Asymptotics for linear spectral statistics of sample covariance matrices

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
an der Fakultät für Mathematik
der Ruhr-Universität Bochum

vorgelegt von

Nina Dörnemann

März 2022

In Erinnerung an meine geliebte "Oma Mausi".

Acknowledgements

It is my great pleasure to thank all the people who contributed in one way or another to this unbelievable journey which shaped me both as a researcher and a human being.

First and foremost, I would like to express my sincere gratitude to Prof. Holger Dette for the possibility to pursue my PhD studies under his supervision. I am extremely grateful for his continuous support, patience and many fruitful discussions throughout the last years. His expertise and passion led me through interesting research projects and helped me to overcome scientific difficulties.

My position was supported by German Research Foundation (DFG) through the research training group "High dimensional phenomena in probability – fluctuations and discontinuity" (RTG 2131). Representative for all the people involved, I would like to thank Prof. Peter Eichelsbacher as the speaker of our RTG for his exceptional commitment. Although the pandemic forced me to cancel further travel plans, I enjoyed participating in conferences, seminars, and summer schools thanks to the RTG.

I owe special gratitude to our years-long team assistant Martina Stein for her continuous help and many pleasant chats, as well as to Tanja Schiffmann and Birgit Tomöhlen for their assistance.

During the last year, I really enjoyed working with Dr. Johannes Heiny and would like to thank him for many exciting discussions and the pleasant and productive collaboration.

More generally, I would like to thank all the current and former colleagues from our department and the RTG. I appreciate your openness, helpfulness, mathematical curiosity and the joint breaks. Especially, I would like to thank Theresa Eckle for many interesting discussions (especially on our fun project "Guinea pig" together with Tim), joint tea breaks and even more for being such a reliable friend.

Back in my bachelor studies, I started exploring the fascinating fields of stochastics. I would like to thank Prof. Angelika Rohde and Prof. Christoph Thäle for their excellent lectures and seminars, which raised my interest and led me to specialize in this field.

Looking further backwards, my former math teacher Melanie Reidegeld played a pivotal role in my decision to study mathematics. I would like to thank her for her extraordinary support through many years and "Jugend forscht" competitions.

Words cannot express my gratitude to my family, especially my one-in-a-million parents Andrea and Ralf and my beloved sister Natascha, for their unconditional love and continuous support during my studies (without even knowing exactly what I am doing). I admire your ability to give me the freedom to grow while always standing by. Thank you for always believing in me.

Last but not least, I would like to express my deep gratitude to Tim, my dearest and loving partner-in-crime. Thank you for your undoubtedly invaluable contributions to my studies. Thank you for being with me all the time. Thank you for making our journey so precious.

Nina Dörnemann
Bochum, March 2022

Contents

Chapter 1

Introduction

Over the last decades, the availability of high-dimensional data sets across diverse disciplines such as as biostatistics, wireless communication and finance has transformed statistical practice. Rapid technological developments have led to large amounts of digital data. For example, computing speed and storage capability have exponentially grown, which enables users to collect, store and analyze data sets of very high dimension. While technological advances are helpful for users in innumerable aspects, it is an urgent challenge for statisticians to develop mathematically solid tools for high-dimensional inference. Indeed, traditional multivariate analysis, as outlined in the text books of Anderson (1984) or Muirhead (1982), is developed under the paradigm that the dimension is negligible compared to the sample size and breaks down seriously if this assumption is violated. Such problems have spurred the development of new analysis tools, that work for dimensions of the same order as and even larger than the sample size.
Turning closer to the scope of this work, random matrix theory is concerned with the study of the spectral behavior of various kinds of random matrices under the assumption that their dimension increases. Classical statistical guarantees derived in the case of fixed dimension often fail severely when considering the dimension as a growing parameter. Thus, a new class of limiting results is needed in order to meet the challenges of big data.
Since the middle of the 20th century, the development of random matrix theory was pushed by various applications, especially in the field of quantum mechanics. Simultaneously, mathematicians got attracted to the study of random matrices. In his pioneering work (Wigner, 1958), Wigner showed that the expected spectral distribution of a large Gaussian matrix converges to the semicircular law. Bai and Yin (1988) proved the limiting spectral distribution of a sample covariance matrix to be the semicircular law when the dimension is asymptotically negligible in comparison to the sample size. For the high-dimensional case, the work of Marčenko and Pastur (1967) is often considered as a breakthrough, establishing the limiting spectral distribution for a general class of sample covariance matrices. This work laid the foundation for a variety of follow-up works, including Bai et al. (1986); Grenander and Silverstein (1977); Wachter (1978, 1980); Yin (1986); Yin and Krishnaiah (1983); Yin et al. (1983), which assumed weaker conditions on the matrix entries

and showed that other types of random matrices obtain nonrandom limiting spectral distributions.
The study of random matrices is also vital from a statistical point of view (see the review of Paul and Aue, 2014). For a statistician, second-order limit theorems such as a central limit theorem for linear spectral statistics are of particular interest, and have therefore attracted increasing attention in the last two decades. Indeed, linear spectral statistics are frequently used to construct tests for various hypotheses. Here, the spectral properties of the sample covariance matrix are of particular interest. Given a sample from a high-dimensional data set, an important indicator for the interaction of the data is the sample covariance matrix. Estimation and inference for this crucial object are fundamental tasks of statistical analysis with numerous applications in biostatistics, wireless communications and finance (see, e.g., Fan and Li (2006), Johnstone (2006) and the references therein). In fact, many test statistics rely on a linear spectral statistic of the sample covariance matrix. For example, Mauchly (1940) proposed a likelihood ratio test for the hypothesis of sphericity (of a normal distribution), which has been extended by Gupta and Xu (2006) to the non-normal case and by Bai et al. (2009) and Wang and Yao (2013) to the high-dimensional case, where the dimension p is of the same order as the sample size n, that is $p/n \to y \in (0,1)$ as $p, n \to \infty$ (see also Theorem 9.12 in the monograph of Yao et al. (2015) for a further extension). Alternative tests based on distances between the sample covariance matrix and a multiple of the identity matrix have been considered in Ledoit and Wolf (2002) and Chen et al. (2010) among others. Fisher et al. (2010) suggested a generalization of John's test for sphericity, which is based on a ratio of arithmetic means of the eigenvalues of different powers of the sample covariance matrix. Among other testing problems such as sphericity, Jiang and Yang (2013) considered some classical q-sample testing problems under normality in a high-dimensional setting, which are further generalized in Jiang and Qi (2015); Qi et al. (2019); Dette and Dörnemann (2020); Guo and Qi (2021); Dörnemann (2022).
Because of its importance in statistics, numerous authors have investigated the asymptotic properties of linear spectral statistics from a more general perspective. An early reference is Jonsson (1982), who established a central limit theorem for the sum of the eigenvalues of a Wishart matrix raised to some power. In their pioneering paper, Bai and Silverstein (2004) proved a central limit theorem for linear spectral statistics of the form

$$\sum_{i=1}^{p} f(\lambda_i(\mathbf{B}_n))$$

of the sample covariance matrices $\mathbf{B}_n = \frac{1}{n}\sum_{i=1}^n \mathbf{T}_n^{1/2}\mathbf{x}_i\mathbf{x}_i^\star\mathbf{T}_n^{1/2}$ under rather general conditions, where $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are independent p-dimensional random vectors with independent real or complex valued (centered) entries x_{ij}, $\mathbf{T}_n$ is a $p \times p$ (non-random) Hermitian nonnegative definite matrix and $\lambda_1(\mathbf{B}_n) \leq \ldots \leq \lambda_p(\mathbf{B}_n)$ are the ordered eigenvalues of the matrix $\mathbf{B}_n$. Here, $\mathbf{x}_i^\star$ denotes the conjugate transpose of $\mathbf{x}_i$ for $1 \leq i \leq n$. While in the original work (Bai and Silverstein, 2004), the random variables x_{ij} were assumed to be independent and identically distributed, the assumption of identical distribution was weakened in Bai and Silverstein (2010). Several authors have followed this line of research and tried to relax the assumptions for such state-

ments (see Pan and Zhou, 2008; Lytova and Pastur, 2009; Shcherbina, 2011; Pan, 2014; Zheng et al., 2015b; Najim and Yao, 2016; Zou et al., 2021, among others). Other authors focused on linear spectral statistics of F-matrices (see, for example, Zheng, 2012; Zheng et al., 2017; Bodnar et al., 2019), auto-cross covariance (Jin et al., 2014), information-plus-noise matrices (Banna et al., 2020), large-dimensional matrices with bivariate dependence measures as entries (Bao et al., 2015a; Li et al., 2019) or sample correlation matrices (Gao et al., 2017; Parolya et al., 2021; Heiny and Parolya, 2021). Other recent works include Ji and Lee (2020) on deformed Wigner matrices, Li et al. (2021) on Kendall's rank correlation matrices and Wang and Yao (2021) on block-Wigner-type matrices, among many others. While most of this work deals with the case that the sample size is asymptotically proportional to the dimension ($p/n \to y \in (0,\infty)$), a CLT for linear spectral statistics of a rescaled version of the sample covariance matrix has been established for the ultra high-dimension case $p/n \to \infty$ (Chen and Pan, 2015; Qiu et al., 2021).
In the main part of this work, we will take a different point of view on linear spectral statistics and study these objects from a sequential perspective. More precisely, we consider a sequential version of the empirical covariance estimator

$$\mathbf{B}_{n,t} = \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} \mathbf{T}_n^{1/2}\mathbf{x}_i\mathbf{x}_i^\star\mathbf{T}_n^{1/2} \ , \quad 0 \leq t \leq 1, \tag{1.1}$$

and investigate the probabilistic properties of the stochastic process corresponding to linear spectral statistics of $\mathbf{B}_{n,t}$, that is,

$$S_t = \frac{1}{p}\sum_{i=1}^{p} f(\lambda_i(\mathbf{B}_{n,t})) \ , \quad 0 \leq t \leq 1, \tag{1.2}$$

where $\lambda_1(\mathbf{B}_{n,t}) \leq \ldots \leq \lambda_p(\mathbf{B}_{n,t})$ are the ordered eigenvalues of the matrix $\mathbf{B}_{n,t}$. In particular, we prove that for any $0 < t_0 < 1$, an appropriately normalized and centered version of the process $(S_t)_{t\in[t_0,1]}$ converges weakly to a non-standard Gaussian process.
Our interest in these processes is partially motivated by the central role of sequential covariance estimators in the detection of second-order structural breaks (see Aue et al., 2009; Dette and Gösmann, 2020, among others). In this field, various functionals of the process $(\mathbf{B}_{n,t})_{0\leq t\leq 1}$ have been studied in the case of fixed dimension, and we expect that results on the weak convergence of the process $(S_t)_{t\in[t_0,1]}$ will be useful in the context of change-point analysis for high-dimensional covariance matrices. In fact, we use the probabilistic results presented in this work to develop a procedure for monitoring deviations from sphericity, see Chapter 5 for more details.
Surprisingly, sequential processes of the form (1.2) have not found much attention in the literature. To our best knowledge, we are only aware of the works of D'Aristotile (2000) and Nagel (2020), who considered sequential aspects of large dimensional random matrices from a different point of view. More precisely, D'Aristotile (2000) studied a sequential process generated from the first $\lfloor nt\rfloor$ diagonal elements of a random matrix chosen according to the Haar measure on the unitary group of $n \times n$

matrices and showed that this process converges weakly to a standard complex-valued Brownian motion (see also D'Aristotile et al., 2003, for similar results). Recently, Nagel (2020) proved a functional central limit theorem for the sum of the first $\lfloor nt \rfloor$ diagonal elements of an $n \times n$ matrix $f(Z)$, where Z has an orthogonal or unitarily invariant distribution such that $\operatorname{tr}\big(f(Z)\big)$ satisfies a CLT. Compared to these results, the contributions of the present work are conceptually different, because, in contrast to the cited references, the parameter t used in the definition of the process (1.1) also appears in the eigenvalues $\lambda_i(\mathbf{B}_{n,t})$. This "non-linearity" results in a substantially more complicated structure of the problem. In particular, the limiting processes of $(S_t)_{t\in[t_0,1]}$ are non-standard Gaussian processes (except for the simplest case $f(x) = x$ and $\mathbf{T}_n = \mathbf{I}$), and the proofs of our results (in particular the proof of tightness) require an extended machinery, which has so far not been considered in the literature on linear spectral statistics. As a consequence, we provide a substantial generalization of the classical CLT for linear spectral statistics (see, for example, Bai and Silverstein, 2010), which is obtained from the process convergence of $(S_t)_{t\in[t_0,1]}$ (appropriately standardized) via continuous mapping.

This thesis is structured as follows. In Chapter 2, we lay the mathematical foundation for the study of linear spectral statistics and recall important results from the literature. Chapter 3 is dedicated to the main result about linear spectral statistics of sequential sample covariance matrices, which is applied for some special cases in Chapter 4. A statistical application to a change-point problem for large covariance matrices can be found in Chapter 5. The proof of the main result formulated in Chapter 3 is challenging and deferred to Chapters 6 and 7. A strategy for the proof can be found in Section 6.1. The applications are proven in Chapter 8. Moreover, a central limit theorem for the diagonal entries of the inverse sample covariance matrix and its connection to linear spectral statistics are presented in Chapter 9.

In agreements with the doctoral regulations of the Faculty of Mathematics of the Ruhr-University Bochum, parts of this thesis are still under review by a journal. More precisely, the mathematical theory presented in Chapters 3 to 8 is based on Dörnemann and Dette (2021) and was submitted for publication. Chapter 9 is based on a manuscript, which has not been submitted so far. Both projects are based on joint work with my supervisor Holger Dette.

Chapter 2

Linear spectral statistics

In this section, we lay the mathematical foundation for the study of linear spectral statistics. We introduce the most important objects and present well-known results on the spectrum of the sample covariance matrix, including the Marčenko–Pastur regime.

2.1 Mathematical foundations

Spectral norm and diagonal matrix

For any matrix $\mathbf{A} \in \mathbb{C}^{p\times p}$, the spectral (or operator) norm $\|\mathbf{A}\|$ is the square root of the largest eigenvalue of $\mathbf{A}\mathbf{A}^\star$. Moreover, $\operatorname{diag}(\mathbf{A})$ denotes the diagonal matrix which has the same diagonal as $\mathbf{A}$.

