{\|X.\|_{\psi_\alpha}^{12} \log^{\kappa_1}((Bp) \vee n)}{n} \right)^{\frac{1}{6}}, \quad (1.7.19)$$

for  $\kappa_1 = \max\{7, 2\alpha + 6\}$ . By Lemma 19 again, it follows that

$$|\Sigma_m - \Sigma|_\infty \lesssim \sqrt{\frac{m}{B}} \Phi_4^4. \quad (1.7.20)$$

Lemma 3.1 in [22] implies

$$d_{T,B,p}^{\dagger(2)} \lesssim \left(\sqrt{\frac{m}{B}} \Phi_4^4\right)^{1/3} \log^{2/3}(Bp), \quad (1.7.21)$$

By Lemma 36, for every  $y > 0$ .

$$\sup_{x \in \mathbb{R}} \mathbb{P}(|\mathbf{Z}|_\infty - x| \leq y) \lesssim y \sqrt{\log(Bp)}. \quad (1.7.22)$$

Choosing

$$y = c_\alpha \frac{\|X\|_{\psi_\alpha}^2 \log^{2\alpha+3/2}(Bp)}{B^{1/2}}$$

with a sufficiently large  $c_\alpha$ , the proof is completed by combining the results (1.7.15) – (1.7.22).

To deal with  $\tilde{\mathbf{F}}$ , the Gaussian approximation result in Lemma 26 concerning the class of all hyperrectangles also holds for  $\tilde{\mathbf{F}}$  approximated by the distribution of  $|\tilde{\mathbf{Z}}|_\infty$ . The case of  $\tilde{\mathbf{F}}$  corresponds to the class of  $s$ -sparsely convex sets with  $s = 2$  in the framework of [24]. We can apply the same arguments of Proposition 4.2 in their paper to obtain the Gaussian approximation result. And the anti-concentration inequality still holds for  $|\tilde{\mathbf{Z}}|_\infty$  in view of (1.7.22) and  $|\mathbf{Z}|_\infty \leq |\tilde{\mathbf{Z}}|_\infty \leq \sqrt{2}|\mathbf{Z}|_\infty$ .  $\square$

*Proof of Corollary 5.* Note that

$$d_{T,B,p}^* \leq \sup_{x \in \mathbb{R}} |\mathbb{P}(\mathbf{F}^* \leq x) - \mathbb{P}(\mathbf{F} \leq x)| + d_{T,B,p}.$$

By the triangle inequality, for every  $y > 0$ , we have

$$\begin{aligned} & \sup_{x \in \mathbb{R}} |\mathbb{P}(\mathbf{F}^* \leq x) - \mathbb{P}(\mathbf{F} \leq x)| \\ & \leq \mathbb{P}\left(\sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}\hat{F}(\theta) - F(\theta)|_\infty \geq y\right) + \sup_{x \in \mathbb{R}} \mathbb{P}(|\mathbf{F} - x| \leq y). \end{aligned}$$

By (1.7.22) and the definition of  $d_{T,B,p}$ , we have

$$\sup_{x \in \mathbb{R}} \mathbb{P}(|\mathbf{F} - x| \leq y) \lesssim d_{T,B,p} + y\sqrt{\log(Bp)}.$$

By Proposition 2, choosing  $y = C(T/B)^{1/2}B^{-1}\Phi_2^2$  for a sufficiently large  $C$ , condition (1.3.4) guarantees  $y\sqrt{\log(Bp)} \rightarrow 0$ , which together with Theorem 4 implies  $d_{T,B,p}^* \rightarrow 0$ . The case of  $\tilde{\mathbf{F}}^*$  can be verified in a similar way.  $\square$

*Proof of Theorem 6.* By the formula of entries in  $\Sigma$  as given in Section 1.3.1, the triangle inequality and (1.9.9), we have

$$|\hat{\Sigma} - \Sigma|_\infty \lesssim \Phi_2^2 \max_{\theta \in \Theta} |\hat{F}_{ij}(\theta) - F_{ij}(\theta)|_\infty. \quad (1.7.23)$$

Theorem 1 and Proposition 2 imply

$$\max_{\theta \in \Theta} |\hat{F}_{ij}(\theta) - F_{ij}(\theta)|_\infty \lesssim \sqrt{\frac{B}{T}} \log^{2\alpha+1}(Bp) \|X\|_{\psi_\alpha}^2 + \frac{\|X\|_{\psi_\alpha}^2}{B}.$$

Following the arguments of Theorem 3.1 in [22], for every  $x > 0$ ,

$$\sup_{\theta_0 \in (0,1)} |\mathbb{P}(\mathbf{F}^* \geq \hat{c}_{1-\theta_0}) - \theta_0| \lesssim d_{T,B}^* + \mathcal{W}(x) + \mathbb{P}(|\hat{\Sigma} - \Sigma|_\infty \geq x),$$

where  $\mathcal{W}(x) = x^{1/3} \log^{2/3}(Bp)$  for  $x > 0$ . Then (1.3.6) follows under the additional condition (1.3.5) and the conditions of Corollary 5.  $\square$

*Proof of Theorem 7.* For  $y \in \mathbb{R}$ , let

$$w(y) = [1 - \min\{0, \max(y, -1)\}]^2$$

. For any  $a > 1$  and  $b \in \mathbb{R}$ , let  $w_{a,b}(y) = w(a(y - b))$ . It is easy to observe that

$$w_{a,b+a^{-1}}(y) \leq \mathbf{1}\{y > b\} \leq w_{a,b}(y) \leq \mathbf{1}\{y > b - a^{-1}\}.$$

Then we have for any  $x > 0$ ,

$$\begin{aligned} \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} w_{a,x+a^{-1}}(\mathbf{F}_\ell) - \mathbb{P}(|\mathbf{Z}|_\infty > x) &\leq \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} \mathbf{1}\{\mathbf{F}_\ell > x\} - \mathbb{P}(|\mathbf{Z}|_\infty > x) \\ &\leq \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} w_{a,x}(\mathbf{F}_\ell) - \mathbb{P}(|\mathbf{Z}|_\infty > x). \end{aligned}$$

It leads to

$$\sup_x |\mathbf{Q}(x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \leq \sup_{x>0} \mathbb{P}(x < |\mathbf{Z}|_\infty \leq x + a^{-1}) + \Delta,$$

where

$$\Delta = \sup_{x>0} \left| \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} w_{a,x}(\mathbf{F}_\ell) - \mathbb{P}(|\mathbf{Z}|_\infty > x) \right|.$$

By Lemma 36, we have

$$\sup_{x>0} \mathbb{P}(x < |\mathbf{Z}|_\infty \leq x + a^{-1}) \leq Ca^{-1} \sqrt{\log(Bp^2)} \lesssim a^{-1} \sqrt{\log(Bp)}.$$

We choose  $a$  such that  $a^{-1} \sqrt{\log(Bp)} = o(1)$ . It suffices to prove  $\Delta \xrightarrow{\mathbb{P}} 0$ . Denote  $\Sigma_0 = \text{diag}(\Sigma)$ . Let  $x_0 = 4\sqrt{|\Sigma_0|_\infty \log(Bp^2)}$  and  $x_1 = x_0 + 1$ . Consider

$$\Delta_1 = \sup_{x>x_1} \left| \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} w_{a,x}(\mathbf{F}_\ell) - \mathbb{P}(|\mathbf{Z}|_\infty > x) \right|,$$

$$\Delta_2 = \sup_{0 < x \leq x_1} \left| \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} w_{a,x}(\mathbf{F}_\ell) - \mathbb{P}(|\mathbf{Z}|_\infty > x) \right|.$$

We firstly deal with  $\Delta_1$ . By the Bonferroni inequality,

$$\mathbb{P}(|\mathbf{Z}|_\infty > x_1) \leq Bp^2 \cdot \max_j \mathbb{P}(|Z_j| > x_1) \leq \frac{1}{Bp^2} \rightarrow 0. \quad (1.7.24)$$

Also, we can obtain

$$\mathbb{E} \sup_{x > x_1} \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} w_{a,x}(\mathbf{F}_\ell) \leq \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} \mathbb{E} w_{a,x_1}(\mathbf{F}_\ell) \leq \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} \mathbb{P}(\mathbf{F}_\ell > x_1 - a^{-1}). \quad (1.7.25)$$

For each  $1 \leq \ell \leq \mathcal{L}$ , define

$$\mathbf{F}_\ell^\star = \sqrt{\frac{M}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\widehat{F}_\ell(\theta) - F(\theta))|_\infty, |\operatorname{Im}(\widehat{F}_\ell(\theta) - F(\theta))|_\infty \}.$$

Recall that

$$\mathbf{F}^\star = \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\widehat{F}(\theta) - F(\theta))|_\infty, |\operatorname{Im}(\widehat{F}(\theta) - F(\theta))|_\infty \}. \quad (1.7.26)$$

Since  $a > 1$ , by the triangle inequality and Corollary 5, it follows that

$$\begin{aligned} \mathbb{P}(\mathbf{F}_\ell > x_1 - a^{-1}) &\leq \mathbb{P}(\mathbf{F}_\ell^\star > x_0/2) + \mathbb{P}(\mathbf{F}^\star > \sqrt{T/M} \cdot x_0/2) \\ &\leq 2\mathbb{P}(|\mathbf{Z}|_\infty > x_0/2) + d_{M,B,p} + d_{T,B,p} \\ &\leq \frac{2}{Bp^2} + d_{M,B,p} + d_{T,B,p} \rightarrow 0. \end{aligned} \quad (1.7.27)$$

By (1.7.24), (1.7.25) and (1.7.27), it follows that  $\Delta_1 \xrightarrow{\mathbb{P}} 0$ . To deal with  $\Delta_2$ , we discretize the interval  $[0, x_1]$  by evenly spaced points. Let  $K$  be large such that  $K^{-1}\Phi_2^2 \log(Bp) = o(1)$

and  $\tilde{x}_k = k \cdot x_1/K$  for  $k = 0, 1, \dots, K$ . By the triangle inequality,

$$\Delta_2 \leq \Delta_2^{(1)} + \Delta_2^{(2)} + \Delta_2^{(3)},$$

where

$$\begin{aligned} \Delta_2^{(1)} &= \max_{1 \leq k \leq K} \mathbb{P}(\tilde{x}_{k-1} < |\mathbf{Z}|_\infty \leq \tilde{x}_k), \\ \Delta_2^{(2)} &= \max_{0 \leq k \leq K} \left| \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} [\mathbb{E} w_{a, \tilde{x}_k}(\mathbf{F}_\ell) - \mathbb{P}(|\mathbf{Z}|_\infty > \tilde{x}_k)] \right|, \\ \Delta_2^{(3)} &= \max_{0 \leq k \leq K} \left| \frac{1}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}} [w_{a, \tilde{x}_k}(\mathbf{F}_\ell) - \mathbb{E} w_{a, \tilde{x}_k}(\mathbf{F}_\ell)] \right|. \end{aligned}$$

By Theorem 3 and (1.9.9),  $|\Sigma_0|_\infty \leq C\Phi_2^4$ . By Lemma 36,

$$\Delta_2^{(1)} \leq Cx_1 K^{-1} \sqrt{\log(Bp^2)} \lesssim K^{-1} \sqrt{|\Sigma_0|_\infty} \log(Bp) \rightarrow 0.$$

Before working on  $\Delta_2^{(2)}$ , we consider

$$d_{M,B}^\ell := \sup_{x>0} |\mathbb{P}(\mathbf{F}_\ell \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| = \sup_{x>0} |\mathbb{P}(\mathbf{F}_\ell > x) - \mathbb{P}(|\mathbf{Z}|_\infty > x)|.$$

By the triangle inequality, Corollary 5 and Lemma 36, for any  $b > 0$ ,

$$\sup_{x>0} |\mathbb{P}(\mathbf{F}_\ell \leq x) - \mathbb{P}(\mathbf{F}_\ell^\star \leq x)| \leq \mathbb{P}(\mathbf{F}^\star > b\sqrt{T/M}) + 2d_{M,B,p} + Cb\sqrt{\log(Bp^2)}.$$

Choosing  $b = 2\sqrt{M|\Sigma_0|_\infty \log(Bp^2)/T}$ , it follows that

$$\mathbb{P}(\mathbf{F}^\star > b\sqrt{T/M}) \leq \mathbb{P}(|\mathbf{Z}|_\infty > b\sqrt{T/M}) + d_{T,B,p} \leq \frac{1}{Bp^2} + d_{T,B,p}.$$

If  $M \log^2(Bp) \Phi_2^4/T \rightarrow 0$ , we have for any  $1 \leq \ell \leq \mathcal{L}$ ,

$$d_{M,B}^\ell \leq \sup_{x>0} |\mathbb{P}(\mathbf{F}_\ell \leq x) - \mathbb{P}(\mathbf{F}_\ell^\star \leq x)| + d_{M,B,p} \rightarrow 0.$$

Now we deal with  $\Delta_2^{(2)}$ . On the one hand,

$$\mathbb{E} w_{a,\tilde{x}_k}(\mathbf{F}_\ell) \geq \mathbb{P}(\mathbf{F}_\ell > \tilde{x}_k) \geq \mathbb{P}(|\mathbf{Z}|_\infty > \tilde{x}_k) - d_{M,B}^\ell.$$

On the other hand,

$$\mathbb{E} w_{a,\tilde{x}_k}(\mathbf{F}_\ell) \leq \mathbb{P}(\mathbf{F}_\ell > \tilde{x}_k - a^{-1}) \leq \mathbb{P}(|\mathbf{Z}|_\infty > \tilde{x}_k - a^{-1}) + d_{M,B}^\ell,$$

where

$$\begin{aligned} \mathbb{P}(|\mathbf{Z}|_\infty > \tilde{x}_k - a^{-1}) &\leq \mathbb{P}(|\mathbf{Z}|_\infty > \tilde{x}_k) + \mathbb{P}(\tilde{x}_k - a^{-1} < |\mathbf{Z}|_\infty \leq \tilde{x}_k) \\ &\leq \mathbb{P}(|\mathbf{Z}|_\infty > \tilde{x}_k) + Ca^{-1} \sqrt{\log(Bp^2)}. \end{aligned}$$

Hence it follows that

$$\Delta_2^{(2)} \leq \max_{\ell} d_{M,B,p}^\ell + Ca^{-1} \sqrt{\log(Bp^2)} \rightarrow 0.$$

It remains to consider  $\Delta_2^{(3)}$ . We are to show  $\Delta_2^{(3)} \xrightarrow{\mathbb{P}} 0$  by verifying  $\mathbb{E}(\Delta_2^{(3)})^2 \rightarrow 0$ . Define

$$\mathcal{U} = \max_{0 \leq k \leq K} \max_{1 \leq \ell, \ell' \leq \mathcal{L}} \left| \text{Cov}(w_{a,\tilde{x}_k}(\mathbf{F}_\ell), w_{a,\tilde{x}_k}(\mathbf{F}_{\ell'})) - \text{Cov}(w_{a,\tilde{x}_k}(\mathbf{F}_\ell^\star), w_{a,\tilde{x}_k}(\mathbf{F}_{\ell'}^\star)) \right|.$$

Since  $0 \leq w(y) \leq 1$  for all  $y \in \mathbb{R}$ , by Hölder's inequality,

$$\mathcal{U} \leq 4 \max_{0 \leq k \leq K} \max_{1 \leq \ell \leq \mathcal{L}} \|w_{a,\tilde{x}_k}(\mathbf{F}_\ell) - w_{a,\tilde{x}_k}(\mathbf{F}_\ell^\star)\|_2.$$

By elementary calculations, we can obtain  $\sup_{y \in \mathbb{R}} |w'_{a,b}(y)| \leq 2a$ , which implies for any  $0 \leq k \leq K$  and  $1 \leq \ell \leq \mathcal{L}$ ,

$$\|w_{a,\tilde{x}_k}(\mathbf{F}_\ell) - w_{a,\tilde{x}_k}(\mathbf{F}_\ell^\star)\|_2 \leq 2a\|\mathbf{F}_\ell - \mathbf{F}_\ell^\star\|_2 \leq 2a\sqrt{\frac{M}{T}}\|\mathbf{F}^\star\|_2.$$

Note that

$$\sqrt{\frac{B}{T}}\|\mathbf{F}^\star\|_2 \leq \|\max_{\theta \in \Theta} |\hat{F}(\theta) - \mathbb{E}\hat{F}(\theta)|_\infty\|_2 + \max_{\theta \in \Theta} |\mathbb{E}\hat{F}(\theta) - F(\theta)|_\infty.$$

By Proposition 2,

$$\max_{\theta \in \Theta} |\mathbb{E}\hat{F}(\theta) - F(\theta)|_\infty \lesssim B^{-1}\Phi_2^2.$$

By Theorem 1,

$$\|\max_{\theta \in \Theta} |\hat{F}(\theta) - \mathbb{E}\hat{F}(\theta)|_\infty\|_2 \lesssim \sqrt{\frac{B}{T}} \log^{2\alpha+1}(Bp) \cdot \|X\|_{\psi_\alpha}^2.$$

It can be obtained that

$$\mathcal{U} \lesssim a\sqrt{\frac{M}{B^3}}\Phi_2^2 + a\sqrt{\frac{M}{T}} \log^{2\alpha+1}(Bp) \cdot \|X\|_{\psi_\alpha}^2. \quad (1.7.28)$$

Then we have

$$\begin{aligned} \mathbb{E}(\Delta_2^{(3)})^2 &\leq \sum_{k=0}^K \frac{1}{\mathcal{L}^2} \sum_{\ell=1}^{\mathcal{L}} \sum_{\ell'=1}^{\mathcal{L}} \text{Cov}(w_{a,\tilde{x}_k}(\mathbf{F}_\ell), w_{a,\tilde{x}_k}(\mathbf{F}_{\ell'})) \\ &\leq \sum_{k=0}^K \frac{1}{\mathcal{L}^2} \sum_{\ell=1}^{\mathcal{L}} \sum_{\ell'=1}^{\mathcal{L}} \text{Cov}(w_{a,\tilde{x}_k}(\mathbf{F}_\ell^\star), w_{a,\tilde{x}_k}(\mathbf{F}_{\ell'}^\star)) + (K+1)\mathcal{U}. \end{aligned} \quad (1.7.29)$$

For any  $m \geq 1$ , let  $\hat{F}_{\ell,m}(\theta)$  be the  $m$ -approximation of  $\hat{F}_\ell(\theta)$  where  $\{X_t\}$  is replaced by

$\{X_{t,m}\}$ . Define

$$\mathbf{F}_{\ell,m}^* = \sqrt{\frac{M}{B}} \max_{\theta \in \Theta} \max \{ |\operatorname{Re}(\widehat{F}_{\ell,m}(\theta) - F(\theta))|_\infty, |\operatorname{Im}(\widehat{F}_{\ell,m}(\theta) - F(\theta))|_\infty \}.$$

Note that  $\mathbf{F}_0^*$  is independent of  $\mathbf{F}_{\ell,m_\ell}^*$  for  $m_\ell = \lfloor (\ell - 1)/2 \rfloor$  and  $\ell \geq 2M - 1$ . Following (1.7.29) and by the stationarity of the process  $\{\mathbf{F}_\ell^*\}_{\ell \in \mathbb{Z}}$ , we have

$$\begin{aligned} \mathbb{E}(\Delta_2^{(3)})^2 &\leq \sum_{k=0}^K \frac{2}{\mathcal{L}} \sum_{\ell=0}^{\mathcal{L}-1} |\operatorname{Cov}(w_{a,\tilde{x}_k}(\mathbf{F}_0^*), w_{a,\tilde{x}_k}(\mathbf{F}_\ell^*))| + 2K\mathcal{U} \\ &\leq \sum_{k=0}^K \frac{2}{\mathcal{L}} \sum_{\ell=1}^{\mathcal{L}-1} |\operatorname{Cov}(w_{a,\tilde{x}_k}(\mathbf{F}_0^*), w_{a,\tilde{x}_k}(\mathbf{F}_\ell^*) - w_{a,\tilde{x}_k}(\mathbf{F}_{\ell,m_\ell}^*))| + \frac{4(K+1)M}{\mathcal{L}} + 2K\mathcal{U} \\ &\leq \frac{2}{\mathcal{L}} \sum_{k=0}^K \sum_{\ell=1}^{\mathcal{L}} \min \{1, \|w_{a,\tilde{x}_k}(\mathbf{F}_\ell^*) - w_{a,\tilde{x}_k}(\mathbf{F}_{\ell,m_\ell}^*)\|_2\} + \frac{4(K+1)M}{\mathcal{L}} + 2K\mathcal{U}. \end{aligned} \tag{1.7.30}$$

By similar arguments in deriving (1.7.28), we have

$$\|w_{a,\tilde{x}_k}(\mathbf{F}_\ell^*) - w_{a,\tilde{x}_k}(\mathbf{F}_{\ell,m_\ell}^*)\|_2 \leq 2a\sqrt{\frac{M}{B}} \|\max_{\theta \in \Theta} |\widehat{F}_\ell(\theta) - \widehat{F}_{\ell,m_\ell}(\theta)|_\infty\|_2.$$

By Lemma 24, we have

$$\|\max_{\theta \in \Theta} |\widehat{F}_\ell(\theta) - \widehat{F}_{\ell,m_\ell}(\theta)|_\infty\|_2 \lesssim \rho^{\ell/2} \sqrt{\frac{B}{M}} \log^{2\alpha+1}(Bp) \|X.\|_{\psi_\alpha}^2.$$

As a result, for any  $0 \leq k \leq K$ ,

$$\|w_{a,\tilde{x}_k}(\mathbf{F}_\ell^*) - w_{a,\tilde{x}_k}(\mathbf{F}_{\ell,m_\ell}^*)\|_2 \lesssim a\rho^{\ell/2} \log^{2\alpha+1}(Bp) \|X.\|_{\psi_\alpha}^2.$$

By (1.7.30) and  $\mathcal{L} \asymp T$ ,  $\mathbb{E}(\Delta_2^{(3)})^2 \rightarrow 0$  follows if we have (i)  $\mathcal{K}aT^{-1} \log^{2\alpha+1}(Bp) \|X.\|_{\psi_\alpha}^2 \rightarrow 0$ , (ii)  $KM/T \rightarrow 0$ , and (iii)  $K\mathcal{U} \rightarrow 0$ , where  $\mathcal{U}$  is bounded in (1.7.28). We recall that the

following conditions are additionally required to control other terms in the previous proof: (iv)  $a^{-1}\sqrt{\log(Bp)} \rightarrow 0$ , (v)  $\mathcal{K}^{-1}\Phi_2^2 \log(Bp) \rightarrow 0$ , (vi)  $M \log^2(Bp)\Phi_2^4/T \rightarrow 0$ . In view of  $\Phi_2 \lesssim \|X\|_{\psi_\alpha}$ , elementary manipulation indicates that these conditions can be reduced to (1.3.7) and (1.3.8), under which the existence of satisfactory  $a$  and  $\mathcal{K}$  is ensured.  $\square$

### 1.7.3 Proofs of Results in Section 1.5

*Proof of Theorem 8.* It is an immediate result of Lemma 30 with  $d = (B/2) + 1$  and Lemma 31 using the Bonferroni technique.  $\square$

*Proof of Proposition 9.* The result can be proved in a manner similar to Proposition 2, in view of  $\Delta_{0,2,i} \leq \Phi_{2,0}$  and  $\Delta_{B,2,j} \leq B^{-\alpha}\Phi_{2,\alpha}$ .  $\square$

*Proof of Theorem 10.* Recall the definitions of  $\widehat{F}(\theta)$ ,  $\widehat{F}_m(\theta)$ ,  $\widehat{H}_{r,m}$  and  $\widehat{S}_m(\theta)$  in the proof of Theorem 3. By Lemma 9 in [111],

$$\|\mathbb{E}_0(\widehat{F}_{ij}(\theta)) - \mathbb{E}_0(\widehat{F}_{ij,m}(\theta))\|_2 \lesssim \sqrt{\frac{B}{T}} m^{-\alpha} \Phi_{4,0} \Phi_{4,\alpha}. \quad (1.7.31)$$

Similarly as (1.9.13), we can obtain

$$\|\mathbb{E}_0(\widehat{F}_{ij,m}(\theta)) - \mathbb{E}_0(\widehat{S}_{ij,m}(\theta))\|_2 \lesssim \sqrt{\frac{m}{T}} \Phi_{4,0}^2. \quad (1.7.32)$$

Choosing  $m \asymp B^{1/(2\alpha+1)}$ , by (1.7.31) and (1.7.32), under  $\mathcal{E}_B \rightarrow 0$ ,

$$\sqrt{\frac{T}{B}} \|\mathbb{E}_0(\widehat{F}_{ij}(\theta)) - \mathbb{E}_0(\widehat{S}_{ij,m}(\theta))\|_2 \lesssim m^{-\alpha} \Phi_{4,0} \Phi_{4,\alpha} + \sqrt{\frac{m}{B}} \Phi_{4,0}^2 \rightarrow 0. \quad (1.7.33)$$

Then Theorem 10 follows similarly as Theorem 3.  $\square$

*Proof of Theorem 11.* By Theorem 15, to have  $d_{T,B} \rightarrow 0$ , we need

$$\mathcal{W}(\mathcal{E}_{B,m}) \rightarrow 0, \quad (1.7.34)$$

$$h_{T,B} \rightarrow 0, \quad (1.7.35)$$

and for some  $y > 0$ ,

$$f_1^*(y) + f_2^*(y) + y\sqrt{\log(Bp)} \rightarrow 0. \quad (1.7.36)$$

Firstly, (1.7.34) requires  $m \gg (\Phi_{2,0}^3 \Phi_{2,\alpha} (\log(Bp))^2)^{1/\alpha}$ ,  $m \ll B(\log(Bp))^{-4} \Phi_{4,0}^{-8}$  and (1.5.6) hold. Next, (1.7.35) requires (1.5.7). Moreover, for  $\alpha > 1/2 - 2/q$ , (1.7.36) needs  $m \gg \max(L_2, (\Phi_{4,\alpha}^2 \log(Bp))^{1/\alpha})$  and  $m \ll \min(R_2, B(\log(Bp))^{-2} \Phi_{4,\alpha}^{-4})$ . Observe that  $\max((\Phi_{2,0}^3 \Phi_{2,\alpha} (\log(Bp))^2)^{1/\alpha}, (\Phi_{4,\alpha}^2 \log(Bp))^{1/\alpha}) \lesssim L_1$  and

$$R_1 \lesssim \min(B(\log(Bp))^{-4} \Phi_{4,0}^{-8}, B(\log(Bp))^{-2} \Phi_{4,\alpha}^{-4}).$$

As a consequence, if

$$\max(L_1, L_2) = o(1) \min(R_1, R_2), \quad (1.7.37)$$

we can always choose  $m$  such that (1.5.9) holds.