Empirical and limiting spectral distribution

For any matrix $\mathbf{A} \in \mathbb{C}^{p\times p}$ with real eigenvalues, we denote its ordered eigenvalues by

$$\lambda_1(\mathbf{A}) \geq \cdots \geq \lambda_p(\mathbf{A})\,.$$

Hence, we have $\|\mathbf{A}\| = \sqrt{\lambda_1(\mathbf{A}\mathbf{A}^\star)}$. The *empirical spectral distribution* of $\mathbf{A}$ is defined by

$$F^{\mathbf{A}} = \frac{1}{p}\sum_{i=1}^{p} \delta_{\lambda_i(\mathbf{A})},$$

where δ_a denotes the Dirac mass at some point $a \in \mathbb{R}$. Let $(F^{\mathbf{A}_p})_{p\in\mathbb{N}}$ be a sequence of empirical spectral distributions for (random) matrices $\mathbf{A}_p \in \mathbb{C}^{p\times p}$, $p \in \mathbb{N}$, with real eigenvalues. The non-random distribution F is called *limiting spectral distribution* of the sequence $(\mathbf{A}_p)_{p\in\mathbb{N}}$ if $(F^{\mathbf{A}_p})_{p\in\mathbb{N}}$ converges weakly to F (almost surely).

Linear spectral statistic

Let $\mathbf{A} \in \mathbb{C}^{p\times p}$ be a random matrix with real eigenvalues and f be a function defined on the support of $F^{\mathbf{A}}$. The statistic

$$\frac{1}{p}\sum_{i=1}^{p} f(\lambda_i(\mathbf{A})) = \int f(x)dF^{\mathbf{A}}(x).$$

is called *linear spectral statistic* of the matrix $\mathbf{A}$.

Stieltjes transform

Let μ be a finite measure on the real line. Its *Stieltjes transform* s_μ is defined as

$$s_\mu(z) = \int \frac{1}{x-z}\mu(dx), \; z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}.$$

If $\mu = F^{\mathbf{A}}$ is an empirical spectral distribution, then its Stieltjes transform has the form

$$s_{F^{\mathbf{A}}}(z) = \frac{1}{p}\,\mathrm{tr}\left\{(\mathbf{A} - z\mathbf{I})^{-1}\right\}, z \in \mathbb{C}^+,$$

where the matrix $(\mathbf{A} - z\mathbf{I})^{-1}$ is called the resolvent.
The Stieltjes transform s_F of a distribution F characterizes F uniquely and F being the limiting spectral distribution of a sequence $(\mathbf{A}_p)_{p\in\mathbb{N}}$ of random matrices with real eigenvalues is equivalent to the convergence

$$\lim_{p\to\infty} s_{F^{\mathbf{A}_p}}(z) = s_F(z) \text{ almost surely}$$

of the corresponding Stieltjes transforms for all $z \in \mathbb{C}^+$. By Cauchy's integral formula (see, e.g., Ahlfors, 1953), we write

$$\int f(x)dG(x) = \frac{1}{2\pi i}\int\int_{\mathcal{C}} \frac{f(z)}{z-x}dzdG(x) = -\frac{1}{2\pi i}\int_{\mathcal{C}} f(z)s_G(z)dz,$$

where G is an arbitrary cumulative distribution function (c.d.f.) with a compact support, f is an arbitrary analytic function on an open set, say O, containing the support of G, $\mathcal{C}$ is a positively oriented contour in O enclosing the support of G. This equation draws an important connection to linear spectral statistics and is usually used when proving a central limit theorem for linear spectral statistics. We will also pursue this approach when proving the main result of this thesis. For a more detailed discussion of this approach, we refer the reader to Section 6.1.

Marčenko–Pastur law

The *Marčenko–Pastur distribution* F^{y,σ^2} with index $y \in (0,\infty)$ and scale parameter $\sigma^2 > 0$ has the density function

$$f(x) = \begin{cases} \frac{1}{2\pi x y \sigma^2}\sqrt{(b-x)(x-a)}, & \text{if } a \le x \le b, \\ 0, & \text{otherwise,} \end{cases}$$

with an additional point mass at the origin of value $1 - 1/y$ if $y > 1$, where

$$a = \sigma^2(1 - \sqrt{y})^2, \ b = \sigma^2(1 + \sqrt{y})^2.$$

Let $y \in (0, \infty)$ and H be a non-random distribution. The so-called *generalized Marčenko-Pastur distribution* $F^{y,H}$ is characterized through its Stieltjes transform $s = s_{F^{y,H}}$, which is the unique solution of the equation

$$s(z) = \int \frac{1}{\lambda(1 - y - yzs(z)) - z} dH(\lambda), \ z \in \mathbb{C}^+. \tag{2.1}$$

The fundamental equation (2.1) is called the Marčenko–Pastur equation for historical reasons. Defining

$$\underline{s}(z) = -\frac{1-y}{z} + ys(z), \ z \in \mathbb{C}^+,$$

then we can write (2.1) equivalently as

$$z = -\frac{1}{\underline{s}(z)} + y \int \frac{1}{1 + \underline{s}(z)\lambda} dH(\lambda),$$

which is referred to as the Silverstein equation.

(Sequential) sample covariance matrix

Let $\mathbf{T}_n \in \mathbb{C}^{p \times p}$ be a non-negative definite Hermitian matrix with Hermitian square root $\mathbf{T}_n^{1/2} \in \mathbb{C}^{p \times p}$ and $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{C}^p$ be a sample of (real or complex valued) random vectors with centered and standardized entries. Then the random matrix

$$\mathbf{B}_{n,t} = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathbf{T}_n^{1/2} \mathbf{x}_i \mathbf{x}_i^\star \mathbf{T}_n^{1/2} \in \mathbb{C}^{p \times p}, \ t \in [0, 1],$$

is called the *sequential sample covariance matrix* of $\mathbf{T}_n^{1/2}\mathbf{x}_1, \ldots, \mathbf{T}_n^{1/2}\mathbf{x}_n$. The matrix $\mathbf{B}_{n,1}$ is called the *sample covariance matrix* of $\mathbf{T}_n^{1/2}\mathbf{x}_1, \ldots, \mathbf{T}_n^{1/2}\mathbf{x}_n$. The matrix $\mathbf{T} = \mathbf{T}_n$ is referred to as the *population covariance matrix.* We define for $\mathbf{B}_{n,t}$ the $(\lfloor nt \rfloor \times \lfloor nt \rfloor)$-dimensional companion matrix through

$$\underline{\mathbf{B}}_{n,t} = \frac{1}{n} \mathbf{X}_{n,t}^\star \mathbf{T}_n \mathbf{X}_{n,t}, \tag{2.2}$$

where the data matrix $\mathbf{X}_{n,t}$ is defined as

$$\mathbf{X}_{n,t} = (\mathbf{x}_1, \ldots, \mathbf{x}_{\lfloor nt \rfloor}) \in \mathbb{C}^{p \times \lfloor nt \rfloor}, \ t \in [0, 1].$$

Note that both matrices $(n/\lfloor nt \rfloor)\mathbf{B}_{n,t}$ and $(n/\lfloor nt \rfloor)\underline{\mathbf{B}}_{n,t}$ have the same non-vanishing eigenvalues and consequently their empirical spectral distributions satisfy

$$\lfloor nt \rfloor F^{(n/\lfloor nt \rfloor)\underline{\mathbf{B}}_{n,t}} - pF^{(n/\lfloor nt \rfloor)\mathbf{B}_{n,t}} = (\lfloor nt \rfloor - p) I_{[0,\infty)}. \tag{2.3}$$

2.2 Limiting spectral distribution of the sample covariance matrix

A variety of important test statistics can be written as a function of linear spectral statistics of the sample covariance matrix (see discussion in Chapter 1). Its limiting spectral distribution was found by Marčenko and Pastur (1967) and this result has been generalized by various authors.

Theorem 2.2.1 (Theorem 2.9 in Yao et al. (2015)) *Let $\mathbf{T} = \mathbf{I}$. Suppose that the entries x_{ij} of the data matrix $\mathbf{X}_{n,1}$ are i.i.d. complex random variables with $\mathbb{E}[x_{11}] = 0$ and $\mathbb{E}|x_{11}|^2 = \sigma^2$ and $p/n \to y \in (0, \infty)$ as $n \to \infty$. Then, almost surely, the empirical spectral distribution $(F^{B_{n,1}})_{n\in\mathbb{N}}$ of the sample covariance matrices $(\mathbf{B}_{n,1})_{n\in\mathbb{N}}$ converges weakly to the Marčenko–Pastur law F^{y,σ^2}.*

The following example illustrates the different asymptotic regimes for linear spectral statistics depending on the dimension-to-sample-size ratio.

Example 2.2.2 Let $\mathbf{x}_1, \ldots, \mathbf{x}_n \sim \mathcal{N}_p(\mathbf{0}, \mathbf{I})$ be a independent sample from a p-dimensional normal distribution with mean vector $\mathbf{0} = (0, \ldots, 0)^\top \in \mathbb{R}^p$ and covariance matrix $\mathbf{I} \in \mathbb{R}^{p\times p}$. For $\mathbf{T} = \mathbf{I}$, consider the statistic

$$T_n = \log|\mathbf{B}_{n,1}| = \sum_{j=1}^{p} \log\left(\lambda_j(\mathbf{B}_{n,1})\right),$$

known as the generalized variance. If the dimension p is fixed, then we get from Example 1.1 in Yao et al. (2015) the convergence

$$\sqrt{\frac{n}{p}} T_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2), \ n \to \infty.$$

However, if the dimension increases with the sample size at the same rate, that is, $p/n \to y \in (0, 1)$ for $n \to \infty$, we observe a completely different asymptotic behavior of T_n. More precisely, using Theorem 2.2.1 given above and Example 2.11 in Yao et al. (2015), it is seen that almost surely

$$\frac{T_n}{p} \to \int_0^\infty \log(x) dF^{y,1}(x) = \frac{y-1}{y} \log(1-y) - 1 < 0,$$

and consequently,

$$\sqrt{\frac{n}{p}} T_n \to -\infty \text{ almost surely}, \ n \to \infty.$$

This toy example indicates that the asymptotic properties of linear spectral statistics depend crucially on the order of the dimension p in comparison to the sample size n. Note that the asymptotic properties of the process $(\log|\mathbf{B}_{n,t}|)_{t\in[t_0,1]}$ for some $t_0 > 0$ are examined in Corollary 4.3.1 in Section 4.3.

Theorem 2.2.1 deals with the simplest case $\mathbf{T} = \mathbf{I}$. The following result provides the limiting spectral distribution for more complex-structured population covariance matrices and is a consequence of Theorem 1.1 in Bai and Zhou (2008).

Theorem 2.2.3 *Assume that $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{C}^p$ are independent random vectors with centered and standardized entries with a common finite fourth moment and $p/n \to y \in (0, \infty)$. Let $\mathbf{T}_n \in \mathbb{C}^{p \times p}$ be a population covariance matrix such that $F^{\mathbf{T}_n}$ converges weakly to some non-random distribution H. Then, almost surely, $F^{\mathbf{B}_{n,1}}$ converges weakly to the generalized Marčenko–Pastur law $F^{y,H}$ for $n \to \infty$.*

The conditions of the previous theorem are adapted to the setting of this thesis, where the fourth moment will be assumed to exist. For example, another version assuming finite variance and i.i.d. data can be found in Yao et al. (2015) (Theorem 2.14). Due to Theorem 2.2.1 and Theorem 2.2.3, the condition $p/n \to y \in (0, \infty)$ is also called the Marčenko–Pastur regime.

If $t \in (0, 1]$ is fixed, we conclude for the limiting spectral distribution of $(\mathbf{B}_{n,t})_{n \in \mathbb{N}}$ from Theorem 2.2.3 that

$$\tilde{F}^{y_t,H}(x) := \lim_{n \to \infty} F^{\mathbf{B}_{n,t}}(x) = F^{y_t,H}(x/t) \text{ almost surely} \tag{2.4}$$

at all points $x \in \mathbb{R}$, where $\tilde{F}^{y_t,H}$ is continuous. Here, H denotes the weak limit of $(H_n)_{n \in \mathbb{N}} = (F^{\mathbf{T}_n})_{n \in \mathbb{N}}$. In other words, H is the limiting spectral distribution of $(\mathbf{T}_n)_{n \in \mathbb{N}}$.

Chapter 3

A sequential perspective on linear spectral statistics

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ be independent p-dimensional random vectors with real or complex entries and covariance matrix given by the identity matrix $\mathbf{I} = \mathbf{I}_p \in \mathbb{R}^{p \times p}$. We use the notation $\mathbf{x}_j = (x_{1j}, \ldots, x_{pj})^\top$ for the components of $\mathbf{x}_j$ and assume that $\mathbb{E}[x_{ij}] = 0$ and $\mathbb{E}[x_{ij}^2] = 1$. When considering asymptotics, the dimension $p = p_n$ of the data is allowed to increase with the sample size $n \to \infty$ at the same order, that is, $p/n \to y \in (0, \infty)$ as $n \to \infty$. Recall the notation of the sequential covariance estimator $\mathbf{B}_{n,t}$ in (1.1) and consider the corresponding linear spectral statistic (as a function of t)

$$S_t = \frac{1}{p} \operatorname{tr} \left(f(\mathbf{B}_{n,t}) \right) = \frac{1}{p} \sum_{j=1}^{p} f \left(\lambda_i(\mathbf{B}_{n,t}) \right), \ t \in [0, 1],$$

where f is an appropriate function defined on a subset of the complex plane. For a given $t_0 \in (0, 1]$, we are interested in the asymptotic properties of the process $(S_t)_{t \in [t_0, 1]}$ and will prove a weak convergence result for an appropriately standardized version of this process in the space $\ell^\infty([t_0, 1])$ of bounded functions defined on the interval $[t_0, 1]$. Note that the random variable S_1 has been studied intensively in the literature (see the discussion in Chapter 1). Before providing preliminaries for the main result, we present an example for the simplest case $f(x) = x$ and $\mathbf{T} = \mathbf{I}$.

Example 3.0.1 (Trace of the sequential sample covariance matrix) Assume that the real valued random variables x_{ij} are independent with $\mathbb{E}[x_{ij}] = 0$, $\mathbb{E}[x_{ij}^2] = 1$, and common fourth moment $\mathbb{E}[x_{ij}^4] = \nu_4 < \infty$, where $1 \leq i \leq p, 1 \leq j \leq n$, and set $\mathbf{T} = \mathbf{I}$. Let $y = \lim_{n \to \infty} p/n \in (0, \infty)$. Consider the linear spectral statistic

$$T_n(t) = \frac{1}{p} \operatorname{tr} \left(\mathbf{B}_{n,t} \right) = \frac{1}{pn} \sum_{i=1}^{\lfloor nt \rfloor} ||\mathbf{x}_i||_2^2 = \frac{1}{pn} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{p} x_{ij}^2.$$

For the mean, we calculate

$$\mathbb{E} T_n(t) = \frac{1}{pn} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{p} \mathbb{E} x_{ij}^2 = \frac{\lfloor nt \rfloor}{n}.$$

Moreover, we may write

$$\frac{\sqrt{n}}{\sqrt{p}}p(T_n(t) - \frac{\lfloor nt \rfloor}{n}) = \frac{\sqrt{n}}{\sqrt{p}}\left(\frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor}\sum_{j=1}^{p}(x_{ij}^2 - 1)\right)$$
$$= \frac{1}{\sqrt{pn}}\sum_{i=1}^{\lfloor nt \rfloor}\sum_{j=1}^{p}(x_{ij}^2 - 1).$$

Note that

$$\operatorname{cov}\left(\frac{1}{\sqrt{pn}}\sum_{i=1}^{\lfloor nt \rfloor}\sum_{j=1}^{p}x_{ij}^2, \frac{1}{\sqrt{pn}}\sum_{k=1}^{\lfloor ns \rfloor}\sum_{l=1}^{p}x_{ij}^2\right) = \frac{(\nu_4 - 1)(\lfloor ns \rfloor \wedge \lfloor nt \rfloor)}{n}.$$

By an application of Donsker's invariance principle for triangular arrays (Theorem 2.12.6, Van Der Vaart and Wellner, 1996), we have for $n \to \infty$ the weak convergence

$$\left(\frac{\sqrt{n}}{\sqrt{p}}p(T_n(t) - \frac{\lfloor nt \rfloor}{n})\right)_{t\in[0,1]} \rightsquigarrow \sqrt{\nu_4 - 1}(W_t)_{t\in[0,1]}$$

in the space $\ell^\infty[0,1]$ of bounded functions, where $(W_s)_{s\in[0,1]}$ denotes a standard Brownian motion. This implies

$$\left(p\left(T_n(t) - \frac{\lfloor nt \rfloor}{n}\right)\right)_{t\in[0,1]} \rightsquigarrow \sqrt{(\nu_4 - 1)y}(W_t)_{t\in[0,1]} \text{ in } \ell^\infty[0,1].$$

Note that the limiting process is centered and admits the following covariance function for $s, t \in [0,1]$

$$\operatorname{cov}(\sqrt{(\nu_4 - 1)y}W_s, \sqrt{(\nu_4 - 1)y}W_t) = (\nu_4 - 1)y(s \wedge t).$$

3.1 Preliminaries

For the following discussion, recall the definition of the $(\lfloor nt \rfloor \times \lfloor nt \rfloor)$-dimensional companion matrix $\underline{\mathbf{B}}_{n,t}$ given in (2.2) and denote the limiting spectral distribution (if it exists) of $F^{\underline{\mathbf{B}}_{n,t}}$ and its corresponding Stieltjes transform by

$$\underline{\tilde{F}}^{y_t,H} \quad \text{and} \quad \underline{\tilde{s}}_t(z) = s_{\underline{\tilde{F}}^{y_t,H}}(z), \tag{3.1}$$

respectively. A straightforward calculation using (2.1) (for details, see Lemma 7.1.6) shows that this Stieltjes transform satisfies the equation

$$z = -\frac{1}{\underline{\tilde{s}}_t(z)} + y\int \frac{\lambda}{1 + \lambda t \underline{\tilde{s}}_t(z)} dH(\lambda). \tag{3.2}$$

Recalling (2.3), we note that $F^{(n/\lfloor nt\rfloor)\mathbf{B}_{n,t}}$ has the limit $F^{y_t,H}$ as $y_{\lfloor nt\rfloor} = p/\lfloor nt\rfloor \to y_t \in (0,\infty)$ if and only if $F^{(n/\lfloor nt\rfloor)\underline{\mathbf{B}}_{n,t}}$ has a limit $\underline{F}^{y_t,H}$ and in this case

$$\underline{F}^{y_t,H} - y_t F^{y_t,H} = (1-y_t)I_{[0,\infty)}.$$

For $z \in \mathbb{C}^+$ we therefore obtain a relation between the corresponding Stieltjes transforms

$$\underline{s}_t(z) = -\frac{1-y_t}{z} + y_t s_t(z). \tag{3.3}$$

As an empirical version of $\underline{F}^{y_t,H}$, we further define the distribution $\underline{F}^{y_{\lfloor nt\rfloor},H_n}$ through

$$\underline{F}^{y_{\lfloor nt\rfloor},H_n} - y_{\lfloor nt\rfloor} F^{y_{\lfloor nt\rfloor},H_n} = (1-y_{\lfloor nt\rfloor})I_{[0,\infty)},$$

and consider its rescaled version

$$\tilde{\underline{F}}^{y_{\lfloor nt\rfloor},H_n}(x) = \underline{F}^{y_{\lfloor nt\rfloor},H_n}\Big(\frac{n}{\lfloor nt\rfloor}x\Big), \ x \in \mathbb{R}. \tag{3.4}$$

3.2 Main result

Our main result provides the asymptotic properties of the process $(X_n(f,t))_{t\in[t_0,1]}$, where $t_0 \in (0,1]$, f is a given function,

$$X_n(f,t) = \int f(x)dG_{n,t}(x), \tag{3.5}$$

the process $G_{n,t}$ is defined by

$$G_{n,t}(x) = p\big(F^{\mathbf{B}_{n,t}}(x) - \tilde{F}^{y_{\lfloor nt\rfloor},H_n}(x)\big), \ t \in [t_0,1],$$

and

$$\tilde{F}^{y_{\lfloor nt\rfloor},H_n}(x) = F^{y_{\lfloor nt\rfloor},H_n}\left(\frac{n}{\lfloor nt\rfloor}x\right), \ x \in \mathbb{R}, \tag{3.6}$$

is a rescaled version of the generalized Marčenko-Pastur distribution defined by (2.1). In the following theorem, we make use of the notion of weak convergence in the space ℓ^∞ of bounded functions. For a detailed definition of this concept, we refer the reader to Chapter 1 in Van Der Vaart and Wellner (1996). (Note that in contrast to this monograph, we sometimes deal with complex-valued functions by considering real and imaginary part.) The proof of the following result is challenging and therefore deferred to Chapter 6 and Chapter 7. An overview about the main steps is given in Section 6.1.

Theorem 3.2.1 *Assume that $p/\lfloor nt\rfloor \to y_t = y/t \in (0,\infty)$ and that the following additional conditions are satisfied:*

(a) For each n, the random variables $x_{ij} = x_{ij}^{(n)}$ are independent with $\mathbb{E}x_{ij} = 0$, $\mathbb{E}|x_{ij}|^2 = 1$, $\max_{i,j,n} \mathbb{E}|x_{ij}|^{12} < \infty$.

(b) $(\mathbf{T}_n)_{n\in\mathbb{N}}$ *is a sequence of* $p \times p$ *Hermitian non-negative definite matrices with bounded spectral norm and the sequence of spectral distributions* $(F^{\mathbf{T}_n})_{n\in\mathbb{N}}$ *converges to a proper c.d.f.* H.

(c) *Let* $t_0 \in (0, 1]$ *and* f_1, f_2 *be functions, which are analytic on an open region containing the interval*

$$\Big[\liminf_{n\to\infty} \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{t_0}) t_0 (1 - \sqrt{y_{t_0}})^2, \limsup_{n\to\infty} \lambda_1(\mathbf{T}_n)(1 + \sqrt{y_{t_0}})^2\Big]. \quad (3.7)$$

(1) *If the random variables* x_{ij} *are real and* $\mathbb{E}[x_{ij}^4] = 3$, *then the process*

$$(X_n(f_1, t), X_n(f_2, t))_{t\in[t_0,1]}$$

converges weakly to a Gaussian process $(X(f_1, t), X(f_2, t))_{t\in[t_0,1]}$ *in the space* $(\ell^\infty([t_0, 1]))^2$ *with means*

$$\mathbb{E}[X(f_i, t)] = -\frac{1}{2\pi i} \int_{\mathcal{C}} f_i(z) \frac{ty \int \frac{\underline{\tilde{s}}_t^3(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^3} dH(\lambda)}{\Big(1 - ty \int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^2} dH(\lambda)\Big)^2} dz \ , \quad i = 1, 2,$$

and covariance kernel

$$\operatorname{cov}(X(f_1, t_1), X(f_2, t_2)) = \frac{1}{2\pi^2} \int_{\mathcal{C}_1} \int_{\mathcal{C}_2} f_1(z_1) \overline{f_2(z_2)} \sigma^2_{t_1,t_2}(z_1, \overline{z_2}) \overline{dz_2} dz_1,$$

where $\mathcal{C}, \mathcal{C}_1, \mathcal{C}_2$ *are arbitrary closed, positively orientated contours in the complex plane enclosing the interval in* (3.7), $\mathcal{C}_1, \mathcal{C}_2$ *are non overlapping and the function* $\sigma^2_{t_1,t_2}(z_1, z_2)$ *is defined in* (6.44).

(2) *If the random variables* x_{ij} *are complex with* $\mathbb{E}x_{ij}^2 = 0$ *and* $\mathbb{E}|x_{ij}|^4 = 2$, *then (1) also holds with means* $\mathbb{E}[X(f_i, t)] = 0$, $i = 1, 2$, *and covariance structure*

$$\operatorname{cov}(X(f_1, t_1), X(f_2, t_2)) = \frac{1}{4\pi^2} \int_{\mathcal{C}_1} \int_{\mathcal{C}_2} f_1(z_1) \overline{f_2(z_2)} \sigma^2_{t_1,t_2}(z_1, \overline{z_2}) \overline{dz_2} dz_1.$$

Remark 3.2.2 While linear spectral statistics have been studied intensively for sample covariance matrices (see, for example, Bai and Silverstein, 2004, 2010), very little effort has been done in a sequential framework so far. In contrast to these "classical" CLTs, the sequential version in Theorem 3.2.1 reveals the asymptotic behavior of the whole process of linear spectral statistics corresponding to the sequential empirical covariance process (1.1) and thus provides a substantial generalization of its one-dimensional versions. In particular, the limiting process is not a standard Gaussian process and the proofs require an extended machinery and some additional assumptions.

(1) While assumptions such as (3.7) and on the spectrum of the population covariance matrix $\mathbf{T}_n$ are common even for a standard CLT of non-sequential

linear spectral statistics, we should have a closer look at the moment assumptions. Among many other technical challenges, the most delicate part of the proof of Theorem 3.2.1 lies in controlling the process $(X_n(f,t))_t$ of linear spectral statistics in terms of (asymptotic) tightness, which enforces higher-order moment conditions in order to find sharper bounds for the concentration of random quadratic forms of the type

$$\mathbf{x}_j^\star \mathbf{A} \mathbf{x}_j - \operatorname{tr}(\mathbf{A}), \tag{3.8}$$

where $\mathbf{A}$ denotes a random $p \times p$ matrix independent of $\mathbf{x}_j$, $j \in \{1, \dots, n\}$. In particular, the existence of the 12th-moment in Theorem 3.2.1 is exclusively needed for the proof of asymptotic tightness and is not used for the proof of convergence of the finite-dimensional distributions (for details, see Section 6.3.3). Strengthening the moment conditions on the underlying random variables appears to be a convenient tool for investigating linear spectral statistics of non-standard random matrices. For example, in the work of Banna et al. (2020), the authors consider linear spectral statistics of random information-plus-noise matrices and assume the existence of the 16th moment for deriving a non-sequential CLT for linear spectral statistics corresponding to this type of random matrices. Consequently, the higher-order moment condition implies stronger bounds for the moments of random quadratic forms of the type (3.8) (see their Lemma A.2 for more details).
Moreover, note that our condition on the 12th moment implies the Lindeberg-type condition (9.7.2) in the work of Bai and Silverstein (2010).

(2) In order to allow for non-centralized data ($\mathbb{E}[x_{ij}] \neq 0$), Zheng et al. (2015b) prove a substitution principle for linear spectral statistics of recentered sample covariance matrices and, thus, weakening the conditions of Bai and Silverstein's CLT. We expect that it is possible to pursue such a generalization of Theorem 3.2.1 combining the tools developed in this work with the methodology used in the proof of Theorem 3.2.1.

(3) Furthermore, it might be of interest to relax the Gaussian-type 4th moment condition. When allowing for a general finite 4th moment, additional terms for the covariance structure and the bias arise, whose convergence is not guaranteed under the assumptions of Theorem 3.2.1. In fact, in this case those terms depend also on the eigenvectors of the population covariance matrix $\mathbf{T}_n$, which are not controlled under the conditions of Theorem 3.2.1. For instance, in the non-sequential case, Najim and Yao (2016) show that the Lévy–Prohorov distance between the linear statistics' distribution and a normal distribution, whose mean and variance may diverge, vanishes asymptotically, while Pan (2014) imposes additional conditions on $\mathbf{T}_n$ in order to ensure convergence of the additional terms for mean and covariance. For the sequential version considered in this thesis, it seems to be promising to derive the convergence of such additional terms under similar conditions on $\mathbf{T}_n$ as used by Pan (2014) for a proof of a "classical" CLT.

Chapter 4

Detailed applications of our main theorem

In general, the calculation of the limiting parameters appearing in Theorem 3.2.1 might be involved, since mean and covariance are given by contour integrals and rely on the Stieltjes transform $\underline{\tilde{s}}_t(z)$, which is defined implicitly by an equation involving the limiting spectral distribution H (see (3.2)) and has in general no closed form. In the case $\mathbf{T}_n = \mathbf{I}$, these integrals can be interpreted as integrals over the unit circle (see Proposition 4.1.1 in Section 4.1), and for specific functions f_1 and f_2 an explicit calculation of the asymptotic expectation and variance in Theorem 3.2.1 is possible. We present some examples in Section 4.2 and Section 4.3.

4.1 How to calculate mean and covariance

The following result provides essential formulas for the calculation of the mean and covariance structure in Theorem 3.2.1 in the case $\mathbf{T}_n = \mathbf{I}$ and is proven in Section 8.2.1. It generalizes the formulas given in Proposition A.1 in Wang and Yao (2013) and Proposition 3.6 in Yao et al. (2015).