For  $\alpha < 1/2 - 2/q$ , by similar arguments, if (1.5.10) is further assumed, (1.7.36) holds and (1.5.9) follows.  $\square$

Theorem 15 characterizes the error rate for the result in Theorem 11.

**Theorem 15.** Assume  $\| |X| \|_{\infty, q, \alpha} < \infty$ , for some  $q > 4$  and  $\alpha > 0$ , and assume

$$\mathcal{E}_{B,m} := m^{-\alpha} \Phi_{2,0}^3 \Phi_{2,\alpha} + B^{-1} \Phi_{4,0}^2 \Psi_{4,2}^2 + \sqrt{\frac{m}{B}} \Phi_{4,0}^4 = o(1).$$

Let Assumption 2 and Assumption 3 be satisfied. Recall that

$$d_{T,B} = \sup_{x \in \mathbb{R}} |\mathbb{P}(\mathbf{F} \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)|.$$

Then for every  $y > 0$ ,

$$d_{T,B} \lesssim f_1^*(y) + f_2^*(y) + y\sqrt{\log(Bp)} + h_{T,B} + \mathcal{W}(\mathcal{E}_{B,m}). \quad (1.7.38)$$

where

$$f_1^*(y) = y^{-q/2}(\log(Bp))^q \| |X|_\infty \|_{q,\alpha}^q F_{T,m} + p^2 B \exp\left(-\frac{y^2 m^{2\alpha}}{C_\alpha \Phi_{4,\alpha}^4}\right),$$

with  $F_{T,B,m} = (B/T)^{q/4-1} m^{-\alpha q/2}$  (resp.  $(B/T)^{q/4-1} m^{-\alpha q/2} + T^{-\alpha q/2}$ ) if  $\alpha > 1/2 - 2/q$  (resp.  $\alpha < 1/2 - 2/q$ ),

$$f_2^*(y) = y^{-q/2}(\log(Bp))^q T^{-q/4+1} B^{-1} m^{q/4} \| |X|_\infty \|_{q,\alpha}^q + p^2 B \exp\left(-\frac{y^2 B}{C \Phi_{4,\alpha}^4 m}\right),$$

$h_{T,B} = h_{T,B}^{(1)} + h_{T,B}^{(2)}$  with

$$\begin{aligned} h_{T,B}^{(1)} &= \left( \Phi_{8,0}^{12} (\log(Bp))^2 \| |X|_\infty \|_{q,\alpha}^4 (\log(Tp))^7 B/T \right)^{\frac{1}{6}}, \\ h_{T,B}^{(2)} &= \left( \Phi_{8,0}^{12} (\log(Bp))^2 \| |X|_\infty \|_{q,\alpha}^4 (\log(Tp))^3 (B/T)^{1-4/q} \right)^{\frac{1}{3}}, \end{aligned}$$

and  $\mathcal{W}(x) = x^{1/3}(1 \vee \log(Bp/x))^{2/3}$  for  $x > 0$ .

*Proof of Theorem 15.* Recall the definitions of  $\widehat{F}_m(\theta)$  and  $\widehat{S}_m(\theta)$  in Section 1.9.2. Note that

$$\begin{aligned} d_{T,B} &\leq \sup_{x \in \mathbb{R}} \left| \mathbb{P}\left(\sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{F}(\theta))|_\infty \geq x\right) - \mathbb{P}\left(\sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{S}_m(\theta))|_\infty \geq x\right) \right| \\ &\quad + \sup_{x \in \mathbb{R}} \left| \mathbb{P}\left(\sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{S}_m(\theta))|_\infty \geq x\right) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x) \right|. \end{aligned}$$

By the triangle inequality, for every  $y > 0$ , we have

$$\begin{aligned}
& \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\hat{F}(\theta))|_\infty \geq x \right) - \mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\hat{S}_m(\theta))|_\infty \geq x \right) \right| \\
& \leq \mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\hat{F}(\theta)) - \mathbb{E}_0(\hat{S}_m(\theta))|_\infty \geq y \right) \\
& \quad + \sup_{x \in \mathbb{R}} \mathbb{P} \left( \left| \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\hat{S}_m(\theta))|_\infty - x \right| \leq y \right).
\end{aligned}$$

By Lemma 29, for every  $y > 0$ , we have

$$\mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\hat{F}(\theta)) - \mathbb{E}_0(\hat{F}_m(\theta))|_\infty \geq y \right) \lesssim f_1^*(y). \quad (1.7.39)$$

And by Lemma 32, for every  $y > 0$ , we have

$$\mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\hat{F}_m(\theta)) - \mathbb{E}_0(\hat{S}_m(\theta))|_\infty \geq y \right) \lesssim f_2^*(y). \quad (1.7.40)$$

It can be easily obtained that

$$\mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\hat{F}(\theta)) - \mathbb{E}_0(\hat{S}_m(\theta))|_\infty \geq y \right) \lesssim f_1^*(y) + f_2^*(y). \quad (1.7.41)$$

Denote  $\theta_k = 2\pi k/B$  for  $k \in \mathbb{Z}$ . Define

$$\mathbf{H}_m = (\text{Re}(\hat{H}_m(\theta_0)), \text{Im}(\hat{H}_m(\theta_0)), \dots, \text{Re}(\hat{H}_m(\theta_{B/2})), \text{Im}(\hat{H}_m(\theta_{B/2}))),$$

where

$$\hat{H}_m(\theta) = (\hat{H}_{ij,m}(\theta))_{i,j=1}^p = \frac{1}{2\pi B} \left( \sum_{t=1}^{B-m} X_{t,m} e^{-it\theta} \right) \left( \sum_{t=1}^{B-m} X_{t,m}^\top e^{it\theta} \right).$$

Let  $\mathbf{vec}_0(\mathbf{H}_m) \in \mathbb{R}^{Bp^2 - (B/2-1)p}$  be the vectorization of  $\mathbf{H}_m$  where the constantly zero entries, i.e.,  $\text{Im}(\hat{H}_{ii,m}(\theta))$  for  $1 \leq i \leq p$  and all the entries in  $\text{Im}(\hat{H}_m(0))$  and  $\text{Im}(\hat{H}_m(\pi))$ , are excluded. Denote  $\mathbf{\Sigma}_m = \text{Cov}(\mathbf{vec}_0(\mathbf{H}_m))$  and  $\mathbf{Z}_m \sim N(0, \mathbf{\Sigma}_m)$ . For  $1 \leq i, j \leq p$  and

$0 \leq \theta \leq \pi$ , denote

$$\begin{aligned}\mathcal{M}_{ij,r,m}(\theta) &= \text{Re}(\widehat{H}_{ij,r,m}(\theta) - \mathbb{E}\widehat{H}_{ij,r,m}(\theta)), \\ \tilde{\mathcal{M}}_{ij,r,m}(\theta) &= \text{Im}(\widehat{H}_{ij,r,m}(\theta) - \mathbb{E}\widehat{H}_{ij,r,m}(\theta)).\end{aligned}$$

Now we deal with  $\mathcal{M}_{ij,r,m}(\theta)$  and the same results will apply for  $\tilde{\mathcal{M}}_{ij,r,m}(\theta)$ . By Lemma 23, under  $\mathcal{E}_{B,m} := m^{-\alpha}\Phi_{2,0}^3\Phi_{2,\alpha} + B^{-1}\Phi_{4,0}^2\Psi_{4,2}^2 + \sqrt{\frac{m}{B}}\Phi_{4,0}^4 = o(1)$  as well as Assumption 2 and Assumption 3, it holds that

$$\frac{1}{n} \sum_{r=1}^n \mathbb{E} \mathcal{M}_{ij,r,m}(\theta)^2 \geq \frac{1}{2} \min\{c_1^2, c_2\}.$$

Applying Lemma 27 (iv), we have  $\max_{1 \leq r \leq n} \mathbb{E} |\mathcal{M}_{ij,r,m}(\theta)|^{2+k} \lesssim \Phi_{8,0}^{6k}$  for  $k = 1, 2$ , and  $\|\max_{\theta \in \Theta} |\mathcal{M}_{ij,r,m}(\theta)|_\infty\|_{q/2} \lesssim \log(Bp) \| |X_\cdot|_\infty \|_{q,\alpha}^2$  for all  $1 \leq i, j \leq p$  and  $\theta \in \Theta$ . Then it follows from Lemma 35 that

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{S}_m(\theta))|_\infty \geq x \right) - \mathbb{P}(|\mathbf{Z}_m|_\infty \leq x) \right| \lesssim h_{T,B}. \quad (1.7.42)$$

We shall find the upper bound of  $\sup_{x \in \mathbb{R}} |\mathbb{P}(|\mathbf{Z}_m|_\infty \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)|$ . By Lemma 23, we have

$$|\Sigma_m - \Sigma|_\infty \lesssim m^{-\alpha}\Phi_{2,0}^3\Phi_{2,\alpha} + B^{-1}\Phi_{4,0}^2\Psi_{4,2}^2 + \sqrt{\frac{m}{B}}\Phi_{4,0}^4 =: \mathcal{E}_{B,m}.$$

By Theorem 2 in [23],

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(|\mathbf{Z}_m|_\infty \leq x) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x)| \lesssim \mathcal{W}(\mathcal{E}_{B,m}), \quad (1.7.43)$$

where  $\mathcal{W}(x) = x^{1/3}(1 \vee \log(Bp/x))^{2/3}$  for  $x > 0$ . By (1.7.42) and (1.7.43), we have

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{S}_m(\theta))|_\infty \geq x \right) - \mathbb{P}(|\mathbf{Z}|_\infty \leq x) \right| \lesssim h_{T,B} + \mathcal{W}(\mathcal{E}_{B,m}). \quad (1.7.44)$$

By Lemma 36, for every  $y > 0$ ,

$$\sup_{x \in \mathbb{R}} \mathbb{P}(|\mathbf{Z}|_\infty - x| \leq y) \lesssim y \sqrt{\log(Bp)}, \quad (1.7.45)$$

which together with (1.7.44) will imply

$$\sup_{x \in \mathbb{R}} \mathbb{P} \left( \left| \sqrt{\frac{T}{B}} \max_{\theta \in \Theta} |\mathbb{E}_0(\widehat{S}_m(\theta))|_\infty - x \right| \leq y \right) \lesssim y \sqrt{\log(Bp)} + h_{T,B} + \mathcal{W}(\mathcal{E}_{B,m}). \quad (1.7.46)$$

The proof is completed by combining the results in (1.7.41), (1.7.44) and (1.7.46).  $\square$

*Proof of Corollary 12.* By choosing  $y = C\sqrt{T/B}(B^{-\alpha}\Phi_{2,0}\Phi_{2,\alpha} + B^{-1}\Phi_{2,0}\Psi_{2,1})$  for a sufficiently large  $C$ , Corollary 12 can be proved in a manner similar to Corollary 5.  $\square$

*Proof of Theorem 13.* By similarly arguments in the proof of Theorem 6, (1.5.13) can be obtained under the condition  $R_{T,B}\Phi_{2,0}^2(\log(Bp))^2 \rightarrow 0$  and the conditions of Corollary 12.  $\square$

*Proof of Theorem 14.* Recall the definitions of  $\mathcal{V}$  and  $R(\mathcal{V})$  in Theorem 14 and note that  $\mathcal{L} \asymp T$ . By similar arguments in the proof of Theorem 7 and elementary calculations, we have

$$\sum_{\ell=1}^{\mathcal{L}} \min \{1, \|w_{a,\tilde{x}_k}(\mathbf{F}_\ell^\star) - w_{a,\tilde{x}_k}(\mathbf{F}_{\ell,\ell-1}^\star)\|_2\} \lesssim a \cdot R(\mathcal{V}).$$

Then  $\mathbb{E}(\Delta_2^{(3)})^2 \rightarrow 0$  follows if  $aKR(\mathcal{V})/\mathcal{L} \rightarrow 0$  and  $K\mathcal{U} \rightarrow 0$ . Recall that  $a^{-1}\sqrt{\log(Bp)} = o(1)$  and  $K^{-1}\Phi_{2,0}^2 \log(Bp) = o(1)$  should also be required. Conditions (i)-(iii) ensure the existence of satisfactory  $a$  and  $K$ .  $\square$

## 1.8 Conclusion

Spectral density is a pivotal component in the analysis of time series data. While a systematic asymptotic theory has been developed for spectral estimates in low-dimensional cases, the distributional theory for high-dimensional time series remains underexplored. Our work endeavors to fill this gap by formulating an inference theory specifically tailored for the high-dimensional spectral density matrix. Our primary focus is on establishing asymptotic normality for point-wise inference and a Gaussian approximation for simultaneous inference. Our distribution theory serves as a powerful tool for addressing a diverse range of challenges in time series inference. Additionally, we introduce two distinct resampling methods designed to facilitate practical implementation of high-dimensional spectral inference. Importantly, we provide theoretical justification for the validity of these methods, enhancing their reliability in real-world applications. We expect that our framework, distribution theory of spectral estimate and implementation methods will be useful in many high-dimensional inference problems and practical applications that involve time series data.

## 1.9 Appendix

In the Appendix, we provide the lemmas that are used in Section 1.7 and their proofs .

### 1.9.1 Useful Lemmas for Expectation and Covariance of Spectral Estimates

**Lemma 16.** *Denote  $\Phi_2 = \max_{1 \leq j \leq p} \|X_{.j}\|_2$ . For  $m \geq 0$ , define*

$$X_{t,m} = \mathbb{E}(X_t | \varepsilon_{t-m}, \varepsilon_{t-m+1}, \dots, \varepsilon_t)$$

and

$$\widehat{G}_{ij,m}(\theta) = \frac{1}{2\pi B} \sum_{t,s=1}^B X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}. \quad (1.9.1)$$

Then for any  $1 \leq i, j \leq p$  and  $0 \leq \theta \leq \pi$ ,

$$|\mathbb{E}\widehat{G}_{ij,m}(\theta) - F_{ij}(\theta)| \lesssim (\rho^m + B^{-1})\Phi_2^2. \quad (1.9.2)$$

*Proof of Lemma 16.* By the definitions of  $\widehat{G}_{ij}(\theta)$  and  $\widehat{G}_{ij,m}(\theta)$ , we have

$$\begin{aligned} (2\pi B)\mathbb{E}\widehat{G}_{ij}(\theta) &= \mathbb{E}\left(\sum_{t=1}^B X_{ti} e^{-\iota t\theta}\right)\left(\sum_{s=1}^B X_{sj} e^{\iota s\theta}\right), \\ (2\pi B)\mathbb{E}\widehat{G}_{ij,m}(\theta) &= \mathbb{E}\left(\sum_{t=1}^B X_{ti,m} e^{-\iota t\theta}\right)\left(\sum_{s=1}^B X_{sj,m} e^{\iota s\theta}\right). \end{aligned}$$

By Hölder's inequality,

$$(2\pi B)|\mathbb{E}\widehat{G}_{ij}(\theta) - \mathbb{E}\widehat{G}_{ij,m}(\theta)| \leq \text{I} + \text{II}, \quad (1.9.3)$$

where

$$\begin{aligned} \text{I} &= \left\| \sum_{t=1}^B (X_{ti} - X_{ti,m}) e^{-\iota t\theta} \right\|_2 \left\| \sum_{s=1}^B X_{sj} e^{\iota s\theta} \right\|_2, \\ \text{II} &= \left\| \sum_{t=1}^B X_{ti,m} e^{-\iota t\theta} \right\|_2 \left\| \sum_{s=1}^B (X_{sj} - X_{sj,m}) e^{\iota s\theta} \right\|_2. \end{aligned}$$

By Lemma 1 in [59], we have  $\left\| \sum_{s=1}^B X_{sj} e^{\iota s\theta} \right\|_2 \lesssim \sqrt{B}\Delta_{0,2,j}$ ,  $\left\| \sum_{t=1}^B X_{ti,m} e^{-\iota t\theta} \right\|_2 \lesssim \sqrt{B}\Delta_{0,2,i}$ ,  $\left\| \sum_{t=1}^B (X_{ti} - X_{ti,m}) e^{-\iota t\theta} \right\|_2 \lesssim \sqrt{B}\Delta_{m+1,2,i}$  and  $\left\| \sum_{s=1}^B (X_{sj} - X_{sj,m}) e^{\iota s\theta} \right\|_2 \lesssim \sqrt{B}\Delta_{m+1,2,j}$ . Since  $\Delta_{0,2,i} \leq \|X_{\cdot i}\|_2$  and  $\Delta_{m+1,2,i} \leq \Delta_{m,2,i} \leq \rho^m \|X_{\cdot i}\|_2$ , it follows that

$$\text{I} \lesssim B\rho^m \|X_{\cdot i}\|_2 \|X_{\cdot j}\|_2, \quad \text{II} \lesssim B\rho^m \|X_{\cdot i}\|_2 \|X_{\cdot j}\|_2. \quad (1.9.4)$$

By the proof of Proposition 2, we have

$$|\mathbb{E}\widehat{G}_{ij}(\theta) - F_{ij}(\theta)| \lesssim B^{-1}\|X_{\cdot i}\|_2\|X_{\cdot j}\|_2. \quad (1.9.5)$$

By (1.9.3), (1.9.4), (1.9.5) and the triangle inequality, (1.9.2) follows by noting  $\Phi_2 = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_2$ .  $\square$

**Lemma 17.** For  $1 \leq i, j, h, l \leq p$  and  $k_1, k_2, k_3 \in \mathbb{Z}$ , denote the 4-th joint cumulant as

$$\gamma_{i,j,h,l}(k_1, k_2, k_3) = \text{Cum}(X_{0i}, X_{k_1j}, X_{k_2h}, X_{k_3l}).$$

Denote  $\Phi_4 = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_4$ . Then for some constant  $C_\rho > 0$ ,

$$\sum_{k_1, k_2, k_3 \in \mathbb{Z}} |\gamma_{i,j,h,l}(k_1, k_2, k_3)| \leq C_\rho \Phi_4^4.$$

*Proof of Lemma 17.* We can apply the similar arguments of proving Theorem 21 in [110] without extra difficulty.  $\square$

**Lemma 18.** Denote  $\Phi_4 = \max_{1 \leq j \leq p} \|X_{\cdot j}\|_4$ . Take any  $\theta, \theta' \in \Theta$  and any  $1 \leq i, j, h, l \leq p$ . Consider  $\widehat{G}_{ij,m}(\theta)$  and  $\widehat{G}_{hl,m}(\theta)$  as given by (1.9.1). Denote  $\mathcal{E}_{B,m}^{(1)} := (\rho^m + B^{-1})\Phi_4^4$ .

(i) If  $\theta \in (0, \pi)$ ,

$$\begin{aligned} |\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}(\theta)) - \mathbb{E}\widehat{G}_{ij,m}(\theta)\mathbb{E}\widehat{G}_{hl,m}(\theta) - F_{il}(\theta)F_{hj}(\theta)| &\lesssim \mathcal{E}_{B,m}^{(1)}, \\ |\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}^H(\theta)) - \mathbb{E}\widehat{G}_{ij,m}(\theta)\mathbb{E}\widehat{G}_{hl,m}^H(\theta) - F_{ih}(\theta)F_{lj}(\theta)| &\lesssim \mathcal{E}_{B,m}^{(1)}. \end{aligned}$$

(ii) If  $\theta = 0$  or  $\pi$ ,

$$|\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}(\theta)) - \mathbb{E}\widehat{G}_{ij,m}(\theta)\mathbb{E}\widehat{G}_{hl,m}(\theta) - F_{ih}(\theta)F_{lj}(\theta) - F_{il}(\theta)F_{hj}(\theta)| \lesssim \mathcal{E}_{B,m}^{(1)}.$$

(iii) If  $\theta \neq \theta'$ ,

$$\begin{aligned} |\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}(\theta')) - \mathbb{E}\widehat{G}_{ij,m}(\theta)\mathbb{E}\widehat{G}_{hl,m}(\theta')| &\lesssim \mathcal{E}_{B,m}^{(1)}, \\ |\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}^H(\theta')) - \mathbb{E}\widehat{G}_{ij,m}(\theta)\mathbb{E}\widehat{G}_{hl,m}^H(\theta')| &\lesssim \mathcal{E}_{B,m}^{(1)}, \end{aligned}$$

*Proof.* For any  $\theta, \theta' \in \Theta$  and any  $1 \leq i, j, h, l \leq p$ , we can write

$$\mathbb{E}\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}(\theta') = \frac{1}{4\pi B^2} \sum_{t,s,u,v=1}^B \mathbb{E}(X_{ti,m}X_{sj,m}X_{uh,m}X_{vl,m}) \times e^{-\iota[(t-s)\theta+(u-v)\theta']}. \quad (1.9.6)$$

As  $\mathbb{E}X_{ti,m} = 0$  for any  $t \in \mathbb{Z}$  and  $1 \leq i \leq p$ , by the decomposition of the 4-th joint cumulant, we have

$$\begin{aligned} &\mathbb{E}(X_{ti,m}X_{sj,m}X_{uh,m}X_{vl,m}) \\ &= \mathbb{E}(X_{ti,m}X_{sj,m})\mathbb{E}(X_{uh,m}X_{vl,m}) \\ &\quad + \mathbb{E}(X_{ti,m}X_{uh,m})\mathbb{E}(X_{sj,m}X_{vl,m}) \\ &\quad + \mathbb{E}(X_{ti,m}X_{vl,m})\mathbb{E}(X_{sj,m}X_{uh,m}) \\ &\quad + \text{Cum}(X_{ti,m}, X_{sj,m}, X_{uh,m}, X_{vl,m}). \end{aligned}$$

Denote

$$\begin{aligned} \text{I} &= \frac{1}{4\pi^2 B^2} \sum_{t,s,u,v=1}^B \mathbb{E}(X_{ti,m}X_{sj,m})\mathbb{E}(X_{uh,m}X_{vl,m}) \times e^{-\iota[(t-s)\theta+(u-v)\theta']}, \\ \text{II} &= \frac{1}{4\pi^2 B^2} \sum_{t,s,u,v=1}^B \mathbb{E}(X_{ti,m}X_{uh,m})\mathbb{E}(X_{sj,m}X_{vl,m}) \times e^{-\iota[(t-s)\theta+(u-v)\theta']}, \\ \text{III} &= \frac{1}{4\pi^2 B^2} \sum_{t,s,u,v=1}^B \mathbb{E}(X_{ti,m}X_{vl,m})\mathbb{E}(X_{sj,m}X_{uh,m}) \times e^{-\iota[(t-s)\theta+(u-v)\theta']}, \\ \text{IV} &= \frac{1}{4\pi^2 B^2} \sum_{t,s,u,v=1}^B \text{Cum}(X_{ti,m}, X_{sj,m}, X_{uh,m}, X_{vl,m}) \times e^{-\iota[(t-s)\theta+(u-v)\theta']}. \end{aligned}$$

It can be easily seen that  $I = (\mathbb{E}\widehat{G}_{ij,m}(\theta))(\mathbb{E}\widehat{G}_{hl,m}(\theta'))$ . Applying Lemma 17 to the process after  $m$ -approximation, we have  $|IV| \lesssim B^{-1}\Phi_4^4$ .

Next we shall consider II and III. For  $\theta = \theta' = 0$  or  $\pi$ ,  $\widehat{G}_{ij,m}(\theta)$  and  $\widehat{G}_{hl,m}(\theta)$  are both real. It follows that

$$\begin{aligned} \text{II} &= \frac{1}{4\pi^2 B^2} \sum_{t,s,u,v=1}^B \mathbb{E}(X_{ti,m}X_{uh,m})\mathbb{E}(X_{sj,m}X_{vl,m}) \times e^{-\iota[(t-s)\theta+(u-v)\theta]} \\ &= \frac{1}{4\pi^2 B^2} \sum_{t,s,u,v=1}^B \mathbb{E}(X_{ti,m}X_{uh,m})\mathbb{E}(X_{sj,m}X_{vl,m}) \times e^{-\iota[(t-s)\theta-(u-v)\theta]} \\ &= \mathbb{E}\widehat{G}_{ih,m}(\theta)\mathbb{E}\widehat{G}_{lj,m}(\theta). \end{aligned}$$

Observe that

$$\text{II} = \frac{1}{4\pi^2 B^2} \left( \sum_{t,u=1}^B \Gamma_{ih,m}(t-u)e^{-\iota(t\theta+u\theta')} \right) \times \left( \sum_{s,v=1}^B \Gamma_{jl,m}(s-v)e^{\iota(s\theta+v\theta')} \right),$$

where  $\Gamma_{ij,m}(k) = \text{Cov}(X_{0i,m}, X_{kj,m})$  for any  $1 \leq i, j \leq p$  and  $k \in \mathbb{Z}$ . We have

$$\begin{aligned} & \sum_{t,u=1}^B \Gamma_{ih,m}(t-u)e^{-\iota(t\theta+u\theta')} \\ &= \sum_{k=-B+1}^{B-1} \Gamma_{ih,m}(k)e^{-\iota k\theta} \sum_{u=1}^{B-|k|} e^{-\iota u(\theta+\theta')}. \end{aligned} \tag{1.9.7}$$

If  $\theta + \theta' \neq 0$  or  $2\pi$ , by the fact that  $\sum_{u=1}^{B-|k|} e^{-\iota u(\theta+\theta')} = 0$ , we have  $|\sum_{u=1}^{B-|k|} e^{-\iota u(\theta+\theta')}| \leq |k|$ .

By similar arguments in (1.7.9) and (1.7.10), we have

$$\sum_{k=1}^{B-1} k|\Gamma_{ih,m}(k)| \leq \sum_{k=1}^{B-1} \sum_{r=0}^{\infty} k\delta_{r,2,i}\delta_{r+k,2,h} \lesssim \Phi_2^2,$$

Dealing with  $\sum_{s,v=1}^B \Gamma_{jl,m}(s-v)e^{\iota(s\theta+v\theta')}$  in the same way, we can obtain  $|\text{II}| \lesssim B^{-2}\Phi_2^4$  if  $\theta + \theta' \neq 0$  or  $2\pi$ .

As for III, if  $\theta = \theta'$ , it can be observed that  $\text{III} = \mathbb{E}\widehat{G}_{il,m}(\theta)(\mathbb{E}\widehat{G}_{hj,m}(\theta))$ . If  $\theta \neq \theta'$ , we can write

$$\text{III} = \frac{1}{4\pi^2 B^2} \left( \sum_{t,v=1}^B \Gamma_{il,m}(t-v) e^{-\iota(t\theta-v\theta')} \right) \times \left( \sum_{s,u=1}^B \Gamma_{jh,m}(s-u) e^{\iota(s\theta-v\theta')} \right),$$

and apply similar arguments as (1.9.7) to get the same bound.

Combining I, II, III and IV and noting that  $\Phi_2 \leq \Phi_4$ , we have

$$|\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}(\theta')) - \mathbb{E}\widehat{G}_{ij,m}(\theta)\mathbb{E}\widehat{G}_{hl,m}(\theta') - V_m| \lesssim B^{-1}\Phi_4^4, \quad (1.9.8)$$

where

$$V_m = \mathbb{E}\widehat{G}_{il,m}(\theta)\mathbb{E}\widehat{G}_{hj,m}(\theta)$$

if  $\theta = \theta' \in (0, \pi)$ ,

$$V_m = \mathbb{E}\widehat{G}_{ih,m}(\theta)\mathbb{E}\widehat{G}_{lj,m}(\theta) + \mathbb{E}\widehat{G}_{il,m}(\theta)\mathbb{E}\widehat{G}_{hj,m}(\theta)$$

if  $\theta = \theta' = 0$  or  $\pi$ , and  $V_m = 0$  if  $\theta \neq \theta'$ . Let  $V = F_{il}(\theta)F_{hj}(\theta)$  if  $\theta = \theta' \in (0, \pi)$ ,  $V = F_{ih}(\theta)F_{lj,m}(\theta) + F_{il,m}(\theta)F_{hj,m}(\theta)$  if  $\theta = \theta' = 0$  or  $\pi$ , and  $V = 0$  if  $\theta \neq \theta'$ . We now account for the bias between  $V_m$  and  $V$ . By (1.7.9), for any  $1 \leq i, j \leq p$  and  $\theta$ ,

$$|F_{ij}(\theta)| \leq 2 \sum_{k=0}^{\infty} |\Gamma_{ij}(k)| \leq 2 \sum_{k=0}^{\infty} \sum_{r=0}^{\infty} \delta_{r,2,i} \delta_{r+k,2,j} \leq 2\Phi_2^2. \quad (1.9.9)$$

By the triangle inequality, Lemma 16 and (1.9.9),

$$|V_m - V| \lesssim (\rho^m + B^{-1})\Phi_2^4. \quad (1.9.10)$$

In view of (1.9.8), (1.9.10) and  $\Phi_2 \leq \Phi_4$ , the desired results for

$$\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}(\theta')) - \mathbb{E}\widehat{G}_{ij,m}(\theta)\mathbb{E}\widehat{G}_{hl,m}(\theta')$$

in (i), (ii) and (iii) can be reached.

When we consider  $\mathbb{E}(\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}^H(\theta'))$ , (1.9.6) becomes

$$\mathbb{E}\widehat{G}_{ij,m}(\theta)\widehat{G}_{hl,m}^H(\theta') = \frac{1}{4\pi B^2} \sum_{t,s,u,v=1}^B \mathbb{E}(X_{ti,m}X_{sj,m}X_{uh,m}X_{vl,m}) \times e^{-[\iota(t-s)\theta - (u-v)\theta']}.$$

We can apply the same manipulation to deal with  $\mathbb{E}(X_{ti,m}X_{sj,m}X_{uh,m}X_{vl,m})$  and define  $\text{I}^\diamond$ ,  $\text{II}^\diamond$ ,  $\text{III}^\diamond$  and  $\text{IV}^\diamond$  with the multiplier  $e^{-\iota[(t-s)\theta - (u-v)\theta']}$  in place of  $e^{-\iota[(t-s)\theta + (u-v)\theta']}$ . It follows that  $\text{I}^\diamond = (\mathbb{E}\widehat{G}_{ij,m}(\theta))(\mathbb{E}\widehat{G}_{hl,m}^H(\theta'))$ . Similar to III, we can derive  $\text{II}^\diamond = \mathbb{E}\widehat{G}_{ih,m}(\theta)\mathbb{E}\widehat{G}_{lj,m}(\theta)$  if  $\theta = \theta'$  and  $|\text{II}^\diamond| \lesssim B^{-2}\Phi_2^4$  if  $\theta \neq \theta'$ . And

$$\text{III}^\diamond = \frac{1}{4\pi^2 B^2} \left( \sum_{t,u=1}^B \Gamma_{il,m}(t-v)e^{-\iota(t\theta + v\theta')} \right) \times \left( \sum_{s=1}^B \Gamma_{jh,m}(s-u)e^{\iota(s\theta + u\theta')} \right)$$

can be dealt with similarly as II. □

**Lemma 19.** *For any  $1 \leq m \leq B$ ,  $1 \leq i, j \leq p$  and  $\theta \in \Theta$ , define*

$$\widehat{H}_{ij,m}(\theta) = \frac{1}{2\pi B} \sum_{t,s=1}^{B-m} X_{ti,m}X_{sj,m}e^{-\iota(t-s)\theta}. \quad (1.9.11)$$

Let

$$\mathcal{E}_{B,m}^{(2)} = (\rho^m + \sqrt{\frac{m}{B}})\Phi_4^4. \quad (1.9.12)$$

Then Lemma 18 (i), (ii) and (iii) hold for  $H_{ij}(\theta)$  and  $H_{hl}(\theta')$  with  $\mathcal{E}_{B,m}^{(1)}$  replaced by  $\mathcal{E}_{B,m}^{(2)}$ .

*Proof of Lemma 19.* Recall that

$$\widehat{G}_{ij,m}(\theta) = \frac{1}{2\pi B} \sum_{t,s=1}^B X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}.$$

Then we have  $2\pi B(\widehat{G}_{ij,m}(\theta) - \widehat{H}_{ij,m}(\theta)) = D_1 + D_2 - D_3$ , where

$$\begin{aligned} D_1 &= \sum_{t=B-m+1}^B \sum_{s=1}^B X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}, \\ D_2 &= \sum_{s=B-m+1}^B \sum_{t=1}^B X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}, \\ D_3 &= \sum_{t,s=B-m+1}^B X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}. \end{aligned}$$

Applying similar arguments of deriving equation (41) in Lemma 8 of [111] to the process  $(X_{ti,m})_t$  and  $(X_{tj,m})_t$ , we have

$$\|\mathbb{E}_0 D_1\|_2 \lesssim \sqrt{Bm} \Phi_4^2, \quad \|\mathbb{E}_0 D_2\|_2 \lesssim \sqrt{Bm} \Phi_4^2, \quad \|\mathbb{E}_0 D_3\|_2 \lesssim m \Phi_4^2.$$

It follows that

$$\|\mathbb{E}_0 \widehat{H}_{ij,m}(\theta) - \mathbb{E}_0 \widehat{G}_{ij,m}(\theta)\|_2 \lesssim \sqrt{\frac{m}{B}} \Phi_4^2. \quad (1.9.13)$$

Applying Lemma 8 of [111] again, it follows that

$$\|\mathbb{E}_0 \widehat{H}_{ij,m}(\theta)\|_2 \lesssim \Phi_4^2, \quad \|\mathbb{E}_0 \widehat{G}_{ij,m}(\theta)\|_2 \lesssim \Phi_4^2.$$

By the triangle inequality and Hölder's inequality, we have

$$\begin{aligned}
& \left| \mathbb{E}(\mathbb{E}_0 \widehat{H}_{ij,m}(\theta) \mathbb{E}_0 \widehat{H}_{hl,m}(\theta')) - \mathbb{E}(\mathbb{E}_0 \widehat{G}_{ij,m}(\theta) \mathbb{E}_0 \widehat{G}_{hl,m}(\theta')) \right| \\
& \leq \|\mathbb{E}_0 \widehat{H}_{ij,m}(\theta)\|_2 \|\mathbb{E}_0 \widehat{G}_{hl,m}(\theta') - \mathbb{E}_0 \widehat{H}_{hl,m}(\theta')\|_2 \\
& \quad + \|\mathbb{E}_0 \widehat{G}_{hl,m}(\theta')\|_2 \|\mathbb{E}_0 \widehat{H}_{ij,m}(\theta) - \mathbb{E}_0 \widehat{G}_{ij,m}(\theta)\|_2 \\
& \lesssim \sqrt{\frac{m}{B}} \Phi_4^4.
\end{aligned} \tag{1.9.14}$$

Similarly, we can obtain

$$\left| \mathbb{E}(\mathbb{E}_0 \widehat{H}_{ij,m}(\theta) \mathbb{E}_0 \widehat{H}_{hl,m}^*(\theta')) - \mathbb{E}(\mathbb{E}_0 \widehat{G}_{ij,m}(\theta) \mathbb{E}_0 \widehat{G}_{hl,m}^*(\theta')) \right| \lesssim \sqrt{\frac{m}{B}} \Phi_4^4.$$

Therefore, Lemma 18 also holds for  $H_{ij}(\theta)$  and  $H_{hl}(\theta')$  with the additional error bound  $\sqrt{m/B} \Phi_4^4$ .  $\square$

**Lemma 20.** Assume  $\Phi_{2,\alpha} < \infty$  and  $\Psi_{2,1} < \infty$ . For  $1 \leq m \leq B$ , define  $X_{t,m} = \mathbb{E}(X_t | \varepsilon_{t-m}, \varepsilon_{t-m+1}, \dots, \varepsilon_t)$  and

$$\widehat{G}_{ij,m}(\theta) = \frac{1}{2\pi B} \sum_{t,s=1}^B X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}.$$

Then for any  $1 \leq i, j \leq p$  and  $0 \leq \theta \leq \pi$ ,

$$\left| \mathbb{E} \widehat{G}_{ij,m}(\theta) - F_{ij}(\theta) \right| \lesssim m^{-\alpha} \Phi_{2,0} \Phi_{2,\alpha} + B^{-1} \Phi_{2,0} \Psi_{2,1}. \tag{1.9.15}$$

*Proof of Lemma 20.* The proof follows by arguments similar to those of Lemma 16.  $\square$

**Lemma 21.** For  $1 \leq i, j, h, l \leq p$  and  $k_1, k_2, k_3 \in \mathbb{Z}$ , denote the 4-th joint cumulant as

$$\gamma_{i,j,h,l}(k_1, k_2, k_3) = \text{Cum}(X_{0i}, X_{k_1j}, X_{k_2h}, X_{k_3l}).$$

Assume that  $\Phi_{4,0} < \infty$  and  $\Psi_{4,2} < \infty$ . Then for some constant  $C > 0$ ,

$$\sum_{k_1, k_2, k_3 \in \mathbb{Z}} |\gamma_{i,j,h,l}(k_1, k_2, k_3)| \leq C \Phi_{4,0}^2 \Psi_{4,2}^2.$$

*Proof of Lemma 21.* We can apply the similar arguments of proving Theorem 21 in [110] without extra difficulty.  $\square$

**Lemma 22.** Assume  $\Phi_{2,\alpha} < \infty$  and  $\Psi_{4,2} < \infty$ . Let  $1 \leq m \leq B$ . Further assume that

$$\mathcal{E}_{B,m}^* := m^{-\alpha} \Phi_{2,0}^3 \Phi_{2,\alpha} + B^{-1} \Phi_{4,0}^2 \Psi_{4,2}^2 = o(1).$$

Take any  $\theta, \theta' \in \Theta$  and any  $1 \leq i, j, h, l \leq p$ . (i) If  $\theta \in (0, \pi)$ ,

$$\begin{aligned} |\mathbb{E}(\widehat{G}_{ij,m}(\theta) \widehat{G}_{hl,m}(\theta)) - \mathbb{E} \widehat{G}_{ij,m}(\theta) \mathbb{E} \widehat{G}_{hl,m}(\theta) - F_{il}(\theta) F_{hj}(\theta)| &\lesssim \mathcal{E}_{B,m}^*, \\ |\mathbb{E}(\widehat{G}_{ij,m}(\theta) \widehat{G}_{hl,m}^H(\theta)) - \mathbb{E} \widehat{G}_{ij,m}(\theta) \mathbb{E} \widehat{G}_{hl,m}^H(\theta) - F_{ih}(\theta) F_{lj}(\theta)| &\lesssim \mathcal{E}_{B,m}^*. \end{aligned}$$

(ii) If  $\theta = 0$  or  $\pi$ ,

$$|\mathbb{E}(\widehat{G}_{ij,m}(\theta) \widehat{G}_{hl,m}(\theta)) - \mathbb{E} \widehat{G}_{ij,m}(\theta) \mathbb{E} \widehat{G}_{hl,m}(\theta) - F_{ih}(\theta) F_{lj}(\theta) - F_{il}(\theta) F_{hj}(\theta)| \lesssim \mathcal{E}_{B,m}^*.$$

(iii) If  $\theta \neq \theta'$ ,

$$\begin{aligned} |\mathbb{E}(\widehat{G}_{ij,m}(\theta) \widehat{G}_{hl,m}(\theta')) - \mathbb{E} \widehat{G}_{ij,m}(\theta) \mathbb{E} \widehat{G}_{hl,m}(\theta')| &\lesssim \mathcal{E}_{B,m}^*, \\ |\mathbb{E}(\widehat{G}_{ij,m}(\theta) \widehat{G}_{hl,m}^H(\theta')) - \mathbb{E} \widehat{G}_{ij,m}(\theta) \mathbb{E} \widehat{G}_{hl,m}^H(\theta')| &\lesssim \mathcal{E}_{B,m}^*, \end{aligned}$$

*Proof of Lemma 22.* The proof follows by arguments similar to those of Lemma 18.  $\square$

**Lemma 23.** Assume  $m = o(B)$  and

$$\mathcal{E}_{B,m} := m^{-\alpha} \Phi_{2,0}^3 \Phi_{2,\alpha} + B^{-1} \Phi_{4,0}^2 \Psi_{4,2}^2 + \sqrt{\frac{m}{B}} \Phi_{4,0}^4 = o(1). \quad (1.9.16)$$

For any  $1 \leq i, j \leq p$  and  $\theta \in \Theta$ , define

$$\widehat{H}_{ij,m}(\theta) = \frac{1}{2\pi B} \sum_{t,s=1}^{B-m} X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}. \quad (1.9.17)$$

Then Lemma 22 (i), (ii) and (iii) hold for  $H_{ij}(\theta)$  and  $H_{hl}(\theta')$  with  $\mathcal{E}_{B,m}^*$  replaced by  $\mathcal{E}_{B,m}$ .

*Proof of Lemma 23.* The proof follows by arguments similar to those of Lemma 19.  $\square$

### 1.9.2 Probability Inequalities for High Dimensional Time Series

Let  $n = T/B$ . Define  $\mathcal{B}_r = \{(r-1)B+1, \dots, rB\}$  and  $\mathcal{D}_r = \{(r-1)B+1, \dots, (r-1)B+B-m\}$  for  $r = 1, \dots, n$ . For  $m \geq 1$ , let  $X_{t,m} = \mathbb{E}(X_t | \varepsilon_{t-m}, \varepsilon_{t-m+1}, \dots, \varepsilon_t)$ . Define

$$\widehat{F}_m(\theta) = (\widehat{F}_{ij,m}(\theta))_{i,j=1}^p = \frac{1}{2\pi T} \sum_{r=1}^n \sum_{t,s \in \mathcal{B}_r} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}, \quad (1.9.18)$$

$$\widehat{H}_{r,m}(\theta) = (\widehat{H}_{ij,r,m}(\theta))_{i,j=1}^p = \frac{1}{2\pi B} \sum_{t,s \in \mathcal{D}_r} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}, \quad (1.9.19)$$

$$\widehat{S}_m(\theta) = (\widehat{S}_{ij,m}(\theta))_{i,j=1}^p = \frac{1}{n} \sum_{r=1}^n \widehat{H}_{r,m}(\theta). \quad (1.9.20)$$

**Lemma 24.** *Let Assumption 1 be satisfied for some  $\alpha \geq 0$ . Then there exists constants  $C_\alpha > 0$  such that for any  $x > 0$ ,*

$$\mathbb{P}(|\mathbb{E}_0(\widehat{F}_{ij}(\theta)) - \mathbb{E}_0(\widehat{F}_{ij,m}(\theta))| \geq x) \lesssim \exp \left\{ -C_\alpha \left( \frac{\sqrt{T}x}{\sqrt{B}\rho^m \|X_{\cdot}\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}.$$

*Proof of Lemma 24.* We follow similar arguments of proving Theorem 1. In particular, we use  $\tilde{R}_{ij}(\theta) = 2\pi T(\widehat{F}_{ij}(\theta) - \widehat{F}_{ij,m}(\theta))$  to replace  $R_{ij}(\theta)$ . By Lemma 9 in [111], we have

$$\begin{aligned} \|\tilde{R}_{ij}(\theta) - \mathbb{E}\tilde{R}_{ij}(\theta)\|_{q/2} &\leq C(q-2)\rho^m \sqrt{TB} \|X_{\cdot i}\|_q \|X_{\cdot j}\|_q \\ &\leq C(q-2)q^{2\alpha}\rho^m \sqrt{TB} \|X_{\cdot}\|_{\psi_\alpha}^2. \end{aligned} \quad (1.9.21)$$

With (1.9.21) in place of (1.7.5) and by similar arguments of Theorem 1, the result can be proved without extra difficulty.  $\square$

**Lemma 25.** *Let Assumption 1 be satisfied for some  $\alpha \geq 0$ . Then there exists constants  $C_\alpha > 0$  such that for any  $x > 0$ ,*

$$\mathbb{P}(|\mathbb{E}_0(\widehat{F}_{ij,m}(\theta)) - \mathbb{E}_0(\widehat{S}_{ij,m}(\theta))| \geq x) \lesssim \exp \left\{ -C_\alpha \left( \frac{\sqrt{T}x}{\sqrt{m}\|X_\cdot\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}. \quad (1.9.22)$$

*Proof of Lemma 25.* Write  $2\pi T(\widehat{F}_{ij,m}(\theta) - \widehat{S}_{ij,m}(\theta)) = \text{I} + \text{II} + \text{III}$ , where

$$\begin{aligned} \text{I} &= \sum_{r=1}^n \sum_{t=rB-m+1}^{rB} \sum_{s=(r-1)B}^{rB} X_{ti,m} X_{sj,m}^\top e^{-\iota(t-s)\theta}, \\ \text{II} &= \sum_{r=1}^n \sum_{s=rB-m+1}^{rB} \sum_{t=(r-1)B}^{rB} X_{ti,m} X_{sj,m}^\top e^{-\iota(t-s)\theta}, \\ \text{III} &= \sum_{r=1}^n \sum_{t,s=rB-m+1}^{rB} X_{ti,m} X_{sj,m}^\top e^{-\iota(t-s)\theta}. \end{aligned}$$

We can adopt similar arguments of proving (1.7.7) to obtain

$$\begin{aligned} \mathbb{P}(|\text{I} - \mathbb{EI}| \geq x) &\lesssim \exp \left\{ -C_\alpha \left( \frac{x}{\sqrt{Tm}\|X_\cdot\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}, \\ \mathbb{P}(|\text{II} - \mathbb{EII}| \geq x) &\lesssim \exp \left\{ -C_\alpha \left( \frac{x}{\sqrt{Tm}\|X_\cdot\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}, \\ \mathbb{P}(|\text{III} - \mathbb{EIII}| \geq x) &\lesssim \exp \left\{ -C_\alpha \left( \frac{\sqrt{B}x}{\sqrt{Tm}\|X_\cdot\|_{\psi_\alpha}^2} \right)^{\frac{1}{2\alpha+1}} \right\}. \end{aligned}$$

Then (1.9.22) can be obtained by combining the three parts.  $\square$

Denote  $\theta_k = 2\pi k/B$  for  $k \in \mathbb{Z}$ . Define

$$\mathbf{H}_m = (\text{Re}(\widehat{H}_m(\theta_0)), \text{Im}(\widehat{H}_m(\theta_0)), \dots, \text{Re}(\widehat{H}_m(\theta_{B/2})), \text{Im}(\widehat{H}_m(\theta_{B/2}))),$$

where

$$\hat{H}_m(\theta) = (\hat{H}_{ij,m}(\theta))_{i,j=1}^p = \frac{1}{2\pi B} \left( \sum_{t=1}^{B-m} X_{t,m} e^{-it\theta} \right) \left( \sum_{t=1}^{B-m} X_{t,m}^\top e^{it\theta} \right).$$

Let  $\tilde{p} = Bp^2 - (B/2 - 1)p$  and let  $\tilde{\mathbf{W}} = (\tilde{W}_1, \tilde{W}_2, \dots, \tilde{W}_{\tilde{p}})^\top = \mathbf{vec}_0(\mathbf{H}_m)$  be the vectorization of  $\mathbf{H}_m$  where the constantly zero entries, i.e.,  $\text{Im}(\hat{H}_{ii,m}(\theta))$  for  $1 \leq i \leq p$  and all the entries in  $\text{Im}(\hat{H}_m(0))$  and  $\text{Im}(\hat{H}_m(\pi))$ , are excluded. Denote  $\Sigma_m = \text{Cov}(\mathbf{vec}_0(\mathbf{H}_m))$  and  $\mathbf{Z}_m \sim N(0, \Sigma_m)$ .

**Lemma 26.** *Suppose Assumption 1 is satisfied for some  $\alpha \geq 0$ . Assume that there exists some constant  $b_0 > 0$  such that  $\min_{1 \leq j \leq \tilde{p}} \text{Var}(\tilde{W}_j) \geq b_0$ . Let  $\mathbf{W}_1, \dots, \mathbf{W}_n$  be i.i.d. copies of  $\tilde{\mathbf{W}}$  and  $\mathbf{Z}_m \sim N(0, \text{Cov}(\mathbf{vec}_0(\mathbf{H}_m)))$ . Then for  $\bar{\mathbf{W}}_n = n^{-1} \sum_{i=1}^n \mathbf{W}_i$ ,*

$$\sup_x |\mathbb{P}(\sqrt{n}|\bar{\mathbf{W}}_n - \mathbb{E}\bar{\mathbf{W}}_n|_\infty \leq x) - \mathbb{P}(|\mathbf{Z}_m|_\infty \leq x)| \lesssim \left( \frac{\|X\|_{\psi_\alpha}^{12} \log^{\kappa_1}(\tilde{p} \vee n)}{n} \right)^{\frac{1}{6}}, \quad (1.9.23)$$

where  $\kappa_1 = \max\{7, 2\alpha + 6\}$ .