Proposition 4.1.1 *Let $h_t = \sqrt{y_t} \in (0, \infty)$ and $\mathbf{T}_n = \mathbf{I}$ and let f_1 and f_2 be functions which are analytic on an open region containing the interval in (3.7). For the process $\big(X(f_1, t_1), X(f_2, t)\big)_{t \in [t_0, 1]}$ given in Theorem 3.2.1, we have the following formulas*

$$\mathbb{E}[X(f_i,t)] = \frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|\xi|=1} f(t(1 + h_t r \xi + h_t r^{-1} \xi^{-1} + h_t^2)) \Big(\frac{\xi}{\xi^2 - r^{-2}} - \frac{1}{\xi} \Big) d\xi,$$

$$\operatorname{cov}(X(f_1,t_1), X(f_2,t_2)) = \frac{1}{2\pi^2} \lim_{\substack{r_2 > r_1, \\ r_1, r_2 \searrow 1}} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} f_1(t_1(1 + h_{t_1} r_1 \xi_1 + h_{t_1} r_1^{-1} \xi_1^{-1} + h_{t_1}^2))$$
$$\times \overline{f_2(t_2(1 + h_{t_2} r_2 \xi_2^{-1} + h_{t_2} r_2^{-1} \xi_2 + h_{t_2}^2))} \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1$$

where $t, t_1, t_2 \in [t_0, 1]$ with $t_2 \leq t_1$

$$g_1(\xi_1,\xi_2) = - \Big(h_1 h_2 r_1 r_2 \Big\{ h_2^4 r_1^2 r_2^2 t_2^2 \xi_1^2 \xi_2^2 + 2 h_2^3 r_1^2 r_2 t_2 \xi_1^2 \xi_2 (r_2^2 t_1 + t_2 \xi_2^2)$$

$$
\begin{aligned}
&- 2h_1h_2r_1r_2t_1\xi_1\xi_2(r_2^2t_1(2+h_1r_1\xi_1)+r_1t_2\xi_1(h_1+2r_1\xi_1)\xi_2^2\Big\} \\
&+ h_1^2t_1\xi_2^2\Big\{r_2^2t_1(1+2h_1r_1\xi_1+3r_1^2\xi_1^2+h_1^2r_1^2\xi_1^2+2h_1r_1^3\xi_1^3)+r_1^2t_2\xi_1^2\ (-1+r_1^2\xi_1^2)\xi_2^2\Big\} \\
&+ h_2^2\Big\{r_2^4t_1^2-r_2^2t_1t_2(1+2h_1r_1\xi_1-3r_1^2\xi_1^2+2h_1^2r_1^2\xi_1^2+2h_1r_1^3\xi_1^3)\xi_2^2+r_1^2t_2^2\xi_1^2\xi_2^4\Big\}\Big),
\end{aligned}
$$

$$
\begin{aligned}
g_2(\xi_1,\xi_2) =& \big(h_2r_2-h_1r_1\xi_1\xi_2\big)^2 \\
&\times \big(h_2^2r_1r_2t_2\xi_1\xi_2-h_1r_2t_1(1+h_1r_1\xi_1+r_1^2\xi_1^2)\xi_2+h_2r_1\xi_1(r_2^2t_1+t_2\xi_2^2)\big)^2.
\end{aligned}
$$

In the complex case, we have $\mathbb{E}[X(f_i,t)] = 0,\ i = 1,2,$ *and the covariance structure is given by* $1/2$ *times the covariance structure for the real case.*

4.2 First applications

We apply Theorem 3.2.1 for the special case $f_1(x) = x, f_2(x) = x^2, \mathbf{T}_n = \mathbf{I}$, which is motivated by the statistical test presented in Chapter 5. A proof can be found in Section 8.2.2.

Corollary 4.2.1 *Let* $t_0 > 0$. *Assume that the random variables* x_{ij} *satisfy condition (a) from Theorem 3.2.1 for* $1 \le i \le p, 1 \le j \le n$ *and* $p/n \to y \in (0,\infty)$ *as* $n \to \infty$. *Then it holds:*

1. *If the variables* x_{ij} *are real and* $\mathbb{E}x_{ij}^4 = 3$, *then the sequence*

$$\big((X_n(f_1,t))_{t\in[t_0,1]},(X_n(f_2,t))_{t\in[t_0,1]}\big)_{n\in\mathbb{N}}$$

with

$$
\begin{aligned}
X_n(f_1,t) &= \operatorname{tr}(\mathbf{B}_{n,t}) - \lfloor nt\rfloor y_n, \\
X_n(f_2,t) &= \operatorname{tr}\left(\mathbf{B}_{n,t}^2\right) - \lfloor nt\rfloor y_n\left(\frac{\lfloor nt\rfloor}{n}+y_n\right),\ t\in[t_0,1],
\end{aligned}
$$

converges weakly to a Gaussian process $\big((X(f,t))_{t\in[t_0,1]},(X(g,t))_{t\in[t_0,1]}\big)$ *in the space* $(\ell^\infty([t_0,1]))^2$ *with means*

$$\mathbb{E}[X(f_1,t)] = 0,\ \mathbb{E}[X(f_2,t)] = ty,$$

and covariance function for $t_1,t_2 \in [t_0,1]$

$$
\begin{aligned}
&\operatorname{cov}(X(f_1,t_1),X(f_1,t_2)) = 2y\min(t_1,t_2), \\
&\operatorname{cov}(X(f_2,t_1),X(f_2,t_2)) = 4\min(t_1,t_2)y\left\{2t_1t_2+[\min(t_1,t_2)+2(t_1+t_2)]\,y+2y^2\right\}, \\
&\operatorname{cov}(X(f_1,t_1),X(f_2,t_2)) = 4\min(t_1,t_2)y(t_2+y).
\end{aligned}
$$

2. *If* x_{ij} *are complex with* $\mathbb{E}x_{ij}^2 = 0$ *and* $\mathbb{E}|x_{ij}|^4 = 2$, *then (1) also holds with means* $\mathbb{E}[X(f,t)] = \mathbb{E}[X(g,t)] = 0$ *and covariance structure given by* $1/2$ *times the covariance structure given for the real case.*

Remark 4.2.2 (Special cases of Corollary 4.2.1) 1. A straight-forward application of Donsker's Theorem yields the convergence $(X_n(f_1,t))_{t\in[0,1]} \rightsquigarrow (X(f_1,t))_{t\in[0,1]}$ in $\ell^\infty[0,1]$ for the real case. Details are given in Example 3.0.1.

2. Considering the special case $t_1 = t_2 = 1$ for the real case, we have the following convergence by Corollary 4.2.1

$$\begin{pmatrix} \operatorname{tr}(\mathbf{B}_{n,1}) - p \\ \operatorname{tr}\left(\mathbf{B}_{n,1}^2\right) - p(1+y_n) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix} 0 \\ y \end{pmatrix}, \begin{pmatrix} 2y & 4y(1+y) \\ 4y(1+y) & 4y(2+5y+2y^2) \end{pmatrix} \right)$$

Using the substitution principle from Zheng et al. (2015b), this implies Proposition 2 in Ledoit and Wolf (2002) for $\alpha = 1$.

We further apply Theorem 3.2.1 for the special case $f_1(x) = x^2, f_2(x) = x^4, \mathbf{T}_n = \mathbf{I}$, which is motivated by a statistical test presented in Chapter 5.

Corollary 4.2.3 *Let $t_0 > 0$. Assume that the random variables x_{ij} satisfy condition (a) from Theorem 3.2.1 for $1 \le i \le p, 1 \le j \le n$ and $p/n \to y \in (0,\infty)$ as $n \to \infty$. Then it holds:*

1. *If the variables x_{ij} are real and $\mathbb{E}x_{ij}^4 = 3$, then the sequence*

$$\left((X_n(f_1,t))_{t\in[t_0,1]}, (X_n(f_2,t))_{t\in[t_0,1]} \right)_{n\in\mathbb{N}}$$

with

$$X_n(f_1,t) = \operatorname{tr}\left(\mathbf{B}_{n,t}^2\right) - \lfloor nt \rfloor y_n \left(\frac{\lfloor nt \rfloor}{n} + y_n \right), \ t \in [t_0,1],$$
$$X_n(f_2,t) = \operatorname{tr}\left(\mathbf{B}_{n,t}^4\right) - \lfloor nt \rfloor y_n \left\{ \left(\frac{\lfloor nt \rfloor}{n} \right)^3 + 6 \left(\frac{\lfloor nt \rfloor}{n} \right)^2 y_n + 6 \frac{\lfloor nt \rfloor}{n} y_n^2 + y_n^3 \right\},$$

converges weakly to a Gaussian process $\left((X(f_1,t))_{t\in[t_0,1]}, (X(f_2,t))_{t\in[t_0,1]} \right)$ in the space $(\ell^\infty([t_0,1]))^2$ with means

$$\mathbb{E}[X(f_1,t)] = ty, \ \mathbb{E}[X(f_2,t)] = ty(6t^2 + 17ty + 6y^2),$$

and covariance function for $t_1, t_2 \in [t_0,1]$

$$\operatorname{cov}(X(f_1,t_1), X(f_1,t_2)) = 4\min(t_1,t_2) y \left\{ 2t_1t_2 + [\min(t_1,t_2) + 2(t_1+t_2)]\, y + 2y^2 \right\},$$
$$\operatorname{cov}(X(f_1,t_1), X(f_2,t_2)) = 8\min(t_1,t_2) y \Big\{ 2t_1t_2^3 + t_2^2(3\min(t_1,t_2) + 2(6t_1+t_2))y$$
$$+ 4t_2(2\min(t_1,t_2) + 3(t_1+t_2))y^2 + (3\min(t_1,t_2) + 2(t_1+6t_2))y^3 + 2y^4 \Big\},$$

and for $t_2 = \min(t_1,t_2)$

$$\operatorname{cov}(X(f_2,t_1), X(f_2,t_2))$$
$$= 8t_2y \Big\{ 4t_1^3t_2^3 + 6t_1^2t_2^2(3t_2 + 4(t_1+t_2))y + 12t_1t_2(t_2^2 + 4t_2(t_1+t_2)$$

$$+ 2(t_1^2 + 6t_1t_2 + t_2^2))y^2 + (t_2^3 + 12t_2^2(t_1 + t_2) + 2t_2(9t_1^2 + 64t_1t_2 + 9t_2^2)$$
$$+ 4(t_1^3 + 36t_1^2t_2 + 36t_1t_2^2 + t_2^3))y^3 + 12(t_2^2 + 4t_2(t_1 + t_2)$$
$$+ 2(t_1^2 + 6t_1t_2 + t_2^2))y^4 + 6(3t_2 + 4(t_1 + t_2))y^5 + 4y^6\Big\}.$$

2. *If x_{ij} are complex with $\mathbb{E}x_{ij}^2 = 0$ and $\mathbb{E}|x_{ij}|^4 = 2$, then (1) also holds with means $\mathbb{E}[X(f,t)] = \mathbb{E}[X(g,t)] = 0$ and covariance structure given by $1/2$ times the covariance structure given for the real case.*

Proof. We omit the proof, since it is very similar to the proof of Corollary 4.2.1. □

4.3 Logarithmic law of the sequential sample covariance matrix

In the following corollary, we study the sequential process corresponding to the log-determinant of $\mathbf{B}_{n,t}$. Note that the log-determinant $\log|\mathbf{B}_{n,1}|$ of the sample covariance matrix is a well-studied object in random matrix theory (Cai et al., 2015; Wang et al., 2018) and has many applications in statistics. Other authors such as Girko (1998), Nguyen and Vu (2014) and Bao et al. (2015b) were interested in the logarithmic law of a random matrix with independent entries or of the sample correlation matrix (Parolya et al., 2021; Heiny and Parolya, 2021). The first appearance of the so-called generalized variance $\log|\mathbf{B}_{n,1}|$ in literature goes back to Frisch (1929) and Wilks (1932). A proof of the following result can be found in Section 8.2.3.

Corollary 4.3.1 *Let $t_0 \in (0,1]$, and assume that condition (a) of Theorem 3.2.1 is satisfied and that $p/n \to y \in (0, t_0)$ as $n \to \infty$.*

1. *If the variables x_{ij} are real and $\mathbb{E}x_{ij}^4 = 3$, then the process*

$$\left(\mathbb{D}_n(t)\right)_{t\in[t_0,1]} = \left(\log|\mathbf{B}_{n,t}| + p + \lfloor nt\rfloor \log(1 - y_{\lfloor nt\rfloor}) - p\log\left(\frac{\lfloor nt\rfloor}{n} - y_n\right)\right)_{t\in[t_0,1]},$$

converges weakly to a Gaussian process $(\mathbb{D}(t))_{t\in[t_0,1]}$ in the space $\ell^\infty([t_0,1])$ with mean

$$\mathbb{E}[\mathbb{D}(t)] = \frac{1}{2}\log(1 - y_t)$$

and covariance kernel

$$\operatorname{cov}(\mathbb{D}(t_1), \mathbb{D}(t_2)) = -2\log(1 - y_{t_1} \wedge y_{t_2}\ ).$$

2. *If x_{ij} are complex with $\mathbb{E}x_{ij}^2 = 0$ and $\mathbb{E}|x_{ij}|^4 = 2$, then (1) also holds with mean $\mathbb{E}[\mathbb{D}(t)] = 0$ and $\operatorname{cov}(\mathbb{D}(t_1), \mathbb{D}(t_2)) = -\log(1 - y_{t_1} \wedge y_{t_2})$.*

Chapter 5

A statistical application: Monitoring sphericity in large dimension

5.1 A change-point test for sphericity

In many statistical problems, an important assumption is sphericity, which means, that the components of a random vector are independent and have common variance. In the present context, the corresponding test problem can be formulated as

$$\mathrm{H}_0 : \mathbf{T}_n = \sigma^2 \mathbf{I}_p \text{ for some } \sigma^2 > 0, \quad \text{vs.} \quad \mathrm{H}_1 : \mathbf{T}_n \neq \sigma^2 \mathbf{I}_p \text{ for all } \sigma^2 > 0. \tag{5.1}$$

In general, it is well-known that the likelihood ratio test statistic for the hypotheses in (5.1) is degenerated if $p > n$ (see Anderson, 1984; Muirhead, 2009). A test statistic which is also applicable in the case $p \geq n$ has been proposed by John (1971) and is based on the statistic

$$\frac{1}{p} \operatorname{tr} \left\{ \left(\frac{\mathbf{B}_{n,1}}{\frac{1}{p} \operatorname{tr} \mathbf{B}_{n,1}} - \mathbf{I} \right)^2 \right\} + 1 = \frac{\frac{1}{p} \operatorname{tr}(\mathbf{B}_{n,1}^2)}{\left(\frac{1}{p} \operatorname{tr} \mathbf{B}_{n,1} \right)^2}.$$

The asymptotic properties of this statistic in the high-dimensional regime are investigated by Ledoit and Wolf (2002) and Yao et al. (2015) in the case $y \in (0, \infty)$ and by Birke and Dette (2005) in the ultra-high dimensional case $y = \infty$.