*Proof of Lemma 26.* Without loss of generality, we may assume that

$$\frac{\|X\|_{\psi_\alpha}^{12} \log^{\kappa_1}(\tilde{p} \vee n)}{n} \leq C, \quad (1.9.24)$$

where  $C$  is a sufficiently small constant. Otherwise, (1.9.23) trivially holds by taking the constant in  $\lesssim$  large enough. We shall apply Theorem 2.1 in [24] to prove this Lemma. Define  $L = \max_{1 \leq j \leq \tilde{p}} \mathbb{E}|\mathbb{E}_0 \tilde{W}_j|^3$ . For  $\phi \geq 1$ , define

$$\begin{aligned} M_{n,W}(\phi) &= \mathbb{E}[|\mathbb{E}_0 \tilde{\mathbf{W}}|_\infty^3 \mathbf{1}\{|\mathbb{E}_0 \tilde{\mathbf{W}}|_\infty > \sqrt{n}/(4\phi \log \tilde{p})\}], \\ M_{n,Z}(\phi) &= \mathbb{E}[|\mathbf{Z}_m|_\infty^3 \mathbf{1}\{|\mathbf{Z}_m|_\infty > \sqrt{n}/(4\phi \log \tilde{p})\}]. \end{aligned}$$

Denote  $M_n(\phi) = M_{n,W}(\phi) + M_{n,Z}(\phi)$ . By Theorem 2.1 in [24], there exist constants

$K_1, K_2 > 0$  depending only on  $b_0$  such that for every constant  $\bar{L} \geq L$ ,

$$\sup_x |\mathbb{P}(\sqrt{n}|\bar{\mathbf{W}}_n - \mathbb{E}\bar{\mathbf{W}}_n|_\infty \leq x) - \mathbb{P}(|\mathbf{Z}_m|_\infty \leq x)| \leq K_1 \left[ \left( \frac{\bar{L}^2 \log^7 \tilde{p}}{n} \right)^{1/6} + \frac{M_n(\phi_n)}{\bar{L}} \right] \quad (1.9.25)$$

with

$$\phi_n = K_2 \left( \frac{\bar{L}^2 \log^4 \tilde{p}}{n} \right)^{-1/6}.$$

Applying Equation (41) in Lemma 8 of [111] for the third moment, we have  $2\pi\|\mathbb{E}_0\tilde{W}_j\|_3 \leq 4\sqrt{10}\Phi_6^2$ . Since  $\Phi_6 \leq 6^\alpha\|X.\|_{\psi_\alpha}$ , we can denote

$$L = \max_{1 \leq j \leq p} \mathbb{E}|\mathbb{E}_0\tilde{W}_j|^3 \leq K_\alpha^3\|X.\|_{\psi_\alpha}^6.$$

Choose

$$\bar{L} = K_\alpha^3 \log^{c_0}(\tilde{p} \vee n) \|X.\|_{\psi_\alpha}^6, \quad (1.9.26)$$

where  $c_0 \geq 0$  and is to be determined later. For  $C \leq (K_2/K_\alpha)^6$ , (1.9.24) implies

$$\phi_n := K_2(n^{-1}\bar{L}^2 \log^4 \tilde{p})^{-1/6} \geq 1.$$

By similar arguments of proving (1.7.7), with  $C_0 = 1/(1 + 2\alpha) \in (0, 1]$ , we have

$$\mathbb{P}(|\mathbb{E}_0\tilde{\mathbf{W}}|_\infty \geq x) \lesssim \tilde{p} \exp \left( - \frac{x^{C_0}}{C_\alpha \|X.\|_{\psi_\alpha}^{2C_0}} \right). \quad (1.9.27)$$

For  $\tau = \sqrt{n}/(4\phi_n \log \tilde{p})$ , we have

$$M_{n,W}(\phi_n) = \tau^3 \mathbb{P}(|\mathbb{E}_0\tilde{\mathbf{W}}|_\infty > \tau) + 3 \int_\tau^\infty x^2 \mathbb{P}(|\mathbb{E}_0\tilde{\mathbf{W}}|_\infty > x) dx. \quad (1.9.28)$$

Using integration by parts, it can be obtained that

$$\begin{aligned}
& \int_{\tau}^{\infty} x^2 \exp\left(-\frac{x^{C_0}}{C_{\alpha}\|X.\|_{\psi_{\alpha}}^{2C_0}}\right) dx \\
&= \frac{C_{\alpha}\|X.\|_{\psi_{\alpha}}^{2C_0} \tau^{3-C_0}}{C_0} \exp\left(-\frac{\tau^{C_0}}{C_{\alpha}\|X.\|_{\psi_{\alpha}}^{2C_0}}\right) \\
&+ \frac{C_{\alpha}(3-C_0)\|X.\|_{\psi_{\alpha}}^{2C_0}}{C_0} \int_{\tau}^{\infty} x^{2-C_0} \exp\left(-\frac{x^{C_0}}{C_{\alpha}\|X.\|_{\psi_{\alpha}}^{2C_0}}\right) dx.
\end{aligned} \tag{1.9.29}$$

Note that  $\tau = (4K_2)^{-1}(n\bar{L}/\log \tilde{p})^{1/3}$ . Let  $c_0 = \max\{0, C_0^{-1} - 2\}$  and let  $C$  in (1.9.24) be small enough such that

$$\frac{\tau^{C_0}}{C_{\alpha}\|X.\|_{\psi_{\alpha}}^{2C_0} \log(\tilde{p} \vee n)} \geq \max\{3, (6 - 2C_0)/C_0\}.$$

By (1.9.27), (1.9.28) and (1.9.29), elementary calculation implies  $M_{n,W}(\phi_n) \lesssim \|X.\|_{\psi_{\alpha}}^6/n$ .

Since  $\mathbf{Z}_m$  is a Gaussian vector and Lemma 8 in [111] suggests

$$\text{Var}(\tilde{W}_j) \leq [(2\pi)^{-1} 4^{2\alpha+1} \sqrt{3} \|X.\|_{\psi_{\alpha}}^2]^2,$$

we have

$$\begin{aligned}
P(|\mathbf{Z}_m|_{\infty} \geq x) &\leq 2 \sum_{j=1}^{\tilde{p}} \int_x^{\infty} \frac{1}{\sqrt{2\pi \text{Var}(\tilde{W}_j)}} \exp\left(-\frac{y^2}{2\text{Var}(\tilde{W}_j)}\right) dy \\
&\lesssim \frac{\tilde{p} \|X.\|_{\psi_{\alpha}}^2}{x} \exp\left(-\frac{x^2}{C'_{\alpha} \|X.\|_{\psi_{\alpha}}^2}\right).
\end{aligned}$$

A similar argument of dealing with  $M_{n,W}(\phi_n)$  also implies  $M_{n,Z}(\phi_n) \lesssim \|X.\|_{\psi_{\alpha}}^6/n$ . By (1.9.25) and (1.9.26), (1.9.23) follows in view of the fact that  $M_n(\phi_n)/\bar{L}$  is negligible compared to the first term.  $\square$

**Lemma 27.** Let  $X_1, \dots, X_T \in \mathbb{R}^p$  be a stationary process generated from (1.1.2) with  $\mathbb{E}X_t = 0$  and  $\Omega_{0,q} < \infty$  for  $q \geq 2$ . Let  $a_{t,\theta}, a_{t,s,\theta} \in \mathbb{C}$  be constants that depend on  $\theta$  for  $\theta \in \Theta$ . Denote  $d = |\Theta|$ ,  $a_T = \left( \sum_{t=1}^T \max_{\theta \in \Theta} |a_{t,\theta}|^2 \right)^{1/2}$  and

$$A_T = \left( \max_{1 \leq t \leq T} \left( \sum_{s=1}^T \max_{\theta \in \Theta} |a_{t,s,\theta}|^2 \right) + \max_{1 \leq s \leq T} \left( \sum_{t=1}^T \max_{\theta \in \Theta} |a_{t,s,\theta}|^2 \right) \right)^{\frac{1}{2}}.$$

Let  $\ell = 1 \vee \log(pd)$ . Then there exists some positive constant  $C_q$  such that

$$\begin{aligned} (i) \quad & \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T a_{t,\theta} (X_t - X_{t,m}) \right|_{\infty} \right\|_q \leq C_q \ell^{\frac{1}{2}} a_T \Omega_{m+1,q}, \\ (ii) \quad & \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T a_{t,\theta} X_t \right|_{\infty} \right\|_q \leq C_q \ell^{\frac{1}{2}} a_T \Omega_{0,q}, \\ (iii) \quad & \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T a_{t,\theta} X_{t,m} \right|_{\infty} \right\|_q \leq C_q \ell^{\frac{1}{2}} a_T \Omega_{0,q}, \\ (iv) \quad & \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} - \mathbb{E} \left( \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} \right) \right|_{\infty} \right\|_q \leq C_q \ell \sqrt{T} A_T \Omega_{0,2q}^2. \end{aligned}$$

*Proof of Lemma 27.* (i) We first write

$$X_t - X_{t,m} = \mathbb{E}(X_t | \mathcal{F}^{-\infty}) - \mathbb{E}(X_t | \mathcal{F}^{t-m}) = \sum_{k=m+1}^{\infty} \mathcal{P}^{t-k} X_t.$$

Then we have

$$\sum_{t=1}^T a_{t,\theta} (X_t - X_{t,m}) = \sum_{t=1}^T a_{t,\theta} \sum_{k=m+1}^{\infty} \mathcal{P}^{t-k} X_t = \sum_{k=m+1}^{\infty} \sum_{t=1}^T a_{t,\theta} \mathcal{P}^{t-k} X_t.$$

We know that for each fixed  $k \geq m+1$ ,  $a_{t,\theta} \mathcal{P}^{t-k} X_t$  is a backward martingale difference with respect to the filtration  $\mathcal{F}^{t-k}$  for  $1 \leq t \leq T$ . By the triangle inequality and Lemma

33, it follows that

$$\begin{aligned} \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T a_{t,\theta} (X_t - X_{t,m}) \right|_{\infty} \right\|_q &\leq \sum_{k=m+1}^{\infty} \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T a_{t,\theta} \mathcal{P}^{t-k} X_t \right|_{\infty} \right\|_q \\ &\leq \sum_{k=m+1}^{\infty} C_q \ell^{\frac{1}{2}} a_{T,\theta} \omega_{k,q} = C_q \ell^{\frac{1}{2}} a_T \Omega_{m+1,q}. \end{aligned}$$

Since  $X_t = \sum_{k=0}^{\infty} \mathcal{P}^{t-k} X_t$  and  $X_{t,m} = \sum_{k=0}^m \mathcal{P}^{t-k} X_t$ , the preceding argument implies (ii) and (iii).

(iv) By martingale decomposition, we have

$$\sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} - \mathbb{E} \left( \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} \right) = \sum_{k=0}^{\infty} \mathcal{P}^{T-k} \left( \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} \right).$$

For each  $k \geq 0$ , by the triangle inequality and Jensen's inequality, we have

$$\begin{aligned} &\left\| \max_{\theta \in \Theta} \left| \mathcal{P}^{T-k} \left( \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} \right) \right|_{\infty} \right\|_q \\ &\leq \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} - \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_{t,\{T-k\}} X_{s,\{T-k\}}^{\top} \right|_{\infty} \right\|_q \\ &\leq \text{I}_k + \text{II}_k, \end{aligned}$$

where

$$\begin{aligned} \text{I}_k &= \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} (X_t - X_{t,\{T-k\}}) X_s^{\top} \right|_{\infty} \right\|_q, \\ \text{II}_k &= \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_{t,\{T-k\}} (X_s - X_{s,\{T-k\}}^{\top}) \right|_{\infty} \right\|_q. \end{aligned}$$

By Jensen's equality, we have  $\left\| \left| X_t - X_{t,\{T-k\}} \right|_{\infty} \right\|_{2q} \leq w_{t-T+k,2q}$ . It follows from Hölder's

inequality and (ii) that

$$\begin{aligned} I_k &\leq \sum_{t=1}^T \left\| \left\| X_t - X_{t, \{T-k\}} \right\|_{\infty} \right\|_{2q} \left\| \max_{\theta \in \Theta} \left| \sum_{s=1}^T a_{t,s,\theta} X_s^{\top} \right|_{\infty} \right\|_{2q} \\ &\leq C_q \ell^{1/2} \max_{1 \leq t \leq T} \left( \sum_{s=1}^T \max_{\theta \in \Theta} |a_{t,s,\theta}|^2 \right)^{\frac{1}{2}} \Omega_{0,2q} \sum_{t=1}^T w_{t-T+k,2q}. \end{aligned}$$

Then

$$\begin{aligned} \sum_{k=0}^{\infty} I_k^2 &\leq \sum_{k=0}^{\infty} C_q \ell \max_{1 \leq t \leq T} \left( \sum_{s=1}^T \max_{\theta \in \Theta} |a_{t,s,\theta}|^2 \right) \Omega_{0,2q}^2 \left( \sum_{t=1}^T w_{t-T+k,2q} \right)^2 \\ &\leq C_q \ell T \max_{1 \leq t \leq T} \left( \sum_{s=1}^T \max_{\theta \in \Theta} |a_{t,s,\theta}|^2 \right) \Omega_{0,2q}^4. \end{aligned}$$

Similarly, we can have

$$\sum_{k=0}^{\infty} \Pi_k^2 \leq C_q \ell T \max_{1 \leq s \leq T} \left( \sum_{t=1}^T \max_{\theta \in \Theta} |a_{t,s,\theta}|^2 \right) \Omega_{0,2q}^4.$$

By Lemma 33,

$$\begin{aligned} &\left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} - \mathbb{E} \left( \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} \right) \right|_{\infty} \right\|_q^2 \\ &\leq C_q \ell \sum_{k=0}^{\infty} \left\| \max_{\theta \in \Theta} \left| \mathcal{P}^{T-k} \left( \sum_{t=1}^T \sum_{s=1}^T a_{t,s,\theta} X_t X_s^{\top} \right) \right|_{\infty} \right\|_q^2 \\ &\leq C_q \ell \sum_{k=0}^{\infty} (I_k^2 + \Pi_k^2) \leq C_q \ell^2 T A_T^2 \Omega_{0,2q}^4. \end{aligned}$$

Thus (iv) follows.  $\square$

In Lemma 28, Lemma 29 and Theorem 30 below, we concern quadratic forms of high dimensional stationary processes generated from (1.1.2) with  $\mathbb{E}X_t = 0$ . For  $B < T$ ,

define the quadratic form indexed by  $\theta \in \Theta$ ,

$$Q_{T,\theta} = \sum_{1 \leq t \leq s \leq T} a_{s,t,\theta} X_t X_s^\top$$

where  $a_{s,t,\theta} \in \mathbb{C}$  are complex numbers such that  $|a_{s,t,\theta}| \leq 1$ ,  $a_{s,t,\theta} = a_{s-t,\theta}$  and  $a_{s,t,\theta} = 0$  if  $s - t > B$ . For  $m \leq B$ , define the  $m$ -approximation of  $Q_{T,\theta}$ ,

$$Q_{T,\theta,m} = \sum_{1 \leq t \leq s \leq T} a_{s,t,\theta} X_{t,m} X_{s,m}^\top.$$

**Lemma 28.** *Assume  $\Omega_{0,q} < \infty$  for some  $q \geq 4$ . Denote  $d = |\Theta|$  and  $\ell = 1 \vee \log(pd)$ .*

*There exists some constant  $C_q > 0$  depending on  $q$  such that*

$$\left\| \max_{\theta \in \Theta} |\mathbb{E}_0 Q_{T,\theta} - \mathbb{E}_0 Q_{T,\theta,m}|_\infty \right\|_{q/2} \leq C_q \sqrt{BT} \ell \Omega_{0,q} \Omega_{m,q}.$$

*Proof of Lemma 28.* Let  $w = \lfloor m/2 \rfloor$ . We first write

$$Q_{T,\theta} = \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} X_t X_s^\top + \sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} X_t X_s^\top =: Q^{(1)} + Q^{(2)},$$

where  $X_t = 0$  for  $t \notin [1, T]$ . We can define the  $m$ -approximations  $Q_m^{(1)}$  and  $Q_m^{(2)}$  accordingly.

By elementary manipulation, we can write

$$\begin{aligned} Q^{(1)} - Q_m^{(1)} &= \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} X_t X_s^\top - \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} X_{t,m} X_{s,m}^\top \\ &= S_1 + S_2 + S_3, \end{aligned}$$

where

$$S_1 = \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} (X_t - X_{t,w})(X_s^\top - X_{s,m}^\top),$$

$$\begin{aligned}
S_2 &= \sum_{t=1}^T (X_t - X_{t,m}) \sum_{s=t}^{t+w} a_{s,t,\theta} X_{s,m}^\top, \\
S_3 &= \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} X_{t,w} (X_s^\top - X_{s,m}^\top).
\end{aligned}$$

For the first term, write  $S_1 - \mathbb{E}S_1 = \sum_{k=-\infty}^T \mathcal{P}^k S_1$ . For  $k \leq T$ ,  $\mathcal{P}^k S_1$  is a backward martingale difference with respect to the filtration  $\mathcal{F}^k$ . Then

$$\left\| \max_{\theta \in \Theta} |\mathcal{P}^k S_1|_\infty \right\|_{q/2} \leq \text{I}_k + \text{II}_k$$

for

$$\begin{aligned}
\text{I}_k &= \left\| \max_{\theta \in \Theta} \left| \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} (X_t - X_{t,w} - X_{t,\{k\}} + X_{t,w,\{k\}}) (X_s^\top - X_{s,m}^\top) \right|_\infty \right\|_{q/2}, \\
\text{II}_k &= \left\| \max_{\theta \in \Theta} \left| \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} (X_{t,\{k\}} - X_{t,w,\{k\}}) (X_s^\top - X_{s,m}^\top - X_{s,\{k\}}^\top + X_{s,m,\{k\}}^\top) \right|_\infty \right\|_{q/2}.
\end{aligned}$$

By Jensen's inequality,

$$\begin{aligned}
\left\| |X_{t,w} - X_{t,w,\{k\}}|_\infty \right\|_q &= \left\| |\mathbb{E}(X_t | \mathcal{F}^{t-w}) - \mathbb{E}(X_{t,\{k\}} | \mathcal{F}^{t-w,\{k\}})|_\infty \right\|_q \\
&= \left\| |\mathbb{E}(X_t - X_{t,\{k\}} | \mathcal{F}^{t-w}, \epsilon'_k)|_\infty \right\|_q \leq w_{t-k,q}.
\end{aligned}$$

By Lemma 33 ,

$$\left\| |X_t - X_{t,w}|_\infty \right\|_q = \left\| \left| \sum_{l=w+1}^\infty \mathcal{P}^{t-l} X_t \right|_\infty \right\|_q \leq C_q \ell^{\frac{1}{2}} \left( \sum_{l=w+1}^\infty w_{l,q}^2 \right)^{\frac{1}{2}}.$$

So we have

$$\left\| |X_t - X_{t,w} - X_{t,\{k\}} + X_{t,w,\{k\}}|_\infty \right\|_q \leq 2 \min \left\{ C_q \ell^{\frac{1}{2}} \left( \sum_{l=w+1}^\infty w_{l,q}^2 \right)^{\frac{1}{2}}, w_{t-k,q} \right\}.$$

It follows from Lemma 27 that

$$\begin{aligned}
I_k &= \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T (X_t - X_{t,w} - X_{t,\{k\}} + X_{t,w,\{k\}}) \sum_{s=t}^{t+w} a_{s,t,\theta} (X_s^\top - X_{s,m}^\top) \right|_\infty \right\|_{q/2} \\
&\leq \sum_{t=1}^T \left\| \left| X_t - X_{t,w} - X_{t,\{k\}} + X_{t,w,\{k\}} \right|_\infty \right\|_q \left\| \max_{\theta \in \Theta} \left| \sum_{s=t}^{t+w} a_{s,t,\theta} (X_s^\top - X_{s,m}^\top) \right|_\infty \right\|_q \\
&\leq 2C_q \sqrt{B\ell} \Omega_{m,q} \sum_{t=1}^T \min \left\{ C_q \ell^{\frac{1}{2}} \left( \sum_{l=w+1}^{\infty} w_{l,q}^2 \right)^{\frac{1}{2}}, w_{t-k,q} \right\}.
\end{aligned}$$

Then

$$\begin{aligned}
\sum_{k=-\infty}^T I_k^2 &= 4C_q^2 B\ell \Omega_{m,q}^2 \sum_{k=-\infty}^T \bar{\Delta}_{w,p,q} \sum_{t=1}^T \min \left\{ C_q \ell^{\frac{1}{2}} \left( \sum_{l=w+1}^{\infty} w_{l,q}^2 \right)^{\frac{1}{2}}, w_{t-k,q} \right\} \\
&\leq 4C_q^2 B T \ell \Omega_{m,q}^2 \bar{\Delta}_{w,q}^2,
\end{aligned}$$

where  $\bar{\Delta}_{m,q} = \sum_{k=0}^{\infty} \min \{ C_q \ell^{\frac{1}{2}} (\sum_{l=m}^{\infty} w_{l,q}^2)^{\frac{1}{2}}, w_{k,q} \}$ . Similarly we have

$$\Pi_k \leq 2C_q \sqrt{B\ell} \Omega_{w,q} \sum_{s=1}^T \min \left\{ C_q \ell^{\frac{1}{2}} \left( \sum_{l=m+1}^{\infty} w_{l,q}^2 \right)^{\frac{1}{2}}, w_{s-k,q} \right\},$$

and  $\sum_{k=-\infty}^T \Pi_k^2 \leq 4C_q^2 B T \ell \Omega_{w,q}^2 \bar{\Delta}_{m,q}^2$ . By Lemma 33,

$$\begin{aligned}
\left\| \max_{\theta \in \Theta} |S_1 - \mathbb{E} S_1|_\infty \right\|_{q/2}^2 &\leq 2C_{q/2}^2 \ell \sum_{k=-\infty}^T \left\| \max_{\theta \in \Theta} |\mathcal{P}^k S_1|_\infty \right\|_{q/2}^2 \\
&\leq 16C_{q/2}^2 C_q^2 B T \ell^2 (\Omega_{m,q}^2 \bar{\Delta}_{w,q}^2 + \Omega_{w,q}^2 \bar{\Delta}_{m,q}^2).
\end{aligned}$$

In view of  $\bar{\Delta}_{w,q} \leq \Omega_{0,q}$ , we have

$$\left\| \max_{\theta \in \Theta} |S_1 - \mathbb{E} S_1|_\infty \right\|_{q/2} \leq 4C_{q/2} C_q \sqrt{B T \ell} \Omega_{0,q} \Omega_{m,q}. \quad (1.9.30)$$

For the third term, we write

$$S_3 = \sum_{k=m+1}^{\infty} \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} X_{t,w} \mathcal{P}^{s-k} X_s^\top,$$

where  $\sum_{t=s-w}^s a_{s,t,\theta} X_{t,w} \mathcal{P}^{s-k} X_s^\top$  is a backward martingale difference with respect to the filtration  $\mathcal{F}^{s-k}$  for  $1 \leq s \leq T$ . By Lemma 27 and Lemma 33,

$$\begin{aligned} & \left\| \max_{\theta \in \Theta} \left| \sum_{s=1}^T \sum_{t=s-w}^s a_{s,t,\theta} X_{t,w} \mathcal{P}^{s-k} X_s^\top \right|_\infty \right\|_{q/2}^2 \\ & \leq C_{q/2}^2 \ell \sum_{s=1}^T \left\| \max_{\theta \in \Theta} \left| \sum_{t=s-w}^s a_{s,t,\theta} X_{t,w} \mathcal{P}^{s-k} X_s^\top \right|_\infty \right\|_{q/2}^2 \\ & \leq C_{q/2}^2 \ell \sum_{s=1}^T \left\| \max_{\theta \in \Theta} \left| \sum_{t=s-w}^s a_{s,t,\theta} X_{t,w} \right|_\infty \right\|_q^2 \left\| \mathcal{P}^{s-k} X_s^\top \right\|_\infty^2 \\ & \leq C_{q/2}^2 \ell T C_q^2 \ell B \Omega_{0,q}^2 \omega_{k,q}^2. \end{aligned}$$

It follows that

$$\left\| \max_{\theta \in \Theta} |S_3|_\infty \right\|_{q/2} \leq C_{q/2} C_q \sqrt{BT} \ell \Omega_{0,q} \Omega_{m,q}. \quad (1.9.31)$$

Similarly, for the second term, we can have

$$\left\| \max_{\theta \in \Theta} |S_2|_\infty \right\|_{q/2} \leq C_{q/2} C_q \sqrt{BT} \ell \Omega_{0,q} \Omega_{m,q}. \quad (1.9.32)$$

Now we consider  $Q^{(2)}$ . Note that  $Q^{(2)}$  is nonzero only when  $B > w$ . Write

$$\begin{aligned} Q^{(2)} - Q_m^{(2)} &= \sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} X_t X_s^\top - \sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} X_{t,m} X_{s,m}^\top \\ &=: S_4 + S_5 + S_6, \end{aligned}$$

where

$$\begin{aligned}
S_4 &= \sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} (X_t - X_{t,m}) (X_s^\top - X_{s,m}^\top), \\
S_5 &= \sum_{t=1}^T (X_t - X_{t,m}) \sum_{s=t+w}^{t+B} a_{s,t,\theta} X_{s,m}^\top, \\
S_6 &= \sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} X_{t,m} (X_s^\top - X_{s,m}^\top).
\end{aligned}$$

Using the same argument as dealing with  $S_1$ , we have

$$\left\| \max_{\theta \in \Theta} |\mathcal{P}^k S_4|_\infty \right\|_{q/2} \leq 4C_q \sqrt{B\ell} \Omega_{m,q} \sum_{s=1}^T \min \left\{ C_q \ell^{\frac{1}{2}} \left( \sum_{l=m+1}^{\infty} w_{l,q}^2 \right)^{\frac{1}{2}}, w_{s-k,q} \right\},$$

which implies

$$\left\| \max_{\theta \in \Theta} |S_4 - \mathbb{E} S_4|_\infty \right\|_{q/2} \leq 4\sqrt{2} C_{q/2} C_q \sqrt{BT} \ell \Omega_{0,q} \Omega_{m,q}. \quad (1.9.33)$$

The term  $S_5$  can be dealt with similarly as  $S_3$ . It follows that

$$\left\| \max_{\theta \in \Theta} |S_5|_\infty \right\|_{q/2} \leq C_{q/2} C_q \sqrt{BT} \ell \Omega_{0,q} \Omega_{m,q}. \quad (1.9.34)$$

The last term  $S_6$  can be written as

$$\begin{aligned}
S_6 &= \sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} \sum_{l=0}^m \mathcal{P}^{t-l} X_t \sum_{k=m+1}^{\infty} \mathcal{P}^{s-k} X_s^\top \\
&= \sum_{k=m+1}^{\infty} \sum_{l=0}^m \sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top.
\end{aligned}$$

Then we can write  $\sum_{s=1}^T \sum_{t=s-B}^{s-w-1} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top = \text{III}_{k,l} + \text{IV}_l$ , where

$$\begin{aligned} \text{III}_{k,l} &= \sum_{s=1}^T \sum_{s-B \leq t \leq s-w-1, t-l \neq s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top, \\ \text{IV}_l &= \sum_{s=1}^T \sum_{s-B \leq t \leq s-w-1, t-l = s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top. \end{aligned}$$

Further write

$$\begin{aligned} \text{III}_{k,l} &= \sum_{s=1}^T \sum_{s-B \leq t \leq s-w-1, t-l > s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top \\ &\quad + \sum_{s=1}^T \sum_{s-B \leq t \leq s-w-1, t-l < s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top. \end{aligned}$$

Note that for each  $k$  and  $l$ , both  $\sum_{s-B \leq t \leq s-w-1, t-l < s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top$  and  $a_{s,t,\theta} \mathcal{P}^{t-l} X_t$  are backward martingale differences with respect to the filtration  $\mathcal{F}^{t-k}$  and  $\mathcal{F}^{s-l}$  respectively.