In the following discussion, we will use the results of Chapter 3 to develop a sequential monitoring procedure for the assumption of sphericity.

To be precise, we consider random variables $\mathbf{y}_1, \ldots, \mathbf{y}_n \in \mathbb{R}^p$ with real entries, where

$$\mathbf{y}_i = \mathbf{\Sigma}_i^{\frac{1}{2}} \mathbf{x}_i, \ 1 \leq i \leq n,$$

for symmetric non-negative definite matrices $\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_n \in \mathbb{R}^{p \times p}$ and random variables $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ with real entries satisfying the assumptions stated in Chapter 3. We are interested in monitoring the sphericity assumption

$$\mathrm{H}_0 : \mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_n = \sigma^2 \mathbf{I}_p \text{ for some } \sigma^2 > 0$$

$$\text{vs. } H_1 : \mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_{\lfloor nt_1^\star \rfloor} = \sigma^2 \mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt_1^\star \rfloor + 1} = \ldots = \mathbf{\Sigma}_n \neq \sigma^2 \mathbf{I}_p, \tag{5.2}$$

for some $0 < t_1^\star < 1$. For the construction of a test we consider a sequential version of the statistic proposed by John (1971), that is,

$$U_{n,t} = \frac{\frac{1}{p}\operatorname{tr}(\hat{\mathbf{\Sigma}}_{n,t}^2)}{\left(\frac{1}{p}\operatorname{tr}\hat{\mathbf{\Sigma}}_{n,t}\right)^2}\ , \tag{5.3}$$

and investigate the asymptotic behavior of the stochastic process $U_n = (U_{n,t})_{t\in[t_0,1]}$ under both the null hypothesis and the alternative. Here, $\hat{\mathbf{\Sigma}}_{n,t}$ denotes the sequential sample covariance matrix corresponding to the sample $\mathbf{y}_1, \ldots, \mathbf{y}_{\lfloor nt \rfloor}$, that is,

$$\hat{\mathbf{\Sigma}}_{n,t} = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor} \mathbf{y}_i \mathbf{y}_i^\top = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor} \mathbf{\Sigma}_i^{\frac{1}{2}} \mathbf{x}_i \mathbf{x}_i^\top \mathbf{\Sigma}_i^{\frac{1}{2}}. \tag{5.4}$$

Note that in contrast to tests based on the likelihood ratio principle the dimension may exceed the sample size. Moreover, under the null hypothesis, we have $\mathbf{\Sigma}_i = \sigma^2 \mathbf{I}_p$ ($i = 1, \ldots, n$), and a simple calculation shows that the statistic $U_{n,t}$ is independent of the concrete proportionality constant σ^2.

5.2 Convergence of the test statistic

The following theorem deals with the weak convergence of $(U_n)_{n\in\mathbb{N}}$ considered as a sequence in the space $\ell^\infty([t_0, 1])$ of bounded functions and its proof is postponed to Chapter 8. Recall that the symbol $\rightsquigarrow$ denotes weak convergence of processes and the symbol $\xrightarrow{\mathcal{D}}$ weak convergences of a real-valued random variables.

Theorem 5.2.1 *Let $y \in (0, \infty)$, $t_0 > 0$ and define $y_t = y/t$ for $t \in [t_0, 1]$. If the random variables $\mathbf{x}_1, \ldots, \mathbf{x}_n$ satisfy the assumptions (a) and (1) of Theorem 3.2.1, it follows under the null hypothesis* (5.2) *that*

$$p\left(U_{n,t} - 1 - y_{\lfloor nt \rfloor}\right)_{t\in[t_0,1]} \rightsquigarrow (U_t)_{t\in[t_0,1]} \quad \text{in } \ell^\infty([t_0, 1]),$$

as $n \to \infty$, where $(U_t)_{t\in[t_0,1]}$ denotes a Gaussian process with mean function $\mathbb{E}[U_t] = y_t$ and covariance kernel

$$\operatorname{cov}(U_{t_1}, U_{t_2}) = 4y^2_{\max(t_1,t_2)}\ , \quad t_1, t_2 \in [t_0, 1].$$

Remark 5.2.2

(1) To obtain a test for the hypotheses in (5.2) we note that the continuous mapping theorem implies under the null hypothesis

$$\sup_{t\in[t_0,1]} p\left(U_{n,t} - 1 - y_{\lfloor nt \rfloor}\right) \xrightarrow{\mathcal{D}} \sup_{t\in[t_0,1]} U_t,\ n \to \infty\ . \tag{5.5}$$

Therefore we propose to reject the null hypothesis in (5.2) whenever

$$\sup_{t\in[t_0,1]} p\left(U_{n,t} - 1 - y_{\lfloor nt\rfloor}\right) > c_\alpha, \tag{5.6}$$

where c_α denotes the $(1-\alpha)$-quantile of the statistic $\sup_{t\in[t_0,1]} U_t$. Thus, we have by (5.5)

$$\lim_{n\to\infty} \mathbb{P}_{H_0}\Big(\sup_{t\in[t_0,1]} p\big(U_{n,t} - 1 - y_{\lfloor nt\rfloor}\big) > c_\alpha\Big) = \mathbb{P}\Big(\sup_{t\in[t_0,1]} U_t > c_\alpha\Big) \le \alpha,$$

which means, that the test keeps a nominal level α (asymptotically).

(2) In order to investigate the consistency of the test (5.6) assume that the matrices $\mathbf{\Sigma}_i$ in (5.2) satisfy

$$\mathbf{\Sigma}_i = \begin{cases} \sigma^2 \mathbf{I}_p & \text{if } \ 0 \le i \le \lfloor nt_1^\star\rfloor, \\ \mathbf{\Sigma} & \text{if } \ \lfloor nt_1^\star\rfloor < i \le n, \end{cases}$$

where $\sigma^2 > 0$ and $\mathbf{\Sigma}$ is a $p\times p$ nonnegative definite matrix. We also assume that $\frac{1}{p}\mathrm{tr}\mathbf{\Sigma}$ and $\frac{1}{p}\mathrm{tr}(\mathbf{\Sigma}^2)$ converge to $g > 0$ and $h > 0$, respectively. Furthermore, for the matrix $\mathbf{H} = \mathbf{\Sigma}^{\frac{1}{2}} = (H_{ij})_{i,j=1,\dots,p}$ we have

$$\Big(\frac{1}{p}\sum_{j,l=1}^{p} H_{jl}^2\Big)^2 = \Big(\frac{1}{p}\mathrm{tr}\mathbf{\Sigma}\Big)^2 \to g^2.$$

A straightforward calculation then shows that for $t \in (t_1^\star, 1)$

$$\begin{aligned}
\frac{1}{p}\mathbb{E}\left[\mathrm{tr}\left(\hat{\mathbf{\Sigma}}_{n,t}\right)\right] &\xrightarrow{\mathbb{P}} t_1^\star\sigma^2 + (t - t_1^\star)g, \\
\frac{1}{p}\mathbb{E}\left[\mathrm{tr}\left(\hat{\mathbf{\Sigma}}_{n,t}^2\right)\right] &\xrightarrow{\mathbb{P}} (t_1^\star)^2\sigma^4 + 2t_1^\star\sigma^2(t - t_1^\star)g + (t - t_1^\star)^2 h + y t_1^\star\sigma^4 + y(t - t_1^\star)g.
\end{aligned}$$

Using a martingale decomposition and the estimate (9.9.3) in Bai and Silverstein (2010), one can show that for fixed $t \in (t_1^\star, 1)$

$$\mathbb{E}|s_{F^{\hat{\mathbf{\Sigma}}_{n,t}}}(z) - \mathbb{E}[s_{F^{\hat{\mathbf{\Sigma}}_{n,t}}}(z)]|^2 \to 0,$$

if we assume that the spectral norm $||\mathbf{\Sigma}||$ is uniformly bounded with respect to $n \in \mathbb{N}$. Using (6.1), this implies

$$\frac{1}{p}\,\mathrm{tr}\left(f(\hat{\mathbf{\Sigma}}_{n,t})\right) - \frac{1}{p}\mathbb{E}\left[\mathrm{tr}\left(f(\hat{\mathbf{\Sigma}}_{n,t})\right)\right] \xrightarrow{\mathbb{P}} 0$$

for $f(x) = x$ and $f(x) = x^2$. Consequently,

$$\begin{aligned}
U_{n,t} &\xrightarrow{\mathbb{P}} \frac{(t_1^\star)^2\sigma^4 + 2t_1^\star\sigma^2(t - t_1^\star)g + (t - t_1^\star)^2 h + y t_1^\star\sigma^4 + y(t - t_1^\star)g^2}{(t_1^\star)^2\sigma^4 + \big((t - t_1^\star)g\big)^2 + 2t_1^\star\sigma^2(t - t_1^\star)g} \\
&= 1 + y_t + \Delta_{1,t} + \Delta_{2,t}
\end{aligned}$$

where

$$\Delta_{1,t} = \frac{(t-t_1^\star)^2(h-g^2)}{(t_1^\star)^2\sigma^4 + \left((t-t_1^\star)g\right)^2 + 2t_1^\star\sigma^2(t-t_1^\star)g} \geq 0$$

by construction, and

$$\begin{aligned}
\Delta_{2,t} =& \frac{yt_1^\star\sigma^4 + y(t-t_1^\star)g^2}{(t_1^\star)^2\sigma^4 + \left((t-t_1^\star)g\right)^2 + 2t_1^\star\sigma^2(t-t_1^\star)g} - y_t \\
=& \frac{yt_1^\star\sigma^4 + y(t-t_1^\star)g^2 - y_t\left\{(t_1^\star)^2\sigma^4 + \left((t-t_1^\star)g\right)^2 + 2t_1^\star\sigma^2(t-t_1^\star)g\right\}}{(t_1^\star)^2\sigma^4 + \left((t-t_1^\star)g\right)^2 + 2t_1^\star\sigma^2(t-t_1^\star)g} \\
=& \frac{y_t t_1^\star(t-t_1^\star)\left(\sigma^2-g\right)^2}{(t_1^\star)^2\sigma^4 + \left((t-t_1^\star)g\right)^2 + 2t_1^\star\sigma^2(t-t_1^\star)g} \geq 0.
\end{aligned}$$

Note that under the alternative in (5.2) two types of structural breaks in the covariance structure corresponding to the terms $\Delta_{1,t}$ and $\Delta_{2,t}$ may occur. On the one hand, the diagonal elements in the matrices $\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_n$ might shift from σ^2 to a different variance while the matrices still remain spherical. This structural break is captured by the term $\Delta_{2,t}$. On the other hand, the change in the matrices could violate the sphericity assumption, which corresponds to the term $\Delta_{1,t}$.
Consequently, whenever there exists a parameter $\tilde{t} \in (t_1^\star, 1)$ such that $\Delta_{1,\tilde{t}} > 0$ or $\Delta_{2,\tilde{t}} > 0$, it follows under the additional assumption $y - y_n = o\left(p^{-1}\right)$ that

$$\sup_{t\in[t_0,1]} p(U_{n,t} - 1 - y_{\lfloor nt \rfloor}) \geq p(U_{n,\tilde{t}} - 1 - y_{\lfloor n\tilde{t} \rfloor}) \xrightarrow{\mathbb{P}} \infty,$$

and in this case the test (5.6) rejects the null hypothesis with a probability converging to 1 as $p, n \to \infty$, $p/n \to y \in (0, \infty)$. This is in particular the case for the alternative considered in (5.2).

Fisher et al. (2010) consider several extensions of the classical test introduced by John (1971). Motivated by this work, an alternative test for the hypothesis (5.2) could be based on the test statistic

$$U_{n,t}^{(2)} = \frac{\frac{1}{p}\operatorname{tr}(\hat{\mathbf{\Sigma}}_{n,t}^4)}{\left(\frac{1}{p}\operatorname{tr}\hat{\mathbf{\Sigma}}_{n,t}^2\right)^2},$$

where the matrix $\hat{\mathbf{\Sigma}}_{n,t}$ is defined in (5.4). For $t = 1$, the asymptotic properties of an appropriately centered version of $U_{n,1}^{(2)}$ have been investigated by Fisher et al. (2010) assuming that all arithmetic means of the eigenvalues of the sample covariance up to order 16 converge to the corresponding arithmetic means of the eigenvalues of the population covariance. The following result provides the weak convergence of the corresponding stochastic process $U_n^{(2)} = (U_{n,t}^{(2)})_{t\in[t_0,1]}$ under the null hypothesis. A corresponding asymptotic level-α test and a discussion of its power properties can be obtained by similar arguments as given for the process $\left(U_{n,t}^{(1)}\right)_{t\in[t_0,1]}$ in Remark 5.2.2 and the details are omitted for the sake of brevity.

Theorem 5.2.3 *Under the assumptions of Theorem 5.2.1 we have*

$$p\Big(U_{n,t}^{(2)} - \frac{1+6y_{\lfloor nt\rfloor}+6y_{\lfloor nt\rfloor}^2+y_{\lfloor nt\rfloor}^3}{(1+y_{\lfloor nt\rfloor})^2}\Big)_{t\in[t_0,1]} \rightsquigarrow (U_t^{(2)})_{t\in[t_0,1]} \quad \text{in } \ell^\infty([t_0,1]),$$

where $(U_t^{(2)})_{t\in[t_0,1]}$ *denotes a Gaussian process with mean function*

$$\mathbb{E}[U_t^{(2)}] = \frac{y(4t^2+7ty+4y^2)}{t(t+y)^2}\ , \quad t\in[t_0,1],$$

and covariance kernel

$$\begin{aligned}&\operatorname{cov}(U_{t_1}^{(2)},U_{t_2}^{(2)})\\ =&\frac{8y^2\Big\{4t_1^2(2t_2^2+3t_2y+2y^2)+6t_1y(4t_2^2+5t_2y+2y^2)+y^2(21t_2^2+24t_2y+8y^2)\Big\}}{t_1^2(t_1+y)^2(t_2+y)^2}\end{aligned}$$

for $t_0 \le t_2 \le t_1 \le 1$.