Applying Lemma 33, we have

$$\begin{aligned} &\left\| \max_{\theta \in \Theta} \left| \sum_{s=1}^T \sum_{s-B \leq t \leq s-w-1, t-l > s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top \right|_\infty \right\|_{q/2}^2 \\ &\leq C_{q/2}^2 \ell \sum_{s=1}^T \left\| \max_{\theta \in \Theta} \left| \sum_{s-B \leq t \leq s-w-1, t-l > s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top \right|_\infty \right\|_{q/2}^2 \\ &\leq C_{q/2}^2 \ell \sum_{s=1}^T \left\| \max_{\theta \in \Theta} \left| \sum_{s-B \leq t \leq s-w-1, t-l > s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \right|_\infty \right\|_q^2 \left\| \mathcal{P}^{s-k} X_s^\top \right\|_q^2 \\ &\leq C_{q/2}^2 C_q^2 B T \ell^2 w_{l,q}^2 w_{k,q}^2. \end{aligned}$$

Similarly, it can be obtained that

$$\begin{aligned}
& \left\| \max_{\theta \in \Theta} \left| \sum_{s=1}^T \sum_{s-B \leq t \leq s-w-1, t-l < s-k} a_{s,t,\theta} \mathcal{P}^{t-l} X_t \mathcal{P}^{s-k} X_s^\top \right|_\infty \right\|_{q/2}^2 \\
&= \left\| \max_{\theta \in \Theta} \left| \sum_{t=1}^T \mathcal{P}^{t-l} X_t \sum_{t+w+1 \leq s \leq t+B, t-l < s-k} a_{s,t,\theta} \mathcal{P}^{s-k} X_s^\top \right|_\infty \right\|_{q/2}^2 \\
&\leq C_{q/2}^2 C_q^2 B T \ell^2 w_{l,q}^2 w_{k,q}^2,
\end{aligned}$$

Putting them together, we have

$$\sum_{k=m+1}^{\infty} \sum_{l=0}^m \left\| \max_{\theta \in \Theta} |\text{III}_{k,l}|_\infty \right\|_{q/2} \leq 2C_{q/2} C_q \sqrt{BT} \ell \Omega_{0,q} \Omega_{m,q}. \quad (1.9.35)$$

It remains to deal with the term  $\text{IV}_l$ . On one hand, since  $k \geq m+1$ , it holds that  $t-l = s-k \leq s-m-1$ , which implies  $s \geq t+1+m-l$ . On the other hand,  $s \leq t+B$  by construction. Then

$$\text{IV}_l = \sum_{t=1}^T \mathcal{P}^{t-l} X_t \sum_{s=t+1+(w \vee (m-l))}^{(t+B) \wedge T} a_{s,t,\theta} \mathcal{P}^{t-l} X_s^\top.$$

Since  $\mathcal{P}^{t-l} X_t \sum_{s=t+1+(w \vee (m-l))}^{(t+B) \wedge T} a_{s,t,\theta} \mathcal{P}^{t-l} X_s^\top$  is  $4B$ -independent for  $1 \leq t \leq T$ , we apply the block technique and split the interval  $[1, T]$  into blocks  $L_b = [4(b-1)B+1, 4bB]$ ,  $1 \leq b \leq \lceil \frac{T}{4B} \rceil$ . Define

$$\begin{aligned}
\text{IV}_l^o &= \sum_{b \text{ is odd}} \sum_{t \in L_b} \mathcal{P}^{t-l} X_t \sum_{s=t+w+1}^{t+B} a_{s,t,\theta} \mathcal{P}^{t-l} X_s^\top, \\
\text{IV}_l^e &= \sum_{b \text{ is even}} \sum_{t \in L_b} \mathcal{P}^{t-l} X_t \sum_{s=t+w+1}^{t+B} a_{s,t,\theta} \mathcal{P}^{t-l} X_s^\top.
\end{aligned}$$

By the triangle inequality and Jensen's inequality,

$$\begin{aligned}
& \left\| \max_{\theta \in \Theta} \left| \mathbb{E}_0 \left( \sum_{t \in L_b} \mathcal{P}^{t-l} X_t \sum_{s=t+1+(w \vee (m-l))}^{(t+B) \wedge T} a_{s,t,\theta} \mathcal{P}^{t-l} X_s^\top \right) \right|_\infty \right\|_{q/2} \\
& \leq 2 \left\| \left\| \sum_{t \in L_b} \mathcal{P}^{t-l} X_t \sum_{s=t+1+(w \vee (m-l))}^{(t+B) \wedge T} a_{s,t,\theta} \mathcal{P}^{t-l} X_s^\top \right\|_\infty \right\|_{q/2}^2 \\
& \leq 2 \sum_{t \in L_b} \left\| \left\| \mathcal{P}^{t-l} X_t \right\|_\infty \right\|_q \sum_{s=t+1+(w \vee (m-l))}^{(t+B) \wedge T} \left\| \max_{\theta \in \Theta} |a_{s,t,\theta} \mathcal{P}^{t-l} X_s^\top|_\infty \right\|_q \\
& \leq 8Bw_{l,q} \Omega_{m,q}.
\end{aligned}$$

Then by Lemma 33, we have

$$\left\| \max_{\theta \in \Theta} |IV_l^o|_\infty \right\|_{q/2}^2 \leq C_{q/2}^2 \ell \cdot \left( \frac{T}{8B} + 1 \right) \cdot 64B^2 w_{l,q}^2 \Omega_{m,q}^2 \leq 72C_{q/2}^2 BT \ell \cdot w_{l,q}^2 \Omega_{m,q}^2.$$

A similar inequality holds for  $IV_{k,l}^e$ . Therefore,

$$\sum_{l=0}^m \left\| \max_{\theta \in \Theta} |IV_l|_\infty \right\|_{q/2} \leq 12\sqrt{2}C_{q/2} \sqrt{BT \ell} \Omega_{0,q} \Omega_{m,q}. \quad (1.9.36)$$

Combining (1.9.30) – (1.9.36) together, Lemma 28 follows.  $\square$

**Lemma 29.** Assume  $\left\| |X|_\infty \right\|_{q,\alpha} < \infty$ , for some  $q > 4$  and  $\alpha > 0$ . Denote  $d = |\Theta|$  and  $\ell = 1 \vee \log(pd)$ . Then there exist constants  $C, C_\alpha, C_{q,\alpha} > 0$  such that for any  $x > 0$ ,

$$\begin{aligned}
\mathbb{P} \left( \max_{\theta \in \Theta} |\mathbb{E}_0(Q_{T,\theta}) - \mathbb{E}_0(Q_{T,\theta,m})|_\infty \geq x \right) & \leq C_{q,\alpha} x^{-q/2} \ell^q \left\| |X|_\infty \right\|_{q,\alpha}^q K_{T,B,m} \\
& \quad + Cp^2 d \exp \left( - \frac{x^2 m^{2\alpha}}{C_\alpha \Phi_{4,\alpha}^4 TB} \right),
\end{aligned}$$

where  $K_{T,B,m} = TB^{q/2-1} m^{-\alpha q/2}$  (resp.  $TB^{q/2-1} m^{-\alpha q/2} + T^{q/4-\alpha q/2} B^{q/4}$ ) if  $\alpha > 1/2 - 2/q$  (resp.  $\alpha < 1/2 - 2/q$ ).

*Proof of Lemma 29.* Denote  $M = \lceil T/B \rceil$ . For any  $k = 1, \dots, M$ ,  $\theta \in \Theta$  and  $\tau \geq 0$ , define

$$V_{k,\theta} = \mathbb{E}_0 \sum_{s=(k-1)B+1}^{(kB)\wedge T} \sum_{t=1}^s a_{s-t,\theta} X_t X_s^\top,$$

$$V_{k,\theta,\tau} = \mathbb{E}_0 \sum_{s=(k-1)B+1}^{(kB)\wedge T} \sum_{t=1}^s a_{s-t,\theta} X_{t,\tau} X_{s,\tau}^\top.$$

For  $1 \leq i, j \leq p$ , the  $(i, j)$  entries of  $V_{k,\theta}$  and  $V_{k,\theta,\tau}$  are respectively

$$V_{k,ij,\theta} = \mathbb{E}_0 \sum_{s=(k-1)B+1}^{(kB)\wedge T} \sum_{t=1}^s a_{s-t,\theta} X_{ti} X_{sj},$$

$$V_{k,ij,\theta,\tau} = \mathbb{E}_0 \sum_{s=(k-1)B+1}^{(kB)\wedge T} \sum_{t=1}^s a_{s-t,\theta} X_{ti,\tau} X_{sj,\tau}.$$

Let  $L = \lfloor \log M / \log 2 \rfloor$ ,  $\tau_l = 2^l$  if  $0 \leq l \leq L-1$ ,  $\tau_L = M$ . Define

$$T_{M,l,\theta} = \sum_{k=1}^M (V_{k,\theta,\tau_l B} - V_{k,\theta,\tau_{l-1} B}).$$

We can have the decomposition

$$\mathbb{E}_0(Q_{T,\theta}) - \mathbb{E}_0(Q_{T,\theta,m}) = \sum_{k=1}^M (V_{k,\theta} - V_{k,\theta,MB}) + \sum_{l=1}^L T_{M,l,\theta} + \sum_{k=1}^M (V_{k,\theta,B} - V_{k,\theta,m}). \quad (1.9.37)$$

For the first term, by Lemma 28,

$$\begin{aligned} \left\| \max_{\theta \in \Theta} \left| \sum_{k=1}^M (V_{k,\theta} - V_{k,\theta,MB}) \right|_\infty \right\|_{q/2} &= \left\| \max_{\theta \in \Theta} \left| \mathbb{E}_0 Q_{T,\theta} - \mathbb{E}_0 Q_{T,\theta,MB} \right|_\infty \right\|_{q/2} \\ &\leq C_q \sqrt{BT} \ell \Omega_{0,q} \Omega_{MB,q} \\ &\leq C_{q,\alpha} \ell B^{1/2} T^{1/2-\alpha} \left\| |X|_\infty \right\|_{q,\alpha}^2, \end{aligned}$$

where the last step follows by the fact that  $\Omega_{0,q} \leq \| |X|_\infty \|_{q,\alpha}$  and

$$\Omega_{MB,q} \leq (MB)^{-\alpha} \| |X|_\infty \|_{q,\alpha} \leq C_\alpha T^{-\alpha} \| |X|_\infty \|_{q,\alpha}.$$

By Markov's inequality, we have

$$\begin{aligned} & \mathbb{P}\left(\max_{\theta \in \Theta} \left| \sum_{k=1}^M (V_{k,\theta} - V_{k,\theta,MB}) \right|_\infty \geq x\right) \\ & \leq x^{-q/2} \mathbb{E}\left(\max_{\theta \in \Theta} \left| \sum_{k=1}^M (V_{k,\theta} - V_{k,\theta,MB}) \right|_\infty^{q/2}\right) \\ & \leq C_{q,\alpha} x^{-q/2} \ell^{q/2} T^{q/4 - \alpha q/2} B^{q/4} \| |X|_\infty \|_{q,\alpha}^q. \end{aligned} \quad (1.9.38)$$

For for each  $1 \leq l \leq L$ , define

$$\begin{aligned} Y_{h,l,\theta} &= \sum_{k=2(h-1)\tau_l+1}^{(2h\tau_l) \wedge M} (V_{k,\theta,\tau_l B} - V_{k,\theta,\tau_{l-1}B}), \quad \text{for } 1 \leq h \leq \left\lceil \frac{M}{2\tau_l} \right\rceil, \\ Y_{h,ij,l,\theta} &= \sum_{k=2(h-1)\tau_l+1}^{(2h\tau_l) \wedge M} (V_{k,ij,\theta,\tau_l B} - V_{k,ij,\theta,\tau_{l-1}B}), \quad \text{for } 1 \leq i, j \leq p, \\ R_{M,l,\theta}^e &= \sum_{h \text{ is even}} Y_{h,l,\theta} \quad \text{and} \quad R_{M,l,\theta}^o = \sum_{h \text{ is odd}} Y_{h,l,\theta}. \end{aligned}$$

So we have  $T_{M,l,\theta} = R_{M,l,\theta}^e + R_{M,l,\theta}^o$ . Let  $\lambda_l = l^{-2}/(\pi^2/3)$  if  $1 \leq l \leq L/2$  and  $\lambda_l = (L+1-l)^{-2}/(\pi^2/3)$  if  $2/L \leq l \leq L$ . Then  $\sum_{l=1}^L \lambda_l \leq 1$ . Since  $Y_{h,l,\theta}$  and  $Y_{h',l,\theta}$  are independent for  $|h - h'| > 1$ , by Lemma 34, for any  $x > 0$ ,

$$\begin{aligned} \mathbb{P}(\max_{\theta \in \Theta} |R_{M,l,\theta}^e|_\infty \geq \lambda_l x) &\leq C_q (\lambda_l x)^{-q/2} \ell^{q/2} \sum_{h \text{ is even}} \mathbb{E} \max_{\theta \in \Theta} |Y_{h,l,\theta}|_\infty^{q/2} \\ &\quad + 2p^2 d \exp\left(-\frac{(\lambda_l x)^2}{C \max_{i,j} \max_{\theta \in \Theta} \sum_h \mathbb{E} |Y_{h,ij,l,\theta}|^2}\right). \end{aligned} \quad (1.9.39)$$

By Lemma 28, we have

$$\begin{aligned} \left\| \max_{\theta \in \Theta} |Y_{h,l,\theta}|_{\infty} \right\|_{q/2} &\leq 2C_q \sqrt{B \cdot 2\tau_l B} \ell \Omega_{0,q} \Omega_{\tau_{l-1}B,q} \\ &\leq C_{q,\alpha} B^{1-\alpha} \tau_l^{1/2-\alpha} \ell \left\| |X|_{\infty} \right\|_{q,\alpha}^2. \end{aligned}$$

Since  $\Delta_{0,4,i} \leq \|X_{\cdot i}\|_{4,\alpha} \leq \Phi_{4,\alpha}$  and  $\Delta_{\tau_{l-1}B,4,j} \leq (\tau_{l-1}B)^{-\alpha} \|X_{\cdot j}\|_{4,\alpha} \leq C_{\alpha}(\tau_l B)^{-\alpha} \Phi_{4,\alpha}$ , we also have

$$\|Y_{h,ij,l,\theta}\|_2 \leq C_2 \sqrt{B \cdot 2\tau_l B} \Delta_{0,4,i} \Delta_{\tau_{l-1}B,4,j} \leq C_{\alpha} B^{1-\alpha} \tau_l^{1/2-\alpha} \Phi_{4,\alpha}^2.$$

Then (1.9.39) becomes

$$\begin{aligned} \mathbb{P}\left(\max_{\theta \in \Theta} |R_{M,l,\theta}^e|_{\infty} \geq \lambda_l x\right) &\leq C_{q,\alpha} (\lambda_l x)^{-q/2} \ell^q \frac{M}{\tau_l} B^{q/2-\alpha q/2} \tau_l^{q/4-\alpha q/2} \left\| |X|_{\infty} \right\|_{q,\alpha}^q \\ &\quad + 2p^2 d \exp\left(-\frac{(\lambda_l x)^2}{C_{\alpha} M / \tau_l \cdot B^{2-2\alpha} \tau_l^{1-2\alpha} \Phi_{4,\alpha}^4}\right). \end{aligned}$$

A similar inequality holds for  $R_{M,l,\theta}^o$ . Then

$$\begin{aligned} &\mathbb{P}\left(\max_{\theta \in \Theta} \left| \sum_{l=1}^L T_{M,l,\theta} \right|_{\infty} \geq 2x\right) \\ &\leq \mathbb{P}\left(\max_{\theta \in \Theta} \left| \sum_{l=1}^L R_{M,l,\theta}^e \right|_{\infty} \geq x\right) + \mathbb{P}\left(\max_{\theta \in \Theta} \left| \sum_{l=1}^L R_{M,l,\theta}^o \right|_{\infty} \geq x\right) \\ &\leq \sum_{l=1}^L \mathbb{P}\left(\max_{\theta \in \Theta} |R_{M,l,\theta}^e|_{\infty} \geq \lambda_l x\right) + \sum_{l=1}^L \mathbb{P}\left(\max_{\theta \in \Theta} |R_{M,l,\theta}^o|_{\infty} \geq \lambda_l x\right) \\ &\leq C_{q,\alpha} x^{-q/2} \ell^q T B^{q/2-\alpha q/2-1} \left\| |X|_{\infty} \right\|_{q,\alpha}^q \cdot \text{I} + \text{II} \end{aligned}$$

where

$$\begin{aligned} \text{I} &= \sum_{l=1}^L \frac{\tau_l^c}{\lambda_l^{q/2}}, \quad c = q/4 - \alpha q/2 - 1, \\ \text{II} &= 4p^2 d \sum_{l=1}^L \exp \left( - \frac{x^2 \lambda_l^2 \tau_l^{2\alpha}}{C_\alpha \Phi_{4,\alpha}^4 T B^{1-2\alpha}} \right). \end{aligned}$$

By the definition of  $\tau_l$  and  $\lambda_l$ , elementary calculations show  $\text{I} \leq C$  if  $c < 0$ , and  $\text{II} \leq CM^c$  if  $c > 0$ . Also, since  $\min_{l \geq 1} \lambda_l^2 \tau_l^{2\alpha} > 0$ , we have

$$\text{II} \leq Cp^2 d \exp \left( - \frac{x^2}{C_\alpha \Phi_{4,\alpha}^4 T B^{1-2\alpha}} \right).$$

Therefore, it follows that

$$\begin{aligned} \mathbb{P}(\max_{\theta \in \Theta} \left| \sum_{l=1}^L T_{M,l,\theta} \right|_\infty \geq 2x) &\leq C_{q,\alpha} x^{-q/2} \ell^q \|X\|_{q,\alpha}^q G_{T,B,m} \\ &\quad + Cp^2 d \exp \left( - \frac{x^2 B^{2\alpha}}{C_\alpha \Phi_{4,\alpha}^4 T B} \right). \end{aligned} \quad (1.9.40)$$

where  $G_{T,B,m} = TB^{q/2-\alpha q/2-1}$  (resp.  $T^{q/4-\alpha q/2} B^{q/4}$ ) if  $\alpha > 1/2 - 2/q$  (resp.  $\alpha < 1/2 - 2/q$ ).

For the term  $\sum_{k=1}^M (V_{k,\theta,B} - V_{k,\theta,m})$ , we can observe that  $V_{k,\theta,B} - V_{k,\theta,m}$  and  $V_{k',\theta,B} - V_{k',\theta,m}$  are independent if  $|k - k'| > 2$ . By Lemma 28 and Lemma 34, we have

$$\begin{aligned} &\mathbb{P}(\max_{\theta \in \Theta} \left| \sum_{k=1}^M V_{k,\theta,B} - V_{k,\theta,m} \right| \geq x) \\ &\leq C_q x^{-q/2} \ell^{q/2} \sum_{k=1}^M \mathbb{E} \max_{\theta \in \Theta} |V_{k,\theta,B} - V_{k,\theta,m}|^{q/2} + 6p^2 B \exp \left( - \frac{x^2 m^{2\alpha}}{C_\alpha \Phi_{4,\alpha}^4 T B} \right) \\ &\leq C_q x^{-q/2} \ell^q \|X\|_{q,\alpha}^q T B^{q/2-1} m^{-\alpha q/2} + 6p^2 B \exp \left( - \frac{x^2 m^{2\alpha}}{C_\alpha \Phi_{4,\alpha}^4 T B} \right). \end{aligned} \quad (1.9.41)$$

Putting (1.9.38), (1.9.40) and (1.9.41) together, Theorem 29 follows.  $\square$

**Lemma 30.** Assume  $\|X\|_{q,\alpha} < \infty$ , for some  $q > 4$  and  $\alpha > 0$ . Denote  $d = |\Theta|$  and

$\ell = 1 \vee \log(pd)$ . Then there exist constants  $C, C_\alpha, C_{q,\alpha} > 0$  such that for any  $x > 0$ ,

$$\begin{aligned} & \mathbb{P}\left(\max_{\theta \in \Theta} |Q_{T,\theta} - \mathbb{E}Q_{T,\theta}|_\infty \geq x\right) \\ & \leq C_{q,\alpha} x^{-q/2} \|X\|_{q,\alpha}^q \ell^q V_{T,B} + Cp^2 d \exp\left(-\frac{x^2}{C_\alpha \Phi_{4,\alpha}^4 TB}\right), \end{aligned}$$

where  $V_{T,B} = TB^{q/2-1}$  (resp.  $TB^{q/2-1} + T^{q/4-\alpha q/2} B^{q/4}$ ) if  $\alpha > 1/2 - 2/q$  (resp.  $\alpha < 1/2 - 2/q$ ).

If we consider the quadratic form

$$Q_{T,ij,\theta} = \sum_{1 \leq t \leq s \leq T} a_{s,t,\theta} X_{ti} X_{sj}$$

in terms of the one-dimensional processes  $X_{ti}$  and  $X_{tj}$  for any  $1 \leq i, j \leq p$ , we can follow the proof of Lemma 30 to get Lemma 31 without technical difficulty.

**Lemma 31.** Assume  $\Phi_{q,\alpha} < \infty$ , for some  $q > 4$  and  $\alpha > 0$ . Denote  $d = |\Theta|$  and  $\ell^* = 1 \vee \log d$ . Then there exist constants  $C, C_\alpha, C_{q,\alpha} > 0$  such that for any  $x > 0$ ,

$$\begin{aligned} & \mathbb{P}\left(\max_{\theta \in \Theta} |Q_{T,ij,\theta} - \mathbb{E}Q_{T,ij,\theta}|_\infty \geq x\right) \\ & \leq C_{q,\alpha} x^{-q/2} \|X_{\cdot i}\|_{q,\alpha}^{q/2} \|X_{\cdot j}\|_{q,\alpha}^{q/2} (\ell^*)^q V_{T,B} + Cd \exp\left(-\frac{x^2}{C_\alpha \Phi_{4,\alpha}^4 TB}\right), \end{aligned}$$

where  $V_{T,B}$  is defined the same as in Lemma 30.