Proof. Using Corollary 4.2.3 combined with the functional delta method, the proof of Theorem 5.2.3 is very similar to the proof of Theorem 5.2.1 in Chapter 8 and therefore omitted. □

Example 5.2.4 We conclude this section with a small simulation study illustrating the finite-sample properties of the test (5.6). For this purpose, we generated centered p-dimensional normally distributed data with various covariance structures. To be precise, we consider the the alternatives

$$\mathbf{\Sigma}_1=\ldots=\mathbf{\Sigma}_{\lfloor nt^\star\rfloor}=\mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star\rfloor+1}=\ldots=\mathbf{\Sigma}_n=\mathbf{I}_p+\operatorname{diag}(\underbrace{0,\ldots,0}_{p/2},\underbrace{\delta,\ldots,\delta}_{p/2}), \tag{5.7}$$

$$\mathbf{\Sigma}_1=\ldots=\mathbf{\Sigma}_{\lfloor nt^\star\rfloor}=\mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star\rfloor+1}=\ldots=\mathbf{\Sigma}_n=\mathbf{I}_p+\operatorname{diag}(\underbrace{0,\ldots,0}_{p/2},\underbrace{\delta,\ldots,\delta}_{p/2})+\tilde{\mathbf{S}}(\delta), \tag{5.8}$$

$$\mathbf{\Sigma}_1=\ldots=\mathbf{\Sigma}_{\lfloor nt^\star\rfloor}=\mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star\rfloor+1}=\ldots=\mathbf{\Sigma}_n=(1+\varepsilon)\mathbf{I}_p, \tag{5.9}$$

$$\mathbf{\Sigma}_1=\ldots=\mathbf{\Sigma}_{\lfloor nt^\star\rfloor}=\mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star\rfloor+1}=\ldots=\mathbf{\Sigma}_n=(1+\varepsilon)\mathbf{I}_p+\mathbf{S}(\varepsilon), \tag{5.10}$$

where $\delta,\varepsilon\ge 0$ determine the "deviation" from the null hypothesis (note that the choice $\delta=0$ and $\varepsilon=0$ correspond to the null hypothesis (5.2)). Here, the entries of the $p\times p$ matrix $\mathbf{S}(\varepsilon)$ in (5.10) are given by $S_{j,j-1}(\varepsilon)=S_{j-1,j}(\varepsilon)=\varepsilon,\ 1\le j\le p$, and all other entries are 0. Similarly, the $p\times p$ matrix $\tilde{\mathbf{S}}(\delta)$ in (5.8) has the entries $\tilde{S}_{j,j-1}(\delta)=\tilde{S}_{j-1,j}(\delta)=\delta,\ p/2<j\le p$, and all other entries are 0.

In Figure 5.1 and Figure 5.2, we display the empirical rejection of the test (5.6) for the different alternatives and different values of n and p, where the change point is given by $t^\star=0.6$. For the parameter t_0, we always use $t_0=0.2$, and all results are based on $2,000$ simulation runs. The vertical gray line in each figure defines the nominal level $\alpha=5\%$.

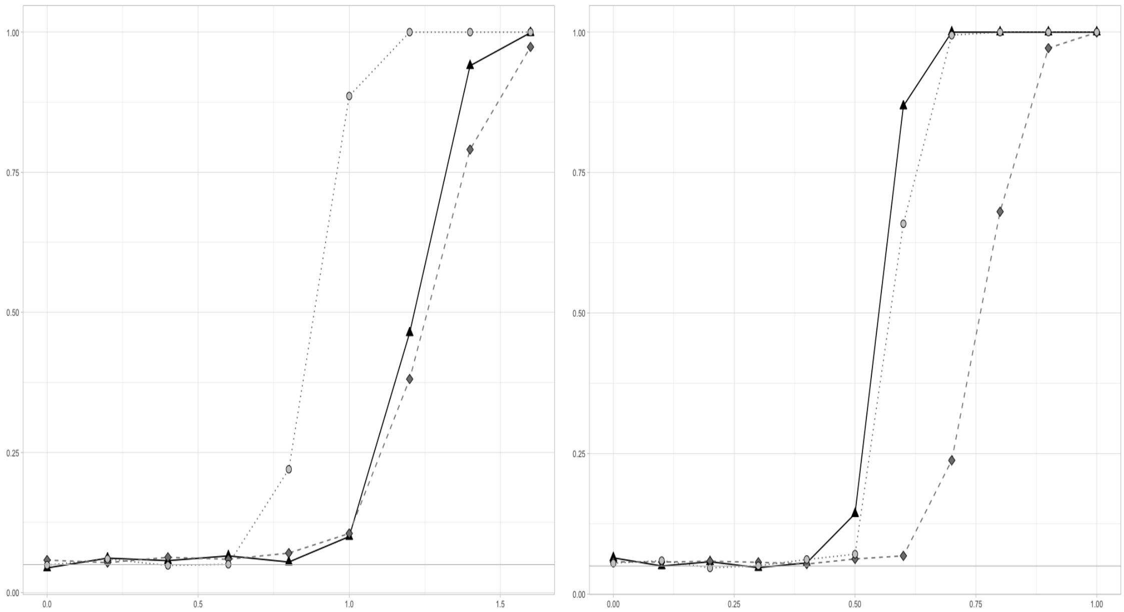

Figure 5.1: *Simulated rejection probabilities of the test* (5.6) *under the null hypothesis (*$\delta = 0$*) and the different alternatives in* (5.7) *(left) and* (5.8) *(right) for* $\delta > 0$ *. The circle indicates* $n = 200, p = 300$*, the triangle* $n = 200, p = 120$ *and the square* $n = 150, p = 300$*.*

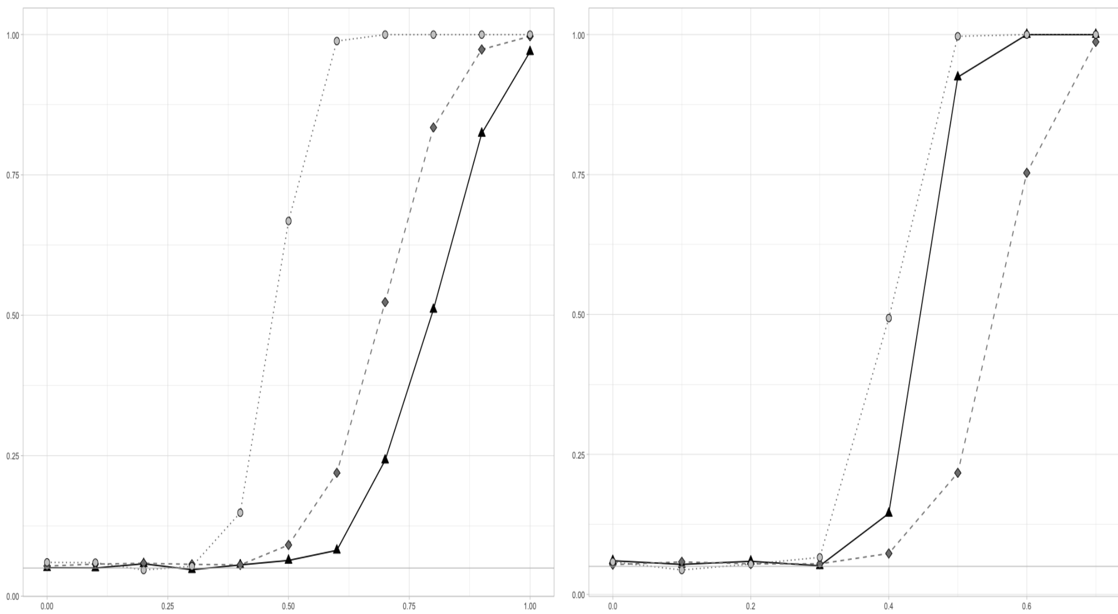

Figure 5.2: *Simulated rejection probabilities of the test* (5.6) *under the null hypothesis (*$\varepsilon = 0$*) and the different alternatives in* (5.9) *(left) and* (5.10) *(right) for* $\varepsilon > 0$*. The circle indicates* $n = 200, p = 300$*, the triangle* $n = 200, p = 120$ *and the square* $n = 150, p = 300$*.*

Note that the choices $\delta = 0$ and $\varepsilon = 0$ correspond to the null hypothesis in

(5.7), (5.8), (5.9) and (5.10), respectively. We observe a good approximation of the nominal level in all cases under consideration. Moreover, the test has power under all considered alternatives, even if the dimension p is substantially larger than the sample size. Note that the test performs better for alternatives of the form (5.8) compared to the alternatives in (5.7). This reflects the intuition that the alternative in (5.7) is somehow closer to sphericity than the alternative (5.8). A similar observation can be made for the alternatives (5.9) and (5.10).

Chapter 6

Proof of Theorem 3.2.1

6.1 Outline of the proof of Theorem 3.2.1

A frequently used powerful tool in random matrix theory is the Stieltjes transform. This is partially explained by the formula

$$\int f(x)dG(x) = \frac{1}{2\pi i}\int\int_{\mathcal{C}} \frac{f(z)}{z-x} dz dG(x) = -\frac{1}{2\pi i}\int_{\mathcal{C}} f(z) s_G(z) dz, \tag{6.1}$$

where G is an arbitrary cumulative distribution function (c.d.f.) with a compact support, f is an arbitrary analytic function on an open set, say O, containing the support of G, $\mathcal{C}$ is a positively oriented contour in O enclosing the support of G and

$$s_G(z) = \int \frac{1}{x-z} dG(x)$$

denotes the Stieltjes transform of G. Note that (6.1) follows from Cauchy's integral formula (see, e.g., Ahlfors, 1953) and Fubini's theorem. Thus invoking the continuous mapping theorem, it may suffice to prove weak convergence for the sequence $(M_n)_{n\in\mathbb{N}}$, where

$$M_n(z,t) = p\left(s_{F^{\mathbf{B}_{n,t}}}(z) - s_{\tilde{F}^{y_{\lfloor nt\rfloor},H_n}}(z)\right), \quad z \in \mathcal{C}. \tag{6.2}$$

Here, $s_{\tilde{F}^{y_{\lfloor nt\rfloor},H_n}}$ denotes the Stieltjes transform of $\tilde{F}^{y_{\lfloor nt\rfloor},H_n}$ given in (3.6) characterized through the equation

$$s_{\tilde{F}^{y_{\lfloor nt\rfloor},H_n}}(z) = \int \frac{1}{\lambda\frac{\lfloor nt\rfloor}{n}\left(1 - y_{\lfloor nt\rfloor} - y_{\lfloor nt\rfloor} z s_{\tilde{F}^{y_{\lfloor nt\rfloor},H_n}}(z)\right) - z} dH_n(\lambda), \tag{6.3}$$

and the contour $\mathcal{C}$ in (6.2) has to be constructed in such a way that it encloses the support of $\tilde{F}^{y_{\lfloor nt\rfloor},H_n}$ and $F^{\mathbf{B_{n,t}}}$ with probability 1 for sufficiently large $n \in \mathbb{N}, t \in [t_0, 1]$. This idea is formalized in the proof of Theorem 3.2.1 in Section 6.2.
In order to prove the weak convergence of (6.2), define a contour $\mathcal{C}$ as follows. Let x_r be any number greater than the right endpoint of the interval (3.7) and $v_0 > 0$

be arbitrary. Let x_l be any negative number if the left endpoint of the interval (3.7) is zero. Otherwise, choose

$$x_l \in \left(0, \liminf_{n\to\infty} \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{t_0}) t_0 (1-\sqrt{y_{t_0}})^2\right).$$

Let $\mathcal{C}_u = \{x + iv_0 : x \in [x_l, x_r]\}$,

$$\mathcal{C}^+ = \{x_l + iv : v \in [0, v_0]\} \cup \mathcal{C}_u \cup \{x_r + iv : v \in [0, v_0]\},$$

and define $\mathcal{C} = \mathcal{C}^+ \cup \overline{\mathcal{C}^+}$, where $\overline{\mathcal{C}^+}$ contains all elements of $\mathcal{C}^+$ complex conjugated. Next, consider a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ converging to zero such that for some $\alpha \in (0,1)$

$$\varepsilon_n \geq n^{-\alpha},$$

define

$$\mathcal{C}_l = \{x_l + iv : v \in [n^{-1}\varepsilon_n, v_0]\}$$
$$\mathcal{C}_r = \{x_r + iv : v \in [n^{-1}\varepsilon_n, v_0]\},$$

and consider the set $\mathcal{C}_n = \mathcal{C}_l \cup \mathcal{C}_u \cup \mathcal{C}_r$. We define an approximation $\hat{M}_n$ of the process M_n for $z = x + iv \in \mathcal{C}^+, t \in [t_0, 1]$ by

$$\hat{M}_n(z,t) = \begin{cases} M_n(z,t) & \text{if } z \in \mathcal{C}_n, \\ M_n(x_r + in^{-1}\varepsilon_n, t) & \text{if } x = x_r,\ v \in [0, n^{-1}\varepsilon_n], \\ M_n(x_l + in^{-1}\varepsilon_n, t) & \text{if } x = x_l,\ v \in [0, n^{-1}\varepsilon_n]. \end{cases} \tag{6.4}$$

In Lemma 6.4.3 in Section 6.4, it is shown that $(\hat{M}_n)_{n\in\mathbb{N}}$ approximates $(M_n)_{n\in\mathbb{N}}$ appropriately in the sense that the corresponding linear spectral statistics

$$-\frac{1}{2\pi i}\int_{\mathcal{C}} f(z) M_n(z,t)dz \quad \text{and} \quad -\frac{1}{2\pi i}\int_{\mathcal{C}} f(z)\hat{M}_n(z,t)dz$$

in (6.1) coincide asymptotically. As a consequence, the weak convergence of the process (6.2) follows from that of $\hat{M}_n$, which is established in the following theorem. The proof is given in Section 6.3.