*Proof of Lemma 30.* We can follow the proof of Lemma 29. Note that we shall write (1.9.37) as

$$\mathbb{E}_0(Q_{T,\theta}) = \sum_{k=1}^M (V_{k,\theta} - V_{k,\theta,MB}) + \sum_{l=1}^L T_{M,l,\theta} + \sum_{k=1}^M V_{k,\theta,B},$$

where the first two terms can be dealt with similarly as (1.9.38) and (1.9.40) and the two bounds apply. For the different term  $\sum_{k=1}^M V_{k,\theta,B}$ , by the definition of  $V_{k,\theta,B}$ , we can observe that  $V_{k,\theta,B}$  and  $V_{k',\theta,B}$  are independent if  $|k - k'| > 2$ . By Lemma 27 (iv) and

Lemma 34, we have

$$\begin{aligned}
& \mathbb{P}(\max_{\theta \in \Theta} |\sum_{k=1}^M V_{k,\theta,B}| \geq x) \\
& \leq C_q x^{-q/2} \ell^{q/2} \sum_{k=1}^M \mathbb{E} \max_{\theta \in \Theta} |V_{k,\theta,B}|^{q/2} + 6p^2 B \exp\left(-\frac{x^2}{C_\alpha \Phi_{4,\alpha}^4 T B}\right) \\
& \leq C_q x^{-q/2} \ell^q \|X\|_{q,\alpha}^q T B^{q/2-1} + 6p^2 B \exp\left(-\frac{x^2}{C_\alpha \Phi_{4,\alpha}^4 T B}\right).
\end{aligned}$$

□

**Lemma 32.** Assume  $\|X\|_{q,\alpha} < \infty$ , for some  $q > 4$  and  $\alpha > 0$ . Let  $n = T/B$  and  $\ell = 1 \vee \log(Bp)$ . Then there exists constants  $C, C_q > 0$  such that for any  $x > 0$ ,

$$\begin{aligned}
& \mathbb{P}(\max_{\theta \in \Theta} |\mathbb{E}_0(\hat{F}_m(\theta)) - \mathbb{E}_0(\hat{S}_m(\theta))|_\infty \geq x) \\
& \leq C_q x^{-q/2} (\log(Bp))^q T^{-q/2+1} B^{q/4-1} m^{q/4} \|X\|_{q,\alpha}^q + C p^2 B \exp\left(-\frac{x^2 T}{C \Phi_{4,\alpha}^4 m}\right).
\end{aligned}$$

*Proof of Lemma 32.* Write  $2\pi T(\hat{F}_m(\theta) - \hat{S}_m(\theta)) = \text{I} + \text{II} + \text{III}$ , where

$$\begin{aligned}
\text{I} &= \sum_{r=1}^n \sum_{t=rB-m+1}^{rB} \sum_{s=(r-1)B}^{rB} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}, \\
\text{II} &= \sum_{r=1}^n \sum_{s=rB-m+1}^{rB} \sum_{t=(r-1)B}^{rB} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}, \\
\text{III} &= \sum_{r=1}^n \sum_{t,s=rB-m+1}^{rB} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}.
\end{aligned}$$

For each  $1 \leq r \leq n$ , define

$$\begin{aligned} Y_r &= \sum_{t=rB-m+1}^{rB} \sum_{s=(r-1)B}^{rB} X_{t,m} X_{s,m}^\top e^{-\iota(t-s)\theta}, \\ Y_{ij,r} &= \sum_{t=rB-m+1}^{rB} \sum_{s=(r-1)B}^{rB} X_{ti,m} X_{sj,m} e^{-\iota(t-s)\theta}, \\ R^o &= \sum_{r \text{ is odd}} Y_r, \quad R^e = \sum_{r \text{ is even}} Y_r. \end{aligned}$$

So we can write  $I = R^o + R^e$ . Since  $Y_r$  and  $Y_{r'}$  are independent for  $|r - r'| > 1$ . By Lemma 34,

$$\begin{aligned} &\mathbb{P}(\max_{\theta \in \Theta} |R^o - \mathbb{E}R^o|_\infty \geq x) \\ &\leq C_q x^{-q/2} (\log(Bp))^{q/2} \sum_{r \text{ is odd}} \mathbb{E} \max_{\theta \in \Theta} |Y_r - \mathbb{E}Y_r|_\infty^{q/2} + 2p^2 B \exp\left(-\frac{x^2}{C\sigma^2}\right), \quad (1.9.42) \end{aligned}$$

where  $\sigma^2 = \max_{i,j} \max_{\theta \in \Theta} \sum_{r \text{ is odd}} \mathbb{E}|Y_{ij,r} - \mathbb{E}Y_{ij,r}|^2$ . By similar arguments in deriving Lemma 27 (iv) and the fact that  $\Omega_{0,q} \leq \| |X|_\infty \|_{q,\alpha}$ , we have

$$\left\| \max_{\theta \in \Theta} |Y_r - \mathbb{E}Y_r|_\infty \right\|_{q/2} \leq C_q \ell \sqrt{Bm} \| |X|_\infty \|_{q,\alpha}^2.$$

Applying similar arguments of deriving equation (41) in Lemma 8 of [111] and the fact that  $\Delta_{0,4,i} \leq \Phi_{4,\alpha}$ , we have

$$\|Y_{ij,r} - \mathbb{E}Y_{ij,r}\|_2 \leq C\sqrt{Bm}\Phi_{4,\alpha}^2.$$

A similar inequality holds for  $R^e$ . Therefore

$$\begin{aligned} &\mathbb{P}(\max_{\theta \in \Theta} |I - \mathbb{E}I|_\infty \geq 2x) \\ &\leq \mathbb{P}(\max_{\theta \in \Theta} |R^o - \mathbb{E}R^o|_\infty \geq x) + \mathbb{P}(\max_{\theta \in \Theta} |R^e - \mathbb{E}R^e|_\infty \geq x) \end{aligned}$$

$$\leq C_q x^{-q/2} (\log(Bp))^q T B^{q/4-1} m^{q/4} \|X\|_{\infty, \alpha}^q + 2p^2 B \exp\left(-\frac{x^2}{C\Phi_{4,\alpha}^4 T m}\right).$$

The same bound also applies to II and III, thus the result follows.  $\square$

**Lemma 33.** *Let  $D_t, 1 \leq t \leq T$ , be  $p$ -dimensional martingale difference vectors with  $p \geq 1$ . Let  $a_{1,\theta}, \dots, a_{T,\theta} \in \mathbb{C}$  that depend on  $\theta$  for  $\theta \in \Theta$ . Denote  $L_{T,\theta} = \sum_{t=1}^T a_{t,\theta} D_t$ ,  $d = |\Theta|$  and  $\ell = 1 \vee \log(pd)$ . Let  $q \geq 2$ . We have*

$$\left\| \max_{\theta \in \Theta} |L_{T,\theta}|_{\infty} \right\|_q^2 \leq C_q \cdot \ell \sum_{t=1}^T \left\| \max_{\theta \in \Theta} |a_{t,\theta} D_t|_{\infty} \right\|_q^2,$$

where  $C_q$  is a positive constant depending on  $q$  only.

*Proof of Lemma 33.* Let  $Z_t = (a_{t,\theta_1} D_t^\top, a_{t,\theta_2} D_t^\top, \dots, a_{t,\theta_d} D_t^\top)^\top$ , then  $Z_t, 1 \leq t \leq T$  are  $(pd)$ -dimensional martingale difference vectors. It follows from Lemma D.3 in [114] that

$$\begin{aligned} \left\| \max_{\theta \in \Theta} |L_{T,\theta}|_{\infty} \right\|_q^2 &= \left\| |Z_1 + Z_2 + \dots + Z_T|_{\infty} \right\|_q^2 \\ &\leq C_q \cdot \ell \sum_{t=1}^T \left\| |Z_t|_{\infty} \right\|_q^2 \\ &\leq C_q \cdot \ell \sum_{t=1}^T \left\| \max_{\theta \in \Theta} |a_{t,\theta} D_t|_{\infty} \right\|_q^2. \end{aligned}$$

$\square$

**Lemma 34.** *Let  $Y_{t,\theta}, 1 \leq t \leq T$ , be independent random vectors in  $\mathbb{R}^p$  indexed by  $\theta \in \Theta$ . Assume  $\mathbb{E}Y_{tj,\theta} = 0$  and  $\|Y_{tj,\theta}\|_q < \infty$  for some  $q \geq 2$  and all  $1 \leq t \leq T, 1 \leq j \leq p, \theta \in \Theta$ . Denote  $S_{T,\theta} = \sum_{t=1}^T Y_{t,\theta}$ ,  $d = |\Theta|$  and  $\ell = 1 \vee \log(pd)$ . Then for any  $x > 0$ ,*

$$\mathbb{P}(\max_{\theta \in \Theta} |S_{T,\theta}|_{\infty} \geq x) \leq C_q x^{-q} \ell^q \sum_{t=1}^T \mathbb{E} \max_{\theta \in \Theta} |Y_{t,\theta}|_q^q + 2pd \exp(-\frac{x^2}{C\sigma^2}),$$

where  $\sigma^2 = \max_{1 \leq j \leq p} \max_{\theta \in \Theta} \sum_{t=1}^T \mathbb{E}|Y_{tj,\theta}|^2$ ,  $C > 0$  is a universal constant and  $C_q > 0$  depends on  $q$  only.

*Proof of Lemma 34.* Let  $U_t = (Y_{t,\theta_1}^\top, Y_{t,\theta_2}^\top, \dots, Y_{t,\theta_d}^\top)^\top$  and  $W_T = \sum_{t=1}^T U_t$ . Then  $U_t, 1 \leq t \leq T$  are independent random vectors in  $\mathbb{R}^{pd}$ . Since  $\max_{\theta \in \Theta} |Y_{t,\theta}|_\infty = |U_t|_\infty$ , by Lemma D.4 in [114], we have

$$\mathbb{P}(|W_T|_\infty \geq x) \leq C_q x^{-q} \ell^q \sum_{t=1}^T \mathbb{E} \max_{\theta \in \Theta} |Y_{t,\theta}|_\infty^q + 2pd \exp\left(-\frac{x^2}{C\sigma^2}\right),$$

where  $\sigma^2 = \max_{1 \leq k \leq pd} \sum_{t=1}^T \mathbb{E} U_{tk}^2 = \max_{1 \leq j \leq p} \max_{\theta \in \Theta} \sum_{t=1}^T \mathbb{E} |Y_{tj,\theta}|^2$ . Thus the result follows in view of  $\max_{\theta \in \Theta} |S_{T,\theta}|_\infty = |W_T|_\infty$ .  $\square$

**Lemma 35.** Let  $Y_1, \dots, Y_n \in \mathbb{R}^p$  be independent random vectors with  $\mathbb{E} Y_{ij} = 0$  and  $\mathbb{E} Y_{ij}^2 < \infty$  for  $1 \leq i \leq n$  and  $1 \leq j \leq p$ . Let  $Z_1, \dots, Z_n \in \mathbb{R}^p$  be independent random vectors such that  $Z_i \sim N(0, \mathbb{E} Y_i Y_i^\top)$ . Define  $\mathcal{A}$  as the collection of all sets  $A$  of the form

$$A = \{v \in \mathbb{R}^p : a_j \leq v_j \leq b_j \text{ for } 1 \leq j \leq p\}.$$

Let  $a \geq 1, b > 0$  be two constants. Assume that  $\frac{1}{n} \sum_{i=1}^n \mathbb{E} Y_{ij}^2 \geq b$  and  $\max_{1 \leq i \leq n} \mathbb{E} |Y_{ij}|^{2+k} \leq a^k$  for  $1 \leq j \leq p$  and  $k = 1, 2$ . Let  $C_n \geq 1$  be a sequence of constants. Further assume that  $\mathbb{E} (\max_{1 \leq j \leq p} |Y_{ij}|)^q \leq c C_n^q$  for some  $q > 2$  and some constant  $c > 0$ , then

$$\sup_{A \in \mathcal{A}} \left| \mathbb{P} \left( \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \in A \right) - \mathbb{P} \left( \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \in A \right) \right| \leq C \left\{ \left( \frac{a^2 C_n^2 (\log(pn))^7}{n} \right)^{1/6} + \left( \frac{a^2 C_n^2 (\log(pn))^3}{n^{1-2/q}} \right)^{1/3} \right\},$$

where  $C$  is a positive constant that depends on  $c, b$  and  $q$ .

*Proof of Lemma 35.* The proof follows from Proposition 2.1 case (E.2) in [24] by setting

$$\bar{L}_n = \left\{ a C_n + \frac{a^2 C_n^2}{n^{1/2-2/q} (\log p)^{1/2}} \right\}.$$

$\square$

**Lemma 36** (Nazarov’s inequality, [65]). *Let  $Z = (Z_1, \dots, Z_p)^\top$  be a centered Gaussian random vector in  $\mathbb{R}^p$  such that  $\mathbb{E}(Z_j^2) \geq c$  for all  $j = 1, \dots, p$  and some constant  $c > 0$ . Then for every  $z \in \mathbb{R}^p$  and  $a > 0$ ,*

$$\mathbb{P}(Z \leq z + a) - \mathbb{P}(Z \leq z) \leq Ca\sqrt{\log p},$$

*where  $C$  is a positive constant depending on  $c$  only.*

## 1.10 Acknowledgements

Chapter 1, in part, has been submitted for publication of the material “Spectral Inference for High Dimensional Time Series” by Chi Zhang and Danna Zhang, as it appears in IEEE Transactions on Information Theory, 2025. The dissertation author was the primary investigator and author of this paper.

# Chapter 2

## Probability and moment inequalities for quadratic forms in independent random variables with fat tails

### 2.1 Introduction

Probability inequalities for quadratic forms play an important role in the study of properties of second-order statistics. The celebrated Hanson-Wright inequality, first established by [45], provides a concentration result for quadratic forms of symmetric sub-Gaussian random variables. [105] relaxed the symmetry assumption on the distributions of random variables. The version widely used in many applications is a refined one presented by [84]. To clarify, let  $\varepsilon_1, \dots, \varepsilon_n$  be independent mean-zero random variables with finite sub-Gaussian norm, i.e.,  $\|\varepsilon_i\|_{\psi_2} := \sup_{q \geq 1} q^{-1/2} \|\varepsilon_i\|_q \leq K$ , for  $1 \leq i \leq n$ , where  $\|\varepsilon_i\|_q = (\mathbb{E}|\varepsilon_i|^q)^{1/q}$ . Let  $A = (a_{ij})_{i,j=1}^n$  be an  $n \times n$  matrix, and write  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^\top$ . Then for every  $y > 0$ , it holds that

$$\mathbb{P}(|\varepsilon^\top A \varepsilon - \mathbb{E} \varepsilon^\top A \varepsilon| \geq y) \leq 2 \exp \left( - \min \left( \frac{y^2}{CK^4 |A|_F^2}, \frac{y}{CK^2 \rho(A)} \right) \right), \quad (2.1.1)$$

where  $C$  is a universal constant,  $|A|_F$  and  $\rho(A)$  are respectively the Frobenius (or Hilbert-Schmidt) norm and spectral norm of  $A$ . The inequality (2.1.1) is sharper than the version in [45] and [105], with  $\rho(A) = \rho((a_{ij})_{i,j})$  replacing the term  $\rho((|a_{ij}|)_{i,j})$  in the earlier results.

[5] provided an alternative proof of (2.1.1) for diagonal-free matrices  $A$  by applying the decoupling technique in [28].

Several variants of concentration inequalities for quadratic forms have been studied recently. For example, [51] considered the tail probability for quadratic forms of sub-Gaussian random vectors, rather than sub-Gaussian components, with the matrix  $A$  being positive-semidefinite. [91] examined the case where the random vector admits certain finite exponential moments. [100] established an exponential probability bound for quadratic forms involving vectors that satisfy the convex concentration property, which was later improved by [1] by eliminating the logarithmic factors. Despite the extensive literature on probability inequalities for quadratic forms and their applications, there has been very limited theoretical investigation into the counterparts when the random variables are non-sub-Gaussian or lack finite exponential moments. Evidence for fat tails has increasingly been observed across a wide variety of research areas. One primary goal of our paper is to establish a sharp probability inequality for quadratic forms in fat-tailed random variables that only allow finite polynomial moments.

For moment inequalities, [4] (cf. Lemma B.26) provided a Rosenthal-type inequality for quadratic forms with i.i.d.  $\varepsilon_i$  and coefficient matrix  $A \in \mathbb{R}^{n \times n}$ :

$$\mathbb{E}|\varepsilon^\top A \varepsilon - \mathbb{E}\varepsilon^\top A \varepsilon|^q \leq C_q \left\{ \|\varepsilon_0\|_4^{2q} [\text{tr}(AA^\top)]^{q/2} + \|\varepsilon_0\|_{2q}^{2q} \text{tr}[(AA^\top)^{q/2}] \right\}, \quad (2.1.2)$$

assuming that  $\|\varepsilon_0\|_{2q} < \infty$  for some  $q \geq 1$  and  $\|\varepsilon_0\|_2 = 1$ , where  $C_q > 0$  is a constant depending on  $q$  only. Let  $s_j^2$  be the eigenvalues of  $A^\top A$  and define  $s = (s_1, \dots, s_n)^\top$ . Since  $\text{tr}(AA^\top) = |A|_F^2$  and  $\text{tr}[(AA^\top)^{q/2}] = \sum_{j=1}^n |s_j|^q$ , (2.1.2) is equivalent to

$$\mathbb{E}|\varepsilon^\top A \varepsilon - \mathbb{E}\varepsilon^\top A \varepsilon|^q \leq C_q \left( \|\varepsilon_0\|_4^{2q} |A|_F^q + \|\varepsilon_0\|_{2q}^{2q} |s|_q^q \right).$$

Theorem 2.1 in [19] presented a similar result by involving the term  $|s|_4$  rather than  $|s|_q$ .

Both of these moment bounds are not sharp for diagonal-free matrices  $A$ , as the moment order is increased to  $2q$  in the bound. Theorem 3.8 of [27] (by letting  $m = 2$  therein) states a sharper version:

$$\mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q \left\{ \mu_2^{2q} |A|_F^q + \mu_2^q \mu_q^q \sum_j \left( \sum_i a_{ij}^2 \right)^{q/2} + \mu_q^{2q} \sum_{i,j} |a_{ij}|^q \right\}, \quad (2.1.3)$$

where  $A$  is symmetric and diagonal-free,  $\varepsilon_i$ 's are independent symmetric random variables with finite  $q$ -th moment  $\mu_q := \max_i \|\varepsilon_i\|_q$ , for  $2 < q < 4$ . In our paper, we shall establish a moment inequality that relaxes the assumption of symmetric distributions on  $\varepsilon_i$  and generalizes the result (2.1.3) to arbitrary order  $q \geq 2$ .

For notational simplicity, we use  $\sum_i$  to indicate the summation sign  $\sum_{i=1}^n$ . Throughout,  $C, C_1, c_1, C_2, c_2, \dots$  denote positive constants whose values may differ from place to place. A constant with a subscript is used to highlight its dependence on that subscript. For two sequences of positive numbers  $(a_n)$  and  $(b_n)$ , we write  $a_n \lesssim b_n$  if there exists some constant  $C > 0$  such that  $a_n \leq C b_n$  for all large  $n$ , and we write  $a_n \gtrsim b_n$  if there exists some constant  $C > 0$  such that  $a_n \geq C b_n$  for all large  $n$ . For any random variable  $X$  and  $q \geq 1$ , we define  $\|X\|_q = (\mathbb{E}|X|^q)^{1/q}$ .

## 2.2 Main Result

Given that  $\varepsilon^\top A \varepsilon = \varepsilon^\top (\frac{1}{2}(A + A^\top)) \varepsilon$ , we assume without loss of generality that  $A$  is symmetric. For any symmetric matrix  $A = (a_{ij})_{i,j=1}^n$  and  $p, q \geq 1$ , we define the entrywise  $L_{p,q}$  norm as

$$\|A\|_{p,q} = \left[ \sum_j \left( \sum_i |a_{ij}|^p \right)^{q/p} \right]^{1/q}.$$

It is straightforward to verify that  $\|A\|_{2,2}$  represents the Frobenius norm of  $A$  and  $\|A\|_{q,q}$  denotes the entrywise  $q$ -norm of  $A$ .

**Theorem 37.** *Let  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  be a sequence of independent random variables, and define*

$\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^\top$ . Suppose  $A = (a_{ij})_{i,j=1}^n$  is a diagonal-free symmetric matrix. Assume  $\mathbb{E}\varepsilon_i = 0$  for all  $1 \leq i \leq n$  and  $\mu_q := \max_{1 \leq i \leq n} \|\varepsilon_i\|_q < \infty$  for some  $q \geq 2$ . Then for any  $y \gtrsim \mu_2 \mu_q \|A\|_{2,2}$ ,

$$\mathbb{P}(|\varepsilon^\top A \varepsilon| \geq y) \lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q + \mu_q^{2q} \|A\|_{q,q}^q}{y^q} + \exp\left(-\frac{y}{C_q \mu_2^2 \|A\|_{2,2}}\right), \quad (2.2.1)$$

where the positive constant  $C_q$  and the constant in  $\lesssim$  both depend on  $q$  only.

Theorem 37 above provides a Nagaev-type inequality, which suggests two types of bounds: If  $y$  is around the order  $\mu_2 \mu_q \|A\|_{2,2}$ , one can use the sub-exponential-type tail bound  $\exp\{-y/(C_q \mu_2^2 \|A\|_{2,2})\}$ . If  $y$  is larger, the algebraic decay tail  $y^{-q}(\mu_2^q \mu_q^q \|A\|_{2,q}^q + \mu_q^{2q} \|A\|_{q,q}^q)$  dominates. If  $\mu_2$  and  $\mu_q$  are of the same order, the terms involving  $\|A\|_{2,q}^q$  dominates the one involving  $\|A\|_{q,q}^q$ .

If  $A$  is not diagonal free, we shall deal with the diagonal sum separately. Assume  $\mu_q < \infty$  for some  $q \geq 2$ . For  $2 \leq q < 4$ , by Corollary 1.6 in [64] and Markov's inequality, we have

$$\begin{aligned} \mathbb{P}\left(\left|\sum_i a_{ii}(\varepsilon_i^2 - \mathbb{E}\varepsilon_i^2)\right| \geq y\right) &\lesssim \sum_i \mathbb{P}(|a_{ii}(\varepsilon_i^2 - \mathbb{E}\varepsilon_i^2)| \geq y/2) + \frac{\mu_q^q \sum_i |a_{ii}|^{q/2}}{y^{q/2}} \\ &\lesssim \frac{\mu_q^q \sum_i |a_{ii}|^{q/2}}{y^{q/2}}. \end{aligned}$$

If  $q \geq 4$ , by Lemma 39, it follows that

$$\mathbb{P}\left(\left|\sum_i a_{ii}(\varepsilon_i^2 - \mathbb{E}\varepsilon_i^2)\right| \geq y\right) \lesssim \frac{\mu_q^q \sum_i |a_{ii}|^{q/2}}{y^{q/2}} + \exp\left(-\frac{y^2}{C_q \mu_4^2 \sum_i a_{ii}^2}\right).$$

**Remark 5.** We shall discuss the sharpness of our result. Consider the special case in which  $\varepsilon_i$ 's are i.i.d. symmetric random variables with  $\|\varepsilon_i\|_2 = 1$  and the tail probability  $\mathbb{P}(\varepsilon_i \geq y) = y^{-q} \ell(y)$ , for  $y \geq y_0$ , where  $\ell(y) = (\log y)^{-2}$  and  $q > 2$ . Then (2.2.1) implies

for  $y \geq n$ ,

$$\mathbb{P}\left(\sum_{i \neq j} \varepsilon_i \varepsilon_j \geq y\right) \lesssim \frac{n^{q/2+1}}{y^q} + \exp\left(-\frac{y}{C_q n}\right). \quad (2.2.2)$$

Next we aim to establish a lower bound of  $\mathbb{P}(\sum_{i \neq j} \varepsilon_i \varepsilon_j \geq y)$ . Denote  $S_n = \sum_{i=1}^n \varepsilon_i$ . By Theorem 1.9 of [64], we have the expansion: for a sequence  $y_n \geq \sqrt{n}$ , as  $n \rightarrow \infty$ ,

$$\frac{\mathbb{P}(S_n \geq y_n)}{ny_n^{-q}\ell(y_n) + 1 - \Phi(y_n/\sqrt{n})} \rightarrow 1, \quad (2.2.3)$$

where  $\Phi(\cdot)$  is the cumulative distribution function of standard normal distribution. Let  $(\varepsilon'_i)_{i=1}^n$  be an i.i.d. copy of  $(\varepsilon_i)_{i=1}^n$  and  $S'_n = \sum_{i=1}^n \varepsilon'_i$ . By Theorem 1 in [28], there exists some constant  $C$  independent of  $n$  such that

$$\mathbb{P}\left(\left|\sum_{i \neq j} \varepsilon_i \varepsilon_j\right| \geq y\right) \geq C\mathbb{P}\left(\left|\sum_{i \neq j} \varepsilon_i \varepsilon'_j\right| \geq y/C\right), \quad (2.2.4)$$

Note that for  $y \geq n$ ,

$$\begin{aligned} \mathbb{P}\left(\sum_{i \neq j} \varepsilon_i \varepsilon'_j \geq y\right) &\geq \mathbb{P}(S_n S'_n \geq 2y) - \mathbb{P}\left(\sum_i \varepsilon_i \varepsilon'_i \geq y\right) \\ &\geq \mathbb{P}(S_n \geq y/\sqrt{n})\mathbb{P}(S'_n \geq 2\sqrt{n}) - \mathbb{P}\left(\sum_i \varepsilon_i \varepsilon'_i \geq y\right). \end{aligned}$$

Applying (2.2.3) to  $S_n$  and  $S'_n$ , we can obtain

$$\mathbb{P}(S_n \geq y/\sqrt{n}) \gtrsim \frac{n^{q/2+1}\ell(y)}{y^q} + 1 - \Phi(y/n) \quad (2.2.5)$$

and  $\mathbb{P}(S'_n \geq 2\sqrt{n}) = (1 - \Phi(2))(1 + o(1))$ . In addition, by elementary calculation, it holds that  $\mathbb{P}(\varepsilon_i \varepsilon'_i \geq y) = C_{q,y_0} y^{-q} \ell(y)(1 + o(1))$  for  $y \geq y_0^2$ , where  $C_{q,y_0}$  is a positive constant depending on  $q$  and  $y_0$ . Applying Theorem 1.9 of [64] to the sum of the i.i.d. symmetric

random variables  $\varepsilon_i \varepsilon'_i$ , we have

$$\mathbb{P}\left(\sum_i \varepsilon_i \varepsilon'_i \geq y\right) \lesssim \frac{n\ell(y)}{y^q} + 1 - \Phi(y/\sqrt{n}),$$

whose order is negligible compared to the right-hand side of (2.2.5) when  $y \geq n$ . Hence, it implies that

$$\mathbb{P}\left(\sum_{i \neq j} \varepsilon_i \varepsilon'_j \geq y\right) \gtrsim \frac{n^{q/2+1}\ell(y)}{y^q} + 1 - \Phi(y/n). \quad (2.2.6)$$

On the other hand, we also have

$$\begin{aligned} \mathbb{P}\left(\sum_{i \neq j} \varepsilon_i \varepsilon'_j \geq y\right) &\geq \mathbb{P}(S_n \geq \sqrt{2y})\mathbb{P}(S'_n \geq \sqrt{2y}) - \mathbb{P}\left(\sum_i \varepsilon_i \varepsilon'_i \geq y\right) \\ &\gtrsim \frac{n^2\ell^2(y)}{y^q} + \left(1 - \Phi(\sqrt{2y/n})\right)^2. \end{aligned} \quad (2.2.7)$$

Combining (2.2.4), (2.2.6) and (2.2.7), it follows

$$\mathbb{P}\left(\sum_{i \neq j} \varepsilon_i \varepsilon_j \geq y\right) \gtrsim \frac{n^{q/2+1}\ell(y)}{y^q} + \left(1 - \Phi(\sqrt{Cy/n})\right)^2,$$

which, compared to (2.2.2), suggests that (2.2.1) is optimal up to a multiplicative logarithmic term.

**Remark 6.** The infinite-dimensional version of (2.2.1) is also true since the right-hand side of (2.2.1) is uniform in  $n$ , provided that the matrix norms  $\|A\|_{2,2}$ ,  $\|A\|_{2,q}$  and  $\|A\|_{q,q}$  converge. This extension together with the discussion on how to deal with the diagonal sum allows us to study the concentration of the autocovariance function for a linear process of the form  $X_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}$ , where  $\sum_{j=0}^{\infty} |b_j| < \infty$  and  $\varepsilon_i$  are i.i.d. mean-zero random variables with  $\mathbb{E}|\varepsilon_i|^q < \infty$  for some  $q \geq 2$ . For example, if we consider the autovariance

estimate at lag 0, i.e.,  $\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n X_t^2$ , it can be constructed that

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=-\infty}^n \sum_{j=-\infty}^n a_{ij} \varepsilon_i \varepsilon_j \text{ where } a_{ij} = \sum_{k=\max\{i,j,1\}}^n b_{k-i} b_{k-j}.$$

By imposing a decay rate of  $b_j$ , the required matrix norms are computable without much difficulty.

**Remark 7.** Another direct application of Theorem 37 is investigating the concentration of sample means for some important nonlinear processes. The Volterra system is a natural model to study nonlinear systems and plays an important role in the nonlinear system theory ([17], [72], [86], [9]). Here we consider the second-order Volterra process

$$X_t = \sum_{u_1=0}^{\infty} h_1(u_1) \varepsilon_{t-u_1} + \sum_{u_1, u_2=0}^{\infty} h_2(u_1, u_2) \varepsilon_{t-u_1} \varepsilon_{t-u_2}.$$

The functions  $h_1, h_2$  are called Volterra kernels. Assume that  $h_1(u) \equiv 0$ ,  $\varepsilon_i$  are i.i.d with mean zero and  $h_2(u_1, u_2)$  is symmetric and it equals zero if  $u_1 = u_2$ . We have

$$\hat{\mu} = \frac{1}{n} \sum_{t=1}^n X_t = \frac{1}{n} \sum_{i=-\infty}^n \sum_{j=-\infty}^n a_{ij} \varepsilon_i \varepsilon_j \text{ where } a_{ij} = \sum_{k=\max\{i,j,1\}}^n h_2(k-i, k-j).$$

and we can use Theorem 37 to study the concentration behavior of  $\hat{\mu}$ . □

Next we shall establish a sharp moment inequality. For any square matrix  $A$ , define the operators

$$\mathcal{D}(A) = A^2 - \text{diag}(A^2), \quad \text{and} \quad \mathcal{D}^k(A) = \underbrace{\mathcal{D} \circ \mathcal{D} \circ \cdots \circ \mathcal{D}}_k(A) \text{ for } k \in \mathbb{N}.$$

Conventionally, we set  $\mathcal{D}^0(A) = A$ . Theorem 38 below delivers a moment inequality for quadratic forms of random variables in the presence of fat tails.

**Theorem 38.** Let  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  be a sequence of independent random variables, and define  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^\top$ . Suppose  $A = (a_{ij})_{i,j=1}^n$  is a diagonal-free symmetric matrix. Assume  $\mathbb{E}\varepsilon_i = 0$  for all  $1 \leq i \leq n$  and  $\mu_q := \max_{1 \leq i \leq n} \|\varepsilon_i\|_q < \infty$  for some  $q \geq 2$ . Let  $K = \lfloor \log_2(q) \rfloor - 1$ . Then we have

$$\begin{aligned} \mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q \sum_{k=0}^K \prod_{h=1}^k \mu_2^{q/2^{h-1}} \cdot \left( \mu_2^{q/2^{k-1}} \|\mathcal{D}^k(A)\|_{2,2}^{q/2^k} + \mu_2^{q/2^k} \mu_{q/2^k}^{q/2^k} \|\mathcal{D}^k(A)\|_{2,q/2^k}^{q/2^k} \right. \\ \left. + \mu_{q/2^k}^{q/2^{k-1}} \|\mathcal{D}^k(A)\|_{q/2^k,q/2^k}^{q/2^k} \right), \end{aligned} \quad (2.2.8)$$

where  $C_q$  is a positive constant depending on  $q$  only.

**Remark 8.** If  $2 \leq q < 4$ , then  $K = 0$  and the result (2.2.8) is reduced to (2.1.3) exactly. Our Theorem 38 extends (2.1.3) to higher moment orders by allowing  $q \geq 2$  in general and in the meantime relaxes the assumption of symmetric random variables. For each doubling of  $q$ , there will be one more layer to add up in the moment bound, each of which is characterized by the operator  $\mathcal{D}^k$  and balances between the  $L_{2,2}$ ,  $L_{2,q/2^k}$  and  $L_{q/2^k,q/2^k}$  norms. If  $\varepsilon_i$  are i.i.d., by a careful check of the proof for the case  $2 \leq q < 4$ , we can conclude (2.2.8) is optimal since the upper bounds in (2.3.19) and (2.3.20) are both optimal.

**Remark 9.** We consider the special case where  $A = (a_{ij})_{i,j=1}^n$  with  $a_{ij} = 1$  if  $i \neq j$  and  $a_{ii} = 0$  for  $1 \leq i \leq n$ . By elementary calculation, we can get that for  $p, q \geq 1$ ,

$$\|\mathcal{D}^k(A)\|_{p,q} \asymp n^{2^k + \frac{1}{p} + \frac{1}{q} - 1}.$$

The moment inequality (2.1.2) suggests  $\mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q \mu_{2q}^{2q} n^q$ , as the second term dominates the first in this case. By Theorem 38, it can be computed that for  $2 \leq q < 4$ ,

$$\mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q (\mu_2^{2q} n^q + \mu_2^q \mu_q^q n^{q/2+1} + \mu_q^{2q} n^2).$$

and for  $4 \leq q < 8$ ,

$$\mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q(\mu_2^{2q} \mu^q + \mu_2^{3q/2} \mu_{q/2}^{q/2} n^{3q/4+1} + \mu_2^q \mu_{q/2}^q n^{q/2+2} + \mu_2^q \mu_q^q n^{q/2+1} + \mu_q^{2q} n^2).$$

The bound for other cases of  $q$  can be presented similarly. In comparison, our result provides a sharper combination of various orders of moments and different norms of the coefficient matrix.

## 2.3 Proofs

In this section, we provide the proofs of Theorem 37 and Theorem 38, beginning with the statement of two useful lemmas.

**Lemma 39** (cf. Corollary 1.8 in [64]). *Let  $X_1, \dots, X_n$  be independent random variables with  $\mathbb{E}X_i = 0$  and  $\|X_i\|_q < \infty$  for some  $q \geq 2$ . Let  $S_n = \sum_{i=1}^n X_i$ . Then for every  $x > 0$ ,*

$$\mathbb{P}(S_n \geq x) \leq \frac{C_1 \sum_{i=1}^n \mathbb{E}|X_i|^q}{x^q} + \exp\left(-\frac{C_2 x^2}{\sum_{i=1}^n \mathbb{E}X_i^2}\right),$$

where  $C_1 = (1 + 2/q)^q$  and  $C_2 = 2(q + 2)^{-2}e^{-q}$ .

**Lemma 40.** *Let  $X_1, \dots, X_n$  be independent random variables with  $\mathbb{E}X_i = 0$  and  $\|X_i\|_q < \infty$  for some  $q \geq 2$ . Let  $S_n = \sum_{i=1}^n X_i$ . Then there exists a constant  $K_q$  depending on  $q$  only such that*

$$\|S_n\|_q \leq K_q \left[ \left( \sum_{i=1}^n \mathbb{E}|X_i|^q \right)^{1/q} + \left( \sum_{i=1}^n \mathbb{E}X_i^2 \right)^{1/2} \right].$$

*Proof of Lemma 40.* For  $q = 2$ , it trivially holds by noticing  $\|S_n\|_2 = (\sum_{i=1}^n \mathbb{E}X_i^2)^{1/2}$ . For  $q > 2$ , it is an adapted version of Theorem 3 in [82].  $\square$

Next we will prove Theorem 37 and Theorem 38.

*Proof of Theorem 37.* By Theorem 1 in [28], we have the decoupling result: for any  $y > 0$ ,

$$\mathbb{P}(|\varepsilon^\top A \varepsilon| \geq y) \leq C \mathbb{P}(|\varepsilon^\top A \varepsilon^*| \geq y/C), \quad (2.3.1)$$

where  $\varepsilon^* = (\varepsilon_1^*, \dots, \varepsilon_n^*)^\top$  is an independent copy of  $\varepsilon$  and  $C$  is a positive constant independent of  $n$ . Now we condition on  $(\varepsilon_i)_{i=1}^n$ . Then  $\varepsilon^\top A \varepsilon^*$  is a linear combination of independent random variables  $\varepsilon_i^*$  with fixed coefficients  $w_i = \sum_j a_{ij} \varepsilon_j$ ,  $1 \leq i \leq n$ , i.e.,  $\varepsilon^\top A \varepsilon^* = \sum_i w_i \varepsilon_i^*$ . Let  $V = \sum_i w_i^2$ . By Lemma 39, we have

$$\mathbb{P}(|\varepsilon^\top A \varepsilon^*| \geq y | (\varepsilon_i)_{i=1}^n) \leq \frac{c_1 \mu_q^q \sum_i |w_i|^q}{y^q} + 2 \exp\left(-\frac{y^2}{c_2 \mu_2^2 V}\right),$$

where  $c_1$  and  $c_2$  are two positive constants depending on  $q$  only. By taking the expectation with respect to  $(\varepsilon_i)_{i=1}^n$ , it follows that

$$\mathbb{P}(|\varepsilon^\top A \varepsilon^*| \geq y) \leq \frac{c_1 \mu_q^q \sum_i \mathbb{E}|w_i|^q}{y^q} + 2 \mathbb{E} \exp\left(-\frac{y^2}{c_2 \mu_2^2 V}\right) =: \text{I} + \text{II}. \quad (2.3.2)$$

To bound the term I, by Lemma 40, we can obtain

$$\mathbb{E}|w_i|^q = \mathbb{E}\left|\sum_j a_{ij} \varepsilon_j\right|^q \lesssim \mu_2^q \left(\sum_j a_{ij}^2\right)^{q/2} + \mu_q^q \sum_j |a_{ij}|^q,$$

where the constant in  $\lesssim$  depends on  $q$  only. It then implies

$$\text{I} \lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q + \mu_q^{2q} \|A\|_{q,q}^q}{y^q}. \quad (2.3.3)$$

Now we deal with the term II. Let  $B = (b_{ij})_{i,j=1}^n = A^2$  and  $\mathcal{D}(A) = A^2 - \text{diag}(A^2)$ . Define

$$V_1 = \sum_i b_{ii}(\varepsilon_i^2 - \mathbb{E}\varepsilon_i^2) \quad \text{and} \quad V_2 = \varepsilon^\top \mathcal{D}(A) \varepsilon.$$

Then we can write  $V = \sum_i b_{ii} \mathbb{E} \varepsilon_i^2 + V_1 + V_2$ . Noticing that  $\sum_i b_{ii} = \|A\|_{2,2}^2$ , we have

$$\Pi \lesssim \exp\left(-\frac{y^2}{C_q \mu_2^4 \|A\|_{2,2}^2}\right) + \mathbb{E} \exp\left(-\frac{y^2}{C_q \mu_2^2 |V_1|}\right) + \mathbb{E} \exp\left(-\frac{y^2}{C_q \mu_2^2 |V_2|}\right) \quad (2.3.4)$$

with  $C_q = 3c_2$ . It remains to deal with  $\mathbb{E} \exp(-y^2/(C_q \mu_2^2 |V_i|))$  for  $i = 1, 2$ .

Case 1:  $q > 4$ . By Lemma 39, for  $x > 0$ ,

$$\mathbb{P}(|V_1| \geq x) \leq \frac{c_1 \mu_q^q \|A\|_{2,q}^q}{x^{q/2}} + 2 \exp\left(-\frac{x^2}{c_2 \mu_4^4 \sum_i b_{ii}^2}\right) =: F(x), \quad (2.3.5)$$

Elementary manipulation indicates

$$\begin{aligned} \mathbb{E} \exp\left(-\frac{y^2}{C_q \mu_2^2 |V_1|}\right) &= -\int_0^\infty \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) d\mathbb{P}(|V_1| \geq x) \\ &\leq -\int_0^\infty \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) F'(x) dx. \end{aligned} \quad (2.3.6)$$

where

$$-F'(x) = \frac{c_3 \mu_q^q \|A\|_{2,q}^q}{x^{q/2+1}} + \frac{x}{c_4 \mu_4^4 \sum_i b_{ii}^2} \exp\left(-\frac{x^2}{c_2 \mu_4^4 \sum_i b_{ii}^2}\right) =: f_1(x) + f_2(x) \quad (2.3.7)$$

with  $c_3 = c_1 \cdot q/2$  and  $c_4 = c_2/4$ . By a change of variables  $t = y^2/(C_q \mu_2^2 x)$  and the definition of Gamma function, we have

$$\begin{aligned} \int_0^\infty \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) f_1(x) dx &= c_3 \mu_q^q \|A\|_{2,q}^q \left(\frac{C_q \mu_2^2}{y^2}\right)^{q/2} \int_0^\infty e^{-t} t^{q/2-1} dt \\ &= \frac{c_3 C_q^{q/2} \mu_2^q \mu_q^q \|A\|_{2,q}^q}{y^q} \cdot \Gamma(q/2) = \frac{C_q' \mu_2^q \mu_q^q \|A\|_{2,q}^q}{y^q} \end{aligned} \quad (2.3.8)$$

Regarding  $f_2(x)$ , we can consider

$$\int_0^\infty \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) f_2(x) dx$$

$$\leq \frac{1}{c_4 \mu_4^4 \sum_i b_{ii}^2} \sup_{x>0} \left\{ x^{q+2} \exp \left( -\frac{x^2}{c_2 \mu_4^4 \sum_i b_{ii}^2} \right) \right\} \int_0^\infty x^{-q-1} \exp \left( -\frac{y^2}{C_q \mu_2^2 x} \right) dx.$$

Note that the function  $x^{q+2} \exp(-x^2/c_2 \mu_4^4 \sum_i b_{ii}^2)$  takes the maximum at

$$x = \sqrt{c_2(q+2)\mu_4^4 \sum_i b_{ii}^2/2}.$$

Similarly as (2.3.8), it follows that

$$\int_0^\infty x^{-q-1} \exp \left( -\frac{y^2}{C_q \mu_2^2 x} \right) dx = \frac{C_q^q \mu_2^{2q}}{y^{2q}} \cdot \Gamma(q).$$

Hence, we have

$$\int_0^\infty \exp \left( -\frac{y^2}{C_q \mu_2^2 x} \right) f_2(x) dx \lesssim \frac{\mu_2^{2q} (\mu_4^4 \sum_i b_{ii}^2)^{q/2}}{y^{2q}} \lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q}{y^q}, \quad (2.3.9)$$

where the last step follows in view of  $y \gtrsim \mu_2 \mu_q \|A\|_{2,2}$ ,  $\mu_q \geq \mu_4$  and

$$\sum_i b_{ii}^2 = \sum_i \left( \sum_j a_{ij}^2 \right)^2 \leq \|A\|_{2,q}^2 \left[ \sum_i \left( \sum_j a_{ij}^2 \right)^{\frac{q}{q-2}} \right]^{\frac{q-2}{q}} \leq \|A\|_{2,q}^2 \|A\|_{2,2}^2.$$

Combining (2.3.6), (2.3.7), (2.3.8) and (2.3.9), we have for  $y \gtrsim \mu_2 \mu_q \|A\|_{2,2}$ ,

$$\mathbb{E} \exp \left( -\frac{y^2}{C_q \mu_2^2 |V_1|} \right) \lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q}{y^q}. \quad (2.3.10)$$

Moreover, we can obtain

$$\begin{aligned} \mathbb{E} \exp \left( -\frac{y^2}{C_q \mu_2^2 |V_2|} \right) &= \mathbb{E} \exp \left( -\frac{y^2}{C_q \mu_2^2 |V_2|} \right) (\mathbf{1}\{|V_2| \leq y \|A\|_{2,2}\} + \mathbf{1}\{|V_2| > y \|A\|_{2,2}\}) \\ &\leq \exp \left( -\frac{y}{C_q \mu_2^2 \|A\|_{2,2}} \right) - \int_{y \|A\|_{2,2}}^\infty \exp \left( -\frac{y^2}{C_q \mu_2^2 x} \right) d\mathbb{P}(|V_2| \geq x), \end{aligned}$$

which together with (2.3.1), (2.3.2), (2.3.3), (2.3.4) and (2.3.10) implies

$$\begin{aligned} \mathbb{P}(|\varepsilon^\top A \varepsilon| \geq y) &\lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q + \mu_q^{2q} \|A\|_{q,q}^q}{y^q} + \exp\left(-\frac{y}{C_q \mu_2^2 \|A\|_{2,2}}\right) \\ &\quad - \int_{y\|A\|_{2,2}}^\infty \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) d\mathbb{P}(|V_2| \geq x). \end{aligned} \quad (2.3.11)$$

By the preceding arguments and noticing  $\|\mathcal{D}(A)\|_{2,2} \leq \|A\|_{2,2}^2$ , we can also obtain for  $x \gtrsim \mu_2 \mu_q \|A\|_{2,2}^2$ ,

$$\begin{aligned} \mathbb{P}(|V_2| \geq x) &\lesssim \frac{\mu_2^q \mu_q^q \|\mathcal{D}(A)\|_{2,q}^q + \mu_q^{2q} \|\mathcal{D}(A)\|_{q,q}^q}{x^q} + \exp\left(-\frac{x}{C_q \mu_2^2 \|A\|_{2,2}^2}\right) \\ &\quad - \int_{x\|A\|_{2,2}^2}^\infty \exp\left(-\frac{x^2}{C_q \mu_2^2 z}\right) d\mathbb{P}(|V_3| \geq z), \end{aligned} \quad (2.3.12)$$

where  $V_3 = \varepsilon^\top \mathcal{D}^2(A) \varepsilon$  with  $\mathcal{D}^2(A) = \mathcal{D} \circ \mathcal{D}(A)$ . We next bound  $\|\mathcal{D}(A)\|_{2,q}^q$ ,  $\|\mathcal{D}(A)\|_{q,q}^q$  and  $\mathbb{P}(|V_3| \geq z)$  respectively. Firstly, by Holder's inequality, we have

$$\begin{aligned} \|\mathcal{D}(A)\|_{2,q}^q &= \sum_i \left[ \sum_{j:j \neq i} \left( \sum_k a_{ik} a_{kj} \right)^2 \right]^{q/2} \leq \|A\|_{2,q}^q \|A\|_{2,2}^q, \\ \|\mathcal{D}(A)\|_{q,q}^q &= \sum_{i \neq j} \left| \sum_k a_{ik} a_{kj} \right|^q \leq \sum_{i \neq j} \left( \sum_k a_{ik}^2 \right)^{q/2} \left( \sum_k a_{kj}^2 \right)^{q/2} \leq \|A\|_{2,q}^{2q}. \end{aligned}$$

As for  $\mathbb{P}(|V_3| \geq z)$ , by Theorem 1 in [28] again and Markov's inequality, we have

$$\mathbb{P}(|V_3| \geq z) \leq C \mathbb{P}(|\varepsilon^\top \mathcal{D}^2(A) \varepsilon^*| \geq z/C) \leq C z^{-q} \mathbb{E} |\varepsilon^\top \mathcal{D}^2(A) \varepsilon^*|^q. \quad (2.3.13)$$

Denote  $\mathcal{D}^2(A) = (d_{ij})_{i,j=1}^n$ . Then we can write  $\varepsilon^\top \mathcal{D}^2(A) \varepsilon^* = \sum_i \varepsilon_i \sum_j d_{ij} \varepsilon_j^*$ . Observe that  $\varepsilon_i \sum_j d_{ij} \varepsilon_j^*$  are uncorrelated for different  $i$ . By the Burkholder inequality for martingales (cf. Theorem 3.2 in [16]) and the fact that  $\|\mathcal{D}(M)\|_{2,2} \leq \|M\|_{2,2}^2$  for any square matrix  $M$ ,

we have

$$\begin{aligned}
\|\varepsilon^\top \mathcal{D}^2(A) \varepsilon^*\|_q^2 &\leq C_q \sum_i \|\varepsilon_i \sum_j d_{ij} \varepsilon_j^*\|_q^2 \leq C_q \mu_q^2 \sum_i \|\sum_j d_{ij} \varepsilon_j^*\|_q^2 \\
&\leq C_q \mu_q^4 \sum_{i,j} d_{ij}^2 = C_q \mu_q^4 \|\mathcal{D}^2(A)\|_{2,2}^2 \leq C_q \mu_q^4 \|A\|_{2,2}^8. \quad (2.3.14)
\end{aligned}$$

By (2.3.13) and (2.3.14), (2.3.12) then becomes

$$\begin{aligned}
P(|V_2| \geq x) &\lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q \|A\|_{2,2}^q + \mu_q^{2q} \|A\|_{2,q}^{2q}}{x^q} + \exp\left(-\frac{x}{C_q \mu_2^2 \|A\|_{2,2}^2}\right) \\
&\quad - \int_{x\|A\|_{2,2}^2}^\infty \exp\left(-\frac{x^2}{C_q \mu_2^2 z}\right) \frac{\mu_q^{2q} \|A\|_{2,2}^{4q}}{z^{q+1}} dz. \quad (2.3.15)
\end{aligned}$$

By similar arguments as (2.3.6) and (2.3.7), let

$$R_i = \int_{y\|A\|_{2,2}}^\infty \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) g_i(x) dx \quad \text{for } i = 1, 2, 3,$$

where

$$\begin{aligned}
g_1(x) &= x^{-q-1} (\mu_2^q \mu_q^q \|A\|_{2,q}^q \|A\|_{2,2}^q + \mu_q^{2q} \|A\|_{2,q}^{2q}), \\
g_2(x) &= (\mu_2^2 \|A\|_{2,2}^2)^{-1} \exp\left(-\frac{x}{C_q \mu_2^2 \|A\|_{2,2}^2}\right), \\
g_3(x) &= x^{-q-1} \mu_q^{2q} \|A\|_{2,2}^{2q} \exp\left(-\frac{x}{C_q \mu_2^2 \|A\|_{2,2}^2}\right). \quad (2.3.16)
\end{aligned}$$

Then we can write

$$- \int_{y\|A\|_{2,2}}^\infty \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) d\mathbb{P}(|V_2| \geq x) \lesssim R_1 + R_2 + R_3.$$

Similarly as (2.3.8) and (2.3.9), it can be obtained that for  $y \geq \mu_2 \mu_q \|A\|_{2,2}$ ,

$$R_1 \lesssim \frac{\mu_2^{3q} \mu_q^q \|A\|_{2,q}^q \|A\|_{2,2}^q + \mu_2^{2q} \mu_q^{2q} \|A\|_{2,q}^{2q}}{y^{2q}} \lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q}{y^q},$$

$$\begin{aligned}
R_2 &\lesssim \exp\left(-\frac{y}{C_q \mu_2^2 \|A\|_{2,2}}\right) \left(\frac{\mu_2^2 \|A\|_{2,2}}{y}\right)^{q-1} \leq \exp\left(-\frac{y}{C_q \mu_2^2 \|A\|_{2,2}}\right), \\
R_3 &\lesssim \exp\left(-\frac{y}{C_q \mu_2^2 \|A\|_{2,2}}\right) \left(\frac{\mu_2 \mu_q \|A\|_{2,2}}{y}\right)^{2q} \leq \exp\left(-\frac{y}{C_q \mu_2^2 \|A\|_{2,2}}\right).
\end{aligned}$$

As a consequence, it implies

$$-\int_{y\|A\|_{2,2}}^{\infty} \exp\left(-\frac{y^2}{C_q \mu_2^2 x}\right) d\mathbb{P}(|V_2| \geq x) \lesssim \frac{\mu_2^q \mu_q^q \|A\|_{2,q}^q}{y^q} + \exp\left(-\frac{y}{C_q \mu_2^2 \|A\|_{2,2}}\right). \quad (2.3.17)$$

Therefore, (2.2.1) follows in view of (2.3.11) and (2.3.17).