Theorem 6.1.1 (Weak convergence for the process of Stieltjes transforms) *Under the assumptions of Theorem 3.2.1, the sequence $(\hat{M}_n)_{n\in\mathbb{N}}$ defined in (6.4) converges weakly to a Gaussian process $(M(z,t))_{z\in\mathcal{C}^+, t\in[t_0,1]}$ in the space $\ell^\infty(\mathcal{C}^+ \times [t_0,1])$.*

The mean of the limiting process M is given by

$$\mathbb{E}M(z,t) = \begin{cases} \dfrac{ty\int \frac{\underline{\tilde{s}}_t^3(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^3}dH(\lambda)}{\left(1 - ty\int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^2}dH(\lambda)\right)^2} & \text{for the real case,} \\ 0 & \text{for the complex case,} \end{cases} \tag{6.5}$$

where $z \in \mathcal{C}^+$, $t \in [t_0,1]$. In the complex case, the covariance kernel of the limiting process M is given by

$$\begin{aligned}\operatorname{cov}(M(z_1,t_1), M(z_2,t_2)) &= \mathbb{E}\left[(M(z_1,t_1) - \mathbb{E}[M(z_1,t_1)])\,\overline{(M(z_2,t_2) - \mathbb{E}[M(z_2,t_2)])}\right] \\ &= \sigma^2_{t_1,t_2}(z_1,\overline{z_2}),\ t_1,t_2 \in [t_0,1],\ z_1,z_2 \in \mathcal{C}^+,\end{aligned}$$

where $\sigma^2_{t_1,t_2}(z_1,z_2)$ is defined in (6.44). In the real case, we have

$$\operatorname{cov}(M(z_1,t_1), M(z_2,t_2)) = 2\sigma^2_{t_1,t_2}(z_1,\overline{z_2}). \tag{6.6}$$

6.2 Proof of Theorem 3.2.1 using Theorem 6.1.1

From (6.1) we obtain

$$-\frac{1}{2\pi i}\int_{\mathcal{C}} f(z)\mathbb{E}s_G(z)dz = -\frac{1}{2\pi i}\mathbb{E}\int_{\mathcal{C}} f(z)s_G(z)dz = \mathbb{E}\int f(x)dG(x). \tag{6.7}$$

We choose v_0, x_r, x_l so that f_1 and f_2 given in Theorem 3.2.1 are analytic on and inside the resulting contour $\mathcal{C}$ and define

$$\mathbf{S}_{n,t} = \frac{1}{n}\mathbf{X}_{n,t}\mathbf{X}_{n,t}^{\star}.$$

The almost sure convergence

$$\lim_{n\to\infty}\lambda_p(\mathbf{S}_{n,t}) = t(1-\sqrt{y_t})^2 I_{(0,1)}(y_t) = (\sqrt{t}-\sqrt{y})^2 I_{(0,1)}(y_t),$$
$$\lim_{n\to\infty}\lambda_1(\mathbf{S}_{n,t}) = t(1+\sqrt{y_t})^2 = (\sqrt{t}+\sqrt{y})^2$$

of the extreme eigenvalues (see, e.g., Theorem 1.1 in Bai and Zhou, 2008) and the inequalities

$$\lambda_1(\mathbf{AB}) \le \lambda_1(\mathbf{A})\lambda_1(\mathbf{B}),\ \lambda_p(\mathbf{AB}) \ge \lambda_p(\mathbf{A})\lambda_p(\mathbf{B})$$

(valid for quadratic Hermitian nonnegative definite matrices $\mathbf{A}$ and $\mathbf{B}$) imply

$$\begin{aligned}\limsup_{n\to\infty}\lambda_1(\mathbf{B}_{n,t}) &\le \limsup_{n\to\infty}\lambda_1(\mathbf{T}_n)\cdot\limsup_{n\to\infty}\lambda_1(\mathbf{S}_{n,t}) \ = \limsup_{n\to\infty}\lambda_1(\mathbf{T}_n)t\left(1+\sqrt{y_t}\right)^2\\ &\le \limsup_{n\to\infty}\lambda_1(\mathbf{T}_n)\left(1+\sqrt{y_{t_0}}\right)^2 < x_r\end{aligned}$$

for each $t\in[t_0,1]$ with probability 1. Similar calculations for x_l show that it holds for all $t\in[t_0,1]$ with probability 1

$$\liminf_{n\to\infty}\min\left(x_r-\lambda_1(\mathbf{B}_{n,t}),\lambda_p(\mathbf{B}_{n,t})-x_l\right) > 0, \tag{6.8}$$

which implies that for sufficiently large n the contour $\mathcal{C}$ encloses the support of $F^{\mathbf{B}_{n,t}}$, $t\in[t_0,1]$, with probability 1 (note that the null set depends on n and t). For every n, there exist only finitely many $t_1,t_2\in[t_0,1]$ such that $\lfloor nt_1\rfloor \ne \lfloor nt_2\rfloor$. Since the countable union of null sets is again a null set, we may choose the above nullset in such a way that $\mathcal{C}$ encloses the support of $F^{\mathbf{B}_{n,t}}$ for sufficiently large n with probability 1 (this null set is independent of n and $t\in[t_0,1]$). From Lemma 6.4.1 in Section 6.4, it follows that the support of $\tilde{F}^{y_{\lfloor nt\rfloor},H_n}$, $t\in[t_0,1]$, is contained in the interval

$$\left[\frac{\lfloor nt_0\rfloor}{n}\lambda_p(\mathbf{T}_n)I_{(0,1)}(y_{\lfloor nt_0\rfloor})(1-\sqrt{y_{\lfloor nt_0\rfloor}})^2, \lambda_1(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt_0\rfloor}})^2\right],$$

which is enclosed by the contour $\mathcal{C}$ for sufficiently large n. Therefore, using (6.1) and (6.7), we have almost surely

$$\left(\left(-\frac{1}{2\pi i}\int_{\mathcal{C}} f_j(z)M_n(z,t)dz\right)_{j=1,2}\right)_{t\in[t_0,1]} = \left((X_n(f_j,t))_{j=1,2}\right)_{t\in[t_0,1]}$$

for sufficiently large n. Moreover, we have with probability 1 (see Lemma 6.4.3 in Section 6.4)

$$\Big| \int_{\mathcal{C}} f_j(z)(M_n(z,t) - \hat{M}_n(z,t))dz \Big| = o(1), \ j = 1, 2,$$

uniformly with respect to $t \in [t_0, 1]$. Let $C(\mathcal{C} \times [t_0, 1])$ and $C([t_0, 1])$ denote the spaces of continuous functions defined on $\mathcal{C} \times [t_0, 1]$ and $[t_0, 1]$, respectively, then the mapping

$$C(\mathcal{C} \times [t_0, 1]) \to (C([t_0, 1]))^2, \ h \mapsto (I_{f_1}(h), I_{f_2}(h))$$

is continuous, where

$$I_{f_j}(h)(\cdot) = -\frac{1}{2\pi i} \int_{\mathcal{C}} f_j(z) h(z, \cdot) dz \in C([t_0, 1]), j = 1, 2.$$

By Corollary 6.3.5 stated in Section 6.3.3 below and (6.5), the limiting process M in Theorem 6.1.1 satisfies $M \in C(\mathcal{C}^+ \times [t_0, 1])$. Invoking the continuous mapping theorem (see Theorem 1.3.6 in Van Der Vaart and Wellner, 1996) and noting that $\overline{M_n(z,t)} = M_n(\overline{z}, t)$, we have

$$\big(I_{f_1}(\hat{M}_n), I_{f_2}(\hat{M}_n)\big) \rightsquigarrow (I_{f_1}(M), I_{f_2}(M)) = \Big(\Big(-\frac{1}{2\pi i} \int_{\mathcal{C}} f_j(z) M(z,t) dz\Big)_{j=1,2}\Big)_{t \in [t_0, 1]}.$$

The fact that this random variable is a Gaussian process follows from the observation that the Riemann sums corresponding to these integrals are multivariate Gaussian and therefore the integral must be Gaussian as well. The limiting expression for the mean and the covariance follow immediately from Theorem 6.1.1. For example, we have for the real case observing (6.6)

$$\begin{aligned}
&\operatorname{cov}\Big(-\frac{1}{2\pi i} \int_{\mathcal{C}} f_1(z) M(z, t_1) dz, -\frac{1}{2\pi i} \int_{\mathcal{C}} f_2(z) M(z, t_2) dz\Big) \\
=&\frac{1}{4\pi^2} \int_{\mathcal{C}_1} \int_{\mathcal{C}_2} f_1(z_1) \overline{f_2(z_2)} \operatorname{cov}\left(M(z_1, t_1), M(z_2, t_2)\right) \overline{dz_2} dz_1 \\
=&\frac{1}{2\pi^2} \int_{\mathcal{C}_1} \int_{\mathcal{C}_2} f_1(z_1) \overline{f_2(z_2)} \sigma^2_{t_1, t_2}(z_1, \overline{z_2}) \overline{dz_2} dz_1.
\end{aligned}$$

6.3 Proof of Theorem 6.1.1

We begin with the usual "truncation" and replace the entries of the matrix $\mathbf{X}_n = (x_{ij})_{i=1,\dots,p, j=1,\dots,n}$ by truncated variables [see Section 9.7.1, Bai and Silverstein (2010)]. More precisely, without loss of generality we assume that

$$|x_{ij}| < \eta_n \sqrt{n}, \ \mathbb{E}[x_{ij}] = 0, \ \mathbb{E}|x_{ij}|^2 = 1, \ \mathbb{E}|x_{ij}|^4 < \infty.$$

Additionally, for the real case (part (1) of Theorem 3.2.1) we may assume that

$$\mathbb{E}|x_{ij}|^4 = 3 + o(1)$$

uniformly in $i \in \{1,\dots,p\}, j \in \{1,\dots,n\}$, and for the complex case (part (2) of Theorem 3.2.1)

$$\mathbb{E}x_{ij}^2 = o\Big(\frac{1}{n}\Big),\ \mathbb{E}|x_{ij}|^4 = 2 + o(1)$$

uniformly in $i \in \{1,\dots,p\}, j \in \{1,\dots,n\}$. Here, $(\eta_n)_{n\in\mathbb{N}}$ denotes a sequence converging to zero with the property

$$\eta_n n^{1/5} \to \infty.$$

We now give a brief outline for the proof of Theorem 6.1.1 describing the important steps, which are carried out in the following sections and chapters. We consider the stochastic processes $(M_n)_{n\in\mathbb{N}}$ and $(\hat{M}_n)_{n\in\mathbb{N}}$ (which is defined in (6.4)) as sequences in the space $\ell^\infty(\mathcal{C}^+ \times [t_0, 1])$ and use the decomposition

$$M_n = M_n^1 + M_n^2 \ , \tag{6.9}$$

where the random part M_n^1 and the deterministic part M_n^2 are given by

$$M_n^1(z,t) = p\left(s_{F^{\mathbf{B}_{n,t}}}(z) - \mathbb{E}\left[s_{F^{\mathbf{B}_{n,t}}}(z)\right]\right), \tag{6.10}$$

$$M_n^2(z,t) = p\left(\mathbb{E}\left[s_{F^{\mathbf{B}_{n,t}}}(z)\right] - s_{\tilde{F}^{y_{\lfloor nt\rfloor}, H_n}}(z)\right), \tag{6.11}$$

the Stieltjes transform $s_{\tilde{F}^{y_{\lfloor nt\rfloor}, H_n}}$ is defined in (6.3) and $s_{F^{\mathbf{B}_{n,t}}}$ denotes the Stieltjes transform of the empirical spectral distribution $F^{\mathbf{B}_{n,t}}$.
Our first result provides the convergence of the finite-dimensional distributions of $(M_n^1)_{n\in\mathbb{N}}$. Its proof relies on a central limit theorem for martingale difference schemes and is carried out in Section 6.3.2.

Theorem 6.3.1 *Under the assumption (1) for the real case or assumption (2) for the complex case from Theorem 3.2.1, it holds for all* $k \in \mathbb{N}, t_1, t_2 \in [0,1]$, $z_1, ..., z_k \in \mathbb{C}$, $\operatorname{Im}(z_i) \neq 0$

$$\begin{aligned}&(M_n^1(z_1,t_1), M_n^1(z_1,t_2), ..., M_n^1(z_k,t_1), M_n^1(z_k,t_2))^\top \\ &\xrightarrow{\mathcal{D}} (M^1(z_1,t_1), M^1(z_1,t_2), ..., M^1(z_k,t_1), M^1(z_k,t_2))^\top \ ,\end{aligned} \tag{6.12}$$

where $M^1(z,t) = M(z,t) - \mathbb{E}[M(z,t)]$ *is the centered version of the Gaussian process defined in Theorem 6.1.1.*

Next, we define the process $\hat{M}_n^1$ in the same way as $\hat{M}_n$ in (6.4) replacing M_n by M_n^1 and show the following tightness result. The main argument in its proof consists of establishing delicate moment inequalities for the increments of the process $(\hat{M}_n^1)_{n\in\mathbb{N}}$, see Lemma 6.3.4 and its proof in Section 7.2.

Theorem 6.3.2 *Under the assumptions of Theorem 3.2.1, the sequence* $(\hat{M}_n^1)_{n\in\mathbb{N}}$ *is asymptotically tight in the space* $\ell^\infty(\mathcal{C}^+ \times [t_0, 1])$.

The third step is an investigation of the deterministic part. In particular, we show that the bias $(M_n^2)_{n\in\mathbb{N}}$ converges in the space $\ell^\infty(\mathcal{C}^+ \times [t_0, 1])$ to the limit given in (6.5). Note that the space of bounded function is equipped with the sup-norm, which demands a uniform convergence of the Stieltjes transform $\mathbb{E}[s_{F^{\mathbf{B}_{n,t}}}(z)]$ with respect to the arguments $t \in [t_0, 1], z \in \mathcal{C}^+$. The latter result is provided in Theorem 6.3.7 in Section 6.3.4.

Theorem 6.3.3 *Under the assumptions of Theorem 3.2.1, it holds*

$$\lim_{n\to\infty} \sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} \left| M_n^2(z,t) - \mathbb{E}[M(z,t)] \right| = 0.$$

The proofs of Theorem 6.3.1, 6.3.2 and 6.3.3 are postponed to Section 6.3.2, 6.3.3 and 6.3.4, respectively. Using these results, we are now in the position to prove Theorem 6.1.1.