Case 2:  $2 \leq q \leq 4$ . It is much easier for this case. By Corollary 1.6 in [64] and Markov's inequality, with  $\sum_i |b_{ii}|^{q/2} = \|A\|_{2,q}^q$ , we have for  $x > 0$ ,

$$\mathbb{P}(|V_1| \geq x) \lesssim \sum_i \mathbb{P}(|b_{ii}(\varepsilon_i^2 - \mathbb{E}\varepsilon_i^2)| \geq x/2) + \frac{\mu_q^q \sum_i |b_{ii}|^{q/2}}{x^{q/2}} \lesssim \frac{\mu_q^q \|A\|_{2,q}^q}{x^{q/2}}, \quad (2.3.18)$$

Replacing (2.3.5) by (2.3.18), due to (2.3.8), we can still have

$$\mathbb{E} \exp\left(-\frac{y^2}{C_q \mu_2^2 |V_1|}\right) \leq \frac{C_q \mu_2^q \mu_q^q \|A\|_{2,q}^q}{y^q}.$$

Other steps are the same as the case  $q > 4$ . □

*Proof of Theorem 38.* Case 1:  $2 \leq q < 4$ . By Corollary 3 in [26], we have

$$\mathbb{E} |\varepsilon^\top A \varepsilon|^q \leq C_q \mathbb{E} \left| \sum_{i \neq j} a_{ij}^2 \varepsilon_i^2 \varepsilon_j^2 \right|^{q/2}. \quad (2.3.19)$$

By Theorem 3.6 in [27], it follows that

$$\mathbb{E} \left| \sum_{i \neq j} a_{ij}^2 \varepsilon_i^2 \varepsilon_j^2 \right|^{q/2} \leq C_q \left( \mu_2^{2q} \|A\|_{2,2}^q + \mu_2^q \mu_q^q \|A\|_{2,q}^q + \mu_q^{2q} \|A\|_{q,q}^q \right). \quad (2.3.20)$$

Hence we have

$$\mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q \left( \mu_2^{2q} \|A\|_{2,2}^q + \mu_2^q \mu_q^q \|A\|_{2,q}^q + \mu_q^{2q} \|A\|_{q,q}^q \right), \quad (2.3.21)$$

which agrees with (2.2.8) for  $K = 0$ .

Case 2:  $q \geq 4$ . By Theorem 1 in [26], we have the decoupling result:

$$\mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q \mathbb{E}|\varepsilon^\top A \varepsilon^*|^q, \quad (2.3.22)$$

where  $\varepsilon^* = (\varepsilon_1^*, \dots, \varepsilon_n^*)^\top$  is an independent copy of  $\varepsilon$ . Now we condition on  $(\varepsilon_i)_{i=1}^n$ . Then  $\varepsilon^\top A \varepsilon^*$  is a linear combination of independent random variables  $\varepsilon_i^*$  with fixed coefficients  $w_i = \sum_j a_{ij} \varepsilon_j$ ,  $1 \leq i \leq n$ , i.e.,  $\varepsilon^\top A \varepsilon^* = \sum_i w_i \varepsilon_i^*$ . By Lemma 40, we have

$$\mathbb{E}(|\varepsilon^\top A \varepsilon^*|^q | (\varepsilon_i)_{i=1}^n) \leq C_q [\mu_q^q \sum_i |w_i|^q + \mu_2^q (\sum_i w_i^2)^{q/2}].$$

By taking the expectation with respect to  $(\varepsilon_i)_{i=1}^n$ , it follows that

$$\mathbb{E}(|\varepsilon^\top A \varepsilon^*|^q) \leq C_q [\mu_q^q \sum_i \mathbb{E}|w_i|^q + \mu_2^q \mathbb{E}(\sum_i w_i^2)^{q/2}]. \quad (2.3.23)$$

Using Lemma 40 again, we have

$$\sum_i \mathbb{E}|w_i|^q \leq C_q (\mu_q^q \|A\|_{q,q}^q + \mu_2^q \|A\|_{2,q}^q). \quad (2.3.24)$$

Hence, it suffices to deal with the term  $\mathbb{E}(\sum_i w_i^2)^{q/2}$ . Let  $B = (b_{ij})_{i,j=1}^n = A^2$  and  $\mathcal{D}(A) = A^2 - \text{diag}(A^2)$ . Then we can write  $\sum_{i=1}^n w_i^2 = \sum_i b_{ii} \varepsilon_i^2 + \varepsilon^\top \mathcal{D}(A) \varepsilon$ . By Lemma 1 in [82],

$$\mathbb{E}|\sum_i b_{ii} \varepsilon_i^2|^{q/2} \leq C_q \mu_q^q \sum_i b_{ii}^{q/2} + C_q \mu_2^q (\sum_i b_{ii})^{q/2}$$

$$= C_q \mu_q^q \|A\|_{2,q}^q + C_q \mu_2^q \|A\|_{2,2}^q. \quad (2.3.25)$$

Combining (2.3.22), (2.3.23), (2.3.24) and (2.3.25), we have

$$\mathbb{E}|\varepsilon^\top A \varepsilon|^q \leq C_q \left( \mu_2^q \mathbb{E}|\varepsilon^\top \mathcal{D}(A) \varepsilon|^{q/2} + \mu_2^{2q} \|A\|_{2,2}^q + \mu_2^q \mu_q^q \|A\|_{2,q}^q + \mu_q^{2q} \|A\|_{q,q}^q \right), \quad (2.3.26)$$

which provides a recursive relationship for the moment. If  $2 \leq q/2 < 4$ , by (2.3.21), we have

$$\mathbb{E}|\varepsilon^\top \mathcal{D}(A) \varepsilon|^{q/2} \leq C_q \left( \mu_2^q \|\mathcal{D}(A)\|_{2,2}^{q/2} + \mu_2^{q/2} \mu_{q/2}^{q/2} \|\mathcal{D}(A)\|_{2,q/2}^{q/2} + \mu_{q/2}^q \|\mathcal{D}(A)\|_{q/2,q/2}^{q/2} \right).$$

If  $q \geq 8$ , we can apply the recursion (2.3.26) for  $K = \lfloor \log_2(q) \rfloor - 1$  times such that  $2 \leq q/2^K < 4$ . Then (2.2.8) follows by adding up all the intermediate terms during the recursion and applying (2.3.21) for  $\mathbb{E}|\varepsilon^\top \mathcal{D}^K(A) \varepsilon|^{q/2^K}$ .  $\square$

## 2.4 Acknowledgements

Chapter 2, in full, has been submitted for publication of the material “Probability and Moment Inequalities for Quadratic Forms in Independent Random Variables with Fat Tails” by Chi Zhang and Danna Zhang, as it appears in *Statistics & Probability Letters*, 2025. The dissertation author was the primary investigator and author of this paper.

# Bibliography

- [1] Radosław Adamczak. A note on the hanson-wright inequality for random vectors with dependencies. *Electronic Communications in Probability*, 20:1–13, 2015.
- [2] Theodore Wilbur Anderson. *The Statistical Analysis of Time Series*. Wiley, 1971.
- [3] ASOS ASOS. Automated surface observing system: Asos user’s guide.[washington, dc]: Us dept. of commerce. *National Oceanic and Atmospheric Administration: Federal Aviation Administration: US Navy: US Dept. of the Air Force*, URL <https://www.weather.gov/media/asos/aum-toc.pdf>, 1998.
- [4] Zhidong Bai and Jack William Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.
- [5] Franck Barthe and Emanuel Milman. Transference principles for log-sobolev and spectral-gap with applications to conservative spin systems. *Communications in Mathematical Physics*, 323(2):575–625, 2013.
- [6] Maurice Stevenson Bartlett. An introduction to stochastic processes. *Cambridge University Press, London*, 1955.
- [7] Andre M. Bastos and Jeroen M. Schoffelen. A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in systems neuroscience*, 9:175, 2016.
- [8] Mats Birger Båth. *Spectral Analysis in Geophysics*. ISSN. Elsevier Science, 2012.
- [9] Julius S. Bendat. *Nonlinear System Analysis and Identification from Random Data*. Wiley-Interscience, 1990.
- [10] Rimantas Y. Bentkus and Rimantas Rudzakis. On the distributions of some statistical estimates of spectral density. *Teoriya Veroyatnostei i ee Primeneniya*, 27(4):739–756, 1982.
- [11] Monika Bhattacharjee and Arup Bose. Estimation of autocovariance matrices for

- infinite dimensional vector linear process. *Journal of Time Series Analysis*, 35(3):262–281, 2014.
- [12] Monika Bhattacharjee and Arup Bose. Large sample behaviour of high dimensional autocovariance matrices. *The Annals of Statistics*, 44(2):598–628, 2016.
  - [13] George E. P. Box and David A. Pierce. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J. Amer. Statist. Assoc.*, 65:1509–1526, 1970.
  - [14] David Ross Brillinger. *Time series: data analysis and theory*. SIAM, 2001.
  - [15] Peter J. Brockwell and Richard A. Davis. *Time series theory and methods*. Springer, 1991.
  - [16] Donald L. Burkholder. Distribution function inequalities for martingales. *The Annals of Probability*, pages 19–42, 1973.
  - [17] John Louis Casti. *Nonlinear System Theory*. Academic Press, 1985.
  - [18] Jinyuan Chang, Qiwei Yao, and Wen Zhou. Testing for high-dimensional white noise using maximum cross-correlations. *Biometrika*, 104(1):111–127, 2017.
  - [19] Xiaohui Chen. A note on moment inequality for quadratic forms. *Statistics & Probability Letters*, 92:83–88, 2014.
  - [20] Xiaohui Chen, Mengyu Xu, and Wei Biao Wu. Covariance and precision matrix estimation for high-dimensional time series. *The Annals of Statistics*, 41(6):2994–3021, 12 2013.
  - [21] Xiaohui Chen, Mengyu Xu, and Wei Biao Wu. Regularized estimation of linear functionals of precision matrices for high-dimensional time series. *IEEE Transactions on Signal Processing*, 64(24):6459–6470, 2016.
  - [22] Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *The Annals of Statistics*, 41(6):2786–2819, 2013.
  - [23] Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Comparison and anti-concentration bounds for maxima of gaussian random vectors. *Probability Theory and Related Fields*, 162(1-2):47–70, 2015.
  - [24] Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Central limit theorems and bootstrap in high dimensions. *The Annals of Probability*, 45(4):2309–2352, 2017.

- [25] Ratnam V. Chitturi. Distribution of residual autocorrelations in multiple autoregressive schemes. *J. Amer. Statist. Assoc.*, 69:928–934, 1974.
- [26] Víctor H. de la Peña. Decoupling and khintchine’s inequalities for U-statistics. *The Annals of Probability*, pages 1877–1892, 1992.
- [27] Víctor H. de la Peña, Rustam Ibragimov, and Serik Sharakhmetov. On extremal distributions and sharp  $L_p$ -bounds for sums of multilinear forms. *The Annals of Probability*, 31(2):630–675, 2003.
- [28] Víctor H. de la Peña and Stephen J. Montgomery-Smith. Decoupling inequalities for the tail probabilities of multivariate U-statistics. *The Annals of Probability*, pages 806–816, 1995.
- [29] Mingzhou Ding, Steven L Bressler, Weiming Yang, and Hualou Liang. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variability assessment. *Biological cybernetics*, 83:35–45, 2000.
- [30] Eugene Duff, Leigh Johnston, Jinhu Xiong, Peter Fox, Iven Mareels, and Gary Egan. The power of spectral density analysis for mapping endogenous bold signal fluctuations. *Human brain mapping*, 29:778–90, 07 2008.
- [31] James Durbin. Tests of serial independence based on the cumulated periodogram. *Bulletin of the International Statistical Institute*, 42:1039–1048, 1967.
- [32] Steven N. Durlauf. Spectral based testing of the martingale hypothesis. *Journal of Econometrics*, 50(3):355–376, 1991.
- [33] Jianqing Fan, Jinchi Lv, and Lei Qi. Sparse high dimensional models in economics. *Annual review of economics*, 3:291–317, 2011.
- [34] Mark Fiecas, Chenlei Leng, Weidong Liu, and Yuan Yu. Spectral analysis of high-dimensional time series. *Electronic Journal of Statistics*, 13(2):4079–4101, 2019.
- [35] Mario Forni, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. The generalized dynamic-factor model: Identification and estimation. *Review of Economics and statistics*, 82(4):540–554, 2000.
- [36] Mario Forni, Marc Hallin, Marco Lippi, and Paolo Zaffaroni. Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis. *Journal of Econometrics*, 199(1):74–92, 2017.
- [37] Luis García Domínguez, Jason Stieben, José Luis Pérez Velázquez, and Stuart

- Shanker. The imaginary part of coherency in autism: differences in cortical functional connectivity in preschool children. *PLoS One*, 8(10):e75941, 2013.
- [38] Michael Ghil, MR Allen, MD Dettinger, K Ide, D Kondrashov, ME Mann, Andrew W Robertson, A Saunders, Y Tian, Fa Varadi, and P. Yiou. Advanced spectral methods for climatic time series. *Reviews of geophysics*, 40(1):3–1, 2002.
  - [39] Finn-Idar Grøtta Giske, Bernt Johan Leira, and Ole Øiseth. Efficient computation of cross-spectral densities in the stochastic modelling of waves and wave loads. *Applied Ocean Research*, 62:70–88, 2017.
  - [40] Ulf Grenander and Murray Rosenblatt. *Statistical analysis of stationary time series*. Chelsea, New York, NY, 1957.
  - [41] Joachim Gross, Jan Kujala, Matti Hämäläinen, Lars Timmermann, Alfons Schnitzler, and Riitta Salmelin. Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences*, 98(2):694–699, 2001.
  - [42] Peter Hall and Biao Jing. On sample reuse methods for dependent data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(4):727–737, 1996.
  - [43] Edward James Hannan. Regression for time series. *Time series analysis*, pages 17–37, 1963.
  - [44] Edward James Hannan. *Multiple time series*. John Wiley & Sons, 1970.
  - [45] Donald L. Hanson and Frank T. Wright. A bound on tail probabilities for quadratic forms in independent random variables. *The Annals of Mathematical Statistics*, 42(3):1079–1083, 1971.
  - [46] Larry D. Haugh. Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach. *Journal of the American Statistical Association*, 71(354):378–385, 1976.
  - [47] Jun He. Coherence and cross-spectral density matrix analysis of random wind and wave in deep water. *Ocean Engineering*, 197:106930, 2020.
  - [48] Felix U. Hohlefeld, Christoph Huchzermeyer, Jochen Huebl, G.H. Schneider, Guido Nolte, Christian Brücke, Thomas Schönecker, Andrea A. Kühn, Guido Curio, and Vadim V. Nikulin. Functional and effective connectivity in subthalamic local field potential recordings of patients with parkinson’s disease. *Neuroscience*, 250:320–332, 2013.

- [49] Yongmiao Hong. Testing for independence between two covariance stationary time series. *Biometrika*, 83(3):615–625, 1996.
- [50] Jonathan R. M. Hosking. The multivariate portmanteau statistic. *J. Amer. Statist. Assoc.*, 75(371):602–608, 1980.
- [51] Daniel Hsu, Sham M. Kakade, and Tong Zhang. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 17(52):1–6, 2012.
- [52] Baisuo Jin, Cheng Wang, Zhidong Bai, Krishnan K. Nair, and Matthew Harding. Limiting spectral distribution of a symmetrized auto-cross covariance matrix. *The Annals of Applied Probability*, 24(3):1199–1225, 2014.
- [53] Patrick D. Koch and Shangjin Yang. A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function. *Journal of the American Statistical Association*, 81(394):533–544, 1986.
- [54] Hans R Künsch. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, pages 1217–1241, 1989.
- [55] Zhentao Li, Clifford Lam, Jianfeng Yao, and Qiwei Yao. On testing for high-dimensional white noise. *The Annals of Statistics*, 47(6):3382–3412, 2019.
- [56] Martin A. Lindquist, Yishuai Xu, Michael B. Nebel, and Brian S. Caffo. Evaluating dynamic bivariate correlations in resting-state fmri: A comparison study and a new approach. *Neuroimage*, 101:531–546, 2014.
- [57] Han Liu, Alexander Aue, and Debashis Paul. On the marčenko–pastur law for linear time series. *The Annals of Statistics*, 43(2):675–712, 2015.
- [58] Regina Y. Liu and Kesar Singh. Moving block jackknife and bootstrap capture weak dependence. *Exploring the limits of bootstrap*, R. Lepage & L. Billard eds, 225:248, 1992.
- [59] Weidong Liu and Wei Biao Wu. Asymptotics of spectral density estimates. *Econometric Theory*, 26(4):1218–1245, 2010.
- [60] Greta M. Ljung and George E. P. Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.
- [61] Helmut Lütkepohl. *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.

- [62] Laura Marzetti, Guido Nolte, Maria Gabriella Perrucci, Giorgio L. Romani, and Claudio Del Gratta. The use of standardized infinity reference in eeg coherency studies. *Neuroimage*, 36(1):48–63, 2007.
- [63] Tucker S. McElroy and Dimitris N. Politis. Optimal linear interpolation of multiple missing values. *Statistical Inference for Stochastic Processes*, 25:1–13, 2022.
- [64] Sergei Vladimirovich Nagaev. Large deviations of sums of independent random variables. *The Annals of Probability*, pages 745–789, 1979.
- [65] Fedor Nazarov. On the maximal perimeter of a convex set in  $r^n$  with respect to gaussian measure, geometric aspects of func. *Anal*, 1807:169–187, 2003.
- [66] Henry R. Neave. A comparison of lag window generators. *Journal of the American Statistical Association*, 67(337):152–158, 1972.
- [67] Guido Nolte, Ovid Bai, Lewis Wheaton, Zoltan Mari, Stefan Vorbach, and Mark Hallett. Identifying true brain interaction from eeg data using the imaginary part of coherency. *Clinical neurophysiology*, 115(10):2292–2307, 2004.
- [68] Dimitris N. Politis and Joseph P. Romano. Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics*, pages 2031–2050, 1994.
- [69] Dimitris N. Politis, Joseph P. Romano, and Michael Wolf. *Subsampling*. Springer, 1999.
- [70] Stephen Portnoy. On the central limit theorem in  $R^p$  when  $p \rightarrow \infty$ . *Probability Theory and Related Fields*, 73(4):571–583, 1986.
- [71] Maurice Bertram Priestley. *Spectral Analysis and Time Series*. Academic Press, 1981.
- [72] Maurice Bertram Priestley. *Non-linear and Non-stationary Time Series Analysis*. Academic Press, 1988.
- [73] Gregory C Reinsel. *Elements of multivariate time series analysis*. Springer Science & Business Media, 2003.
- [74] Emmanuel Rio. Moment inequalities for sums of dependent random variables under projective conditions. *Journal of Theoretical Probability*, 22(1):146–163, 2009.
- [75] Peter A. Robinson, James A. Henderson, Nicolas C. Gabay, Kirsten M. Aquino, Tara Babaie-Janvier, and Xiaogeng Gao. Determination of dynamic brain connectivity

via spectral analysis. *Frontiers in Human Neuroscience*, page 360, 2021.

- [76] Peter Michael Robinson. Non-linear regression for multiple time-series. *Journal of Applied Probability*, 9(4):758–768, 1972.
- [77] Peter Michael Robinson. Automatic frequency domain inference on semiparametric and nonparametric models. *Econometrica: Journal of the Econometric Society*, pages 1329–1363, 1991.
- [78] Murray Rosenblatt. The multidimensional prediction problem. *Proceedings of the National Academy of Sciences*, 43(11):989–992, 1957.
- [79] Murray Rosenblatt. *Markov Processes: Structure and Asymptotic Behavior*. Springer, 1971.
- [80] Murray Rosenblatt. Asymptotic normality, strong mixing and spectral density estimates. *The Annals of Probability*, 12(4):1167–1180, 1984.
- [81] Murray Rosenblatt. *Stationary Sequences and Random Fields*. Springer, 1985.
- [82] Haskell Paul Rosenthal. On the subspaces of  $L^p$  ( $p > 2$ ) spanned by sequences of independent random variables. *Israel Journal of Mathematics*, 8(3):273–303, 1970.
- [83] Antoine Rosuel, Philippe Loubaton, and Pascal Vallet. On the asymptotic distribution of the maximum sample spectral coherence of gaussian time series in the high dimensional regime. *arXiv preprint arXiv:2107.02891*, 2021.
- [84] Mark Rudelson and Roman Vershynin. Hanson-wright inequality and sub-gaussian concentration. *Electronic Communications in Probability*, 18(82):1–9, 2013.
- [85] Rimas Rudzkis. On the distribution of the maximum deviation of the gaussian stationary time series spectral density estimate. *Lithuanian Mathematical Journal*, pages 118–130, 1985.
- [86] Marshall Schetzen. *The Volterra and Wiener Theories of Nonlinear Systems*. Krieger Publishing Company, 1989.
- [87] Xiaofeng Shao and Wei Biao Wu. Asymptotic spectral theory for nonlinear time series. *The Annals of Statistics*, 35(4):1773–1801, 2007.
- [88] Michael Sherman and Edward Carlstein. Replicate histograms. *Journal of the American Statistical Association*, 91(434):566–576, 1996.
- [89] Sean L. Simpson, F. DuBois Bowman, and Paul J. Laurienti. Analyzing complex

functional brain networks: fusing statistics and network science to understand the brain. *Statistics surveys*, 7:1, 2013.

- [90] Christopher A Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, 1980.
- [91] Vladimir Spokoiny and Maryna Zhilova. Sharp deviation bounds for quadratic forms. *Mathematical Methods of Statistics*, 22(2):100–113, 2013.
- [92] Anthony Stathopoulos and Matthew G Karlaftis. Spectral and cross-spectral analysis of urban traffic flows. In *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 01TH8585)*, pages 820–825. IEEE, 2001.
- [93] Martin Stendel, Jennifer Francis, Rachel White, Paul D Williams, and Tim Woollings. The jet stream and climate change. In *Climate change*, pages 327–357. Elsevier, 2021.
- [94] James H Stock and Mark W Watson. Vector autoregressions. *Journal of Economic perspectives*, 15(4):101–115, 2001.
- [95] Yubo Sun, Yiyuan Li, Amy Kuceyeski, and Sumanta Basu. Large spectral density matrix estimation by thresholding. *arXiv preprint arXiv:1812.00532*, 2018.
- [96] Dennis P Todey, DE Herzmman, and ES Takle. The iowa environmental mesonet combining observing systems into a single network. In *Sixth Symposium on Integrated Observing Systems*. Iowa State University Ames, IA, USA, 2002.
- [97] Howell Tong. *Non-linear Time Series: A Dynamical System Approach*. Oxford University Press, 1990.
- [98] Ruey S. Tsay. *Analysis of Financial Time Series*, volume 543. John Wiley & Sons, 2005.
- [99] Ruey S. Tsay. Testing serial correlations in high-dimensional time series via extreme value theory. *Journal of Econometrics*, 216(1):106–117, 2020.
- [100] Vincent Q. Vu and Kejun Wang. Random weighted projections, random quadratic forms and random eigenvectors. *Random Structures & Algorithms*, 47(4):792–821, 2015.
- [101] David A Warde and Sebastián M Torres. The autocorrelation spectral density for doppler-weather-radar signal analysis. *IEEE transactions on geoscience and remote sensing*, 52(1):508–518, 2013.
- [102] Peter Welch. The use of fast fourier transform for the estimation of power spec-

- tra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73, 1967.
- [103] Norbert Wiener. *Nonlinear Problems in Random Theory*. Wiley, New York, 1958.
  - [104] Michael B. Woodroffe and John W. Van Ness. The maximum deviation of sample spectral densities. *The Annals of Mathematical Statistics*, pages 1558–1569, 1967.
  - [105] Frank T. Wright. A bound on tail probabilities for quadratic forms in independent random variables whose distributions are not necessarily symmetric. *The Annals of Probability*, 1(6):1068–1070, 1973.
  - [106] Wei Biao Wu. Nonlinear system theory: another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America*, 102(40):pp. 14150–14154, 2005.
  - [107] Wei Biao Wu and Xiaofeng Shao. Limit theorems for iterated random functions. *Journal of Applied Probability*, pages 425–436, 2004.
  - [108] Wei Biao Wu and Yang Ning Wu. Performance bounds for parameter estimates of high-dimensional linear models with correlated errors. *Electronic Journal of Statistics*, 10(1):352–379, 2016.
  - [109] Wei Biao Wu and Paolo Zaffaroni. Asymptotic theory for spectral density estimates of general multivariate time series. *Econometric Theory*, 34(1):1–22, 2018.
  - [110] Han Xiao and Wei Biao Wu. Asymptotic inference of autocovariances of stationary processes. *arXiv preprint arXiv:1105.3423*, 2011.
  - [111] Han Xiao and Wei Biao Wu. Covariance matrix estimation for stationary time series. *The Annals of Statistics*, 40(1):466–493, 2012.
  - [112] Danna Zhang. Robust estimation of the mean and covariance matrix for high dimensional time series. *Statistica Sinica*, 31(2):797–820, 2021.
  - [113] Danna Zhang and Wei Biao Wu. Gaussian approximation for high dimensional time series. *The Annals of Statistics*, 45(5):1895–1919, 2017.
  - [114] Danna Zhang and Wei Biao Wu. Convergence of covariance and spectral density estimates for high-dimensional locally stationary processes. *The Annals of Statistics*, 49(1):233–254, 2021.
  - [115] Xianyang Zhang and Guang Cheng. Bootstrapping high dimensional time series. *arXiv preprint arXiv:1406.1037*, 2014.

- [116] Xianyang Zhang and Guang Cheng. Gaussian approximation for high dimensional vector under physical dependence. *Bernoulli*, 24:2640–2675, 2018.
- [117] Antoni Zygmund. *Trigonometric Series. Vol I, II*. Third Ed. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2002.