6.3.1 Proof of Theorem 6.1.1

Theorem 6.3.1 yields the convergence of the finite-dimensional distributions of $M_n^1(z,t)$ for $t \in [t_0, 1]$ and $z \in \mathcal{C}$ with $\mathrm{Im}(z) \neq 0$ towards the corresponding finite-dimensional distributions of the centered process $M^1(z,t) = M(z,t) - \mathbb{E}[M(z,t)]$. By the definition in equation (6.4), this implies the convergence of the finite-dimensional distributions of $\hat{M}_n^1(z,t)$ for $t \in [t_0, 1]$ and $z \in \mathcal{C}$ with $\mathrm{Im}(z) \neq 0$ towards the corresponding finite-dimensional distributions of M^1. Since the limiting process $(M^1(z,t))_{z\in\mathcal{C}^+, t\in[t_0,1]}$ is continuous as proven later in this chapter (see Corollary 6.3.5 in Section 6.3.3) and $(\mathcal{C}^+ \setminus \{x_l, x_r\}) \times [t_0, 1]$ is a dense subset of $\mathcal{C}^+ \times [t_0, 1]$, this is sufficient in order to ensure uniqueness of the limiting process. As Theorem 6.3.2 establishes asymptotic tightness, Theorem 6.1.1 follows from the decomposition (6.9), Theorem 1.5.6 in Van Der Vaart and Wellner (1996) and Theorem 6.3.3.

6.3.2 Proof of Theorem 6.3.1

The proof is divided in several steps and demands some auxiliary results, which can be found in Section 7.1. We start by performing some preparations and by introducing notations which will remain crucial for the rest of this work.

Step 0: *Preliminaries and notations*
The convergence in (6.12) is implied by the weak convergence

$$\sum_{i=1}^{k} \left(\alpha_{i,1} M_n^1(z_i, t_1) + \alpha_{i,2}(M_n^1(z_i, t_2)\right) \xrightarrow{\mathcal{D}} \sum_{i=1}^{k} \left(\alpha_{i,1} M^1(z_i, t_1) + \alpha_{i,2}(M^1(z_i, t_2)\right) \tag{6.13}$$

for all $\alpha_{1,1}, \ldots, \alpha_{k,1}, \alpha_{1,2}, \ldots, \alpha_{k,2} \in \mathbb{C}$. We want to show that the limiting random variable on the right-hand side of the display above follows a Gaussian distribution

under the assumption (1) or (2) of Theorem 3.2.1.
Recalling assumption (b) in Theorem 3.2.1, we may assume $||\mathbf{T}_n|| \leq 1$ for convenience ($n \in \mathbb{N}$). Define for $k, j = 1, .., \lfloor nt \rfloor$, $k \neq j$, $t \in (0, 1]$, $z \in \mathbb{C}$ with $\mathrm{Im}(z) \neq 0$

$$\begin{aligned}
\mathbf{r}_j &= \frac{1}{\sqrt{n}} \mathbf{T}_n^{\frac{1}{2}} \mathbf{x}_j \\
\mathbf{B}_{n,t} &= \sum_{j=1}^{\lfloor nt \rfloor} \mathbf{r}_j \mathbf{r}_j^\star, \\
\mathbf{D}_t(z) &= \mathbf{B}_{n,t} - z\mathbf{I}, \\
\mathbf{D}_{j,t}(z) &= \mathbf{D}_t(z) - \mathbf{r}_j \mathbf{r}_j^\star, \\
\mathbf{D}_{k,j,t}(z) &= \mathbf{D}_{j,t}(z) - \mathbf{r}_k \mathbf{r}_k^\star = \mathbf{D}_t(z) - \mathbf{r}_k \mathbf{r}_k^\star - \mathbf{r}_j \mathbf{r}_j^\star, \\
\alpha_{j,t}(z) &= \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-2}(z) \mathbf{r}_j - n^{-1} \operatorname{tr}(\mathbf{D}_{j,t}^{-2}(z) \mathbf{T}_n), \\
\gamma_{j,t}(z) &= \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z) \mathbf{r}_j - n^{-1} \mathbb{E} \operatorname{tr}(\mathbf{D}_{j,t}^{-1}(z) \mathbf{T}_n), \\
\gamma_{k,j,t}(z) &= \mathbf{r}_k^\star \mathbf{D}_{k,j,t}^{-1}(z) \mathbf{r}_k - n^{-1} \mathbb{E} \left[\operatorname{tr} \left(\mathbf{T}_n \mathbf{D}_{k,j,t}^{-1}(z) \right) \right] \\
\hat{\gamma}_{j,t}(z) &= \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z) \mathbf{r}_j - n^{-1} \operatorname{tr}(\mathbf{D}_{j,t}^{-1}(z) \mathbf{T}_n), \\
\beta_{j,t}(z) &= \frac{1}{1 + \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z) \mathbf{r}_j}, \\
\beta_{k,j,t}(z) &= \frac{1}{1 + \mathbf{r}_k^\star \mathbf{D}_{k,j,t}^{-1}(z) \mathbf{r}_k}, \\
\overline{\beta}_{j,t}(z) &= \frac{1}{1 + n^{-1} \operatorname{tr}(\mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z))}, \\
b_{j,t}(z) &= \frac{1}{1 + n^{-1} \mathbb{E} \operatorname{tr}(\mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z))}, \\
b_t(z) &= \frac{1}{1 + n^{-1} \mathbb{E} \operatorname{tr}(\mathbf{T}_n \mathbf{D}_t^{-1}(z))}.
\end{aligned}$$

Note that the terms $\beta_{j,t}(z), \beta_{k,j,t}(z), \overline{\beta}_{j,t}(z), b_{j,t}(z)$ and $b_t(z)$ are bounded in absolute value by $|z|/v$, where $v = \mathrm{Im}(z)$ is assumed to be positive (see (6.2.5) in Bai and Silverstein, 2010). We will denote constants appearing in the following inequalities by $K > 0$ (K may depend on t and z) and K may take on different values from line to line. By the Sherman–Morrison formula we obtain the representation

$$\mathbf{D}_t^{-1}(z) - \mathbf{D}_{j,t}^{-1}(z) = -\mathbf{D}_{j,t}^{-1}(z) \mathbf{r}_j \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z) \beta_{j,t}(z). \tag{6.14}$$

Step 1: *CLT for martingale difference schemes*
In order to prove asymptotic normality of the random variable appearing in (6.13), we show that it can be represented as a suitable martingale difference scheme plus some negligible remainder, which allows us to apply a central limit theorem.
For $j = 1 \ldots, n$, let $\mathbb{E}_j$ denote the conditional expectation with respect to the filtration $\mathcal{F}_{nj} = \sigma(\{\mathbf{r}_1, ..., \mathbf{r}_j\})$ (by $\mathbb{E}_0$ we denote the common expectation). Recalling

the definition (6.10) and using the martingale decomposition, we have

$$\begin{aligned}
M_n^1(z,t) &= \operatorname{tr}(\mathbf{D}_t^{-1}(z) - \mathbb{E}\mathbf{D}_t^{-1}(z)) \\
&= \sum_{j=1}^{\lfloor nt \rfloor} \left(\operatorname{tr}\mathbb{E}_j \mathbf{D}_t^{-1}(z) - \operatorname{tr}\mathbb{E}_{j-1}\mathbf{D}_t^{-1}(z)\right) \\
&= \sum_{j=1}^{\lfloor nt \rfloor} \left(\operatorname{tr}\mathbb{E}_j \left[\mathbf{D}_t^{-1}(z) - \mathbf{D}_{t,j}^{-1}(z)\right] - \operatorname{tr}\mathbb{E}_{j-1}\left[\mathbf{D}_t^{-1}(z) - \mathbf{D}_{j,t}^{-1}(z)\right]\right) \\
&= -\sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-2}(z)\mathbf{r}_j. \qquad (6.15)
\end{aligned}$$

Using the identity

$$\beta_{j,t}(z) = \overline{\beta}_{j,t}(z) - \overline{\beta}_{j,t}^2(z)\hat{\gamma}_{j,t}(z) + \overline{\beta}_{j,t}^2(z)\beta_{j,t}(z)\hat{\gamma}_{j,t}^2(z),$$

we write

$$\begin{aligned}
&(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-2}(z)\mathbf{r}_j \\
&\quad = \mathbb{E}_j\left(\overline{\beta}_{j,t}(z)\alpha_{j,t}(z) - \overline{\beta}_{j,t}^2(z)\hat{\gamma}_{j,t}(z)\frac{1}{n}\operatorname{tr}(\mathbf{T}_n\mathbf{D}_{j,t}^{-2}(z))\right) \\
&\quad - (\mathbb{E}_j - \mathbb{E}_{j-1})\overline{\beta}_{j,t}^2(z)\left(\hat{\gamma}_{j,t}(z)\alpha_{j,t}(z) - \beta_{j,t}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-2}(z)\mathbf{r}_j\hat{\gamma}_{j,t}^2(z)\right).
\end{aligned}$$

By considering the second moment, one can further show that

$$-\sum_{j=1}^{\lfloor nt \rfloor}(\mathbb{E}_j - \mathbb{E}_{j-1})\overline{\beta}_{j,t}^2(z)\left(\hat{\gamma}_{j,t}(z)\alpha_{j,t}(z) - \beta_{j,t}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-2}(z)\mathbf{r}_j\hat{\gamma}_{j,t}^2(z)\right) \xrightarrow{\mathbb{P}} 0.$$

Thus, it is sufficient to prove asymptotic normality for the quantity

$$\sum_{j=1}^{\max(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} Z_{nj}^{t_1,t_2} ,$$

where

$$Z_{nj}^{t_1,t_2} = \sum_{i=1}^{k}\left(\alpha_{i,1}Y_{j,t_1}(z_i) + \alpha_{i,2}Y_{j,t_2}(z_i)\right) , \qquad (6.16)$$

$$Y_{j,t}(z) = -\mathbb{E}_j\left[\overline{\beta}_{j,t}(z)\alpha_{j,t}(z) - \overline{\beta}_{j,t}^2(z)\hat{\gamma}_{j,t}(z)\frac{1}{n}\operatorname{tr}(\mathbf{T}_n\mathbf{D}_{j,t}^{-2}(z))\right] = -\mathbb{E}_j\frac{d}{dz}\overline{\beta}_{j,t}(z)\hat{\gamma}_{j,t}(z) \qquad (6.17)$$

if $j \le \lfloor nt \rfloor$ and $Y_{j,t}(z) = 0$ if $j > \lfloor nt \rfloor$.

For this purpose we verify conditions (5.29) - (5.31) of the central limit theorem for complex-valued martingale difference schemes given in Lemma 5.6 of Najim and

Yao (2016).

Lemma 7.1.1 in Section 7.1 shows that $Z_{nj}^{t_1,t_2}$ forms a martingale difference scheme with respect to the filtration $\mathcal{F}_{nj} = \sigma(\{\mathbf{r}_1, ..., \mathbf{r}_j\})$ and a proof of (5.31) in this reference is given in Lemma 7.1.2 by deriving bounds for the 4th moment of $Y_{j,t}(z)$. For a proof of condition (5.30), we note that

$$\sum_{j=1}^{\max(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \mathbb{E}_{j-1}\left[\left(Z_{nj}^{t_1,t_2}\right)^2\right] = \sum_{i,l=1}^{k} \Big(\sum_{j=1}^{\lfloor nt_1 \rfloor} \alpha_{i,1}\alpha_{l,1}\mathbb{E}_{j-1}[Y_{j,t_1}(z_i)Y_{j,t_1}(z_l)] \\ + \sum_{j=1}^{\min(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \alpha_{i,1}\alpha_{l,2}\mathbb{E}_{j-1}[Y_{j,t_1}(z_i)Y_{j,t_2}(z_l)] \\ + \sum_{j=1}^{\min(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \alpha_{i,2}\alpha_{l,1}\mathbb{E}_{j-1}[Y_{j,t_2}(z_i)Y_{j,t_1}(z_l)] \\ + \sum_{j=1}^{\lfloor nt_2 \rfloor} \alpha_{i,2}\alpha_{l,2}\mathbb{E}_{j-1}[Y_{j,t_2}(z_i)Y_{j,t_2}(z_l)] \Big) .$$

As all terms have the same form, it is sufficient to show that for all $z_1, z_2 \in \mathbb{C}$ with $\mathrm{Im}(z_1), \mathrm{Im}(z_2) \neq 0$ and $t_1, t_2 \in (0, 1]$

$$V_n(z_1, z_2, t_1, t_2) = \sum_{j=1}^{\min(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \mathbb{E}_{j-1}\left[Y_{j,t_1}(z_1)Y_{j,t_2}(z_2)\right] \xrightarrow{\mathbb{P}} \sigma^2_{t_1,t_2}(z_1, z_2) \tag{6.18}$$

for an appropriate function $\sigma^2_{t_1,t_2}(z_1, z_2)$ (see equation (6.44) below for a precise definition). Note that this convergence implies condition (5.29), since

$$\sum_{j=1}^{\min(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \mathbb{E}_{j-1}\big[Y_{j,t_1}(z_1)\overline{Y_{j,t_2}(z_2)}\big] = \sum_{j=1}^{\min(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \mathbb{E}_{j-1}\left[Y_{j,t_1}(z_1)Y_{j,t_2}(\overline{z_2})\right] \xrightarrow{\mathbb{P}} \sigma^2_{t_1,t_2}(z_1, \overline{z_2}),$$

where the equality follows from the fact that the matrices $\mathbf{T}_n, \mathbf{B}_{n,t}, \mathbf{r}_j\mathbf{r}_j^\star$ are Hermitian and $(\overline{\mathbf{D}_{j,t}^{-1}(z)})^T = \mathbf{D}_{j,t}^{-1}(\overline{z})$. Consequently, Lemma 5.6 in Najim and Yao (2016) combined with the Cramer–Wold device yields the weak convergence of the finite-dimensional distributions to a multivariate normal distribution with covariance $\sigma^2_{t_1,t_2}(z_1, \overline{z_2}) = \mathrm{cov}(M^1(z_1, t_1), M^1(z_2, t_2))$.

Step 2: *Calculation of the covariance structure*
The rest of this proof is devoted to the calculation of $\sigma^2_{t_1,t_2}(z_1, z_2)$, which gives us the covariance structure of our process $(M^1(z, t))_{z\in\mathcal{C}^+, t\in[t_0,1]}$. We will first express our random variable of interest in (6.18) through the derivative of another random variable and find a more handy representation for this new random variable. We finally determine its limit in two further sub-steps.

Consider the sum

$$V_n^{(0)}(z_1,z_2,t_1,t_2) = \sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} \mathbb{E}_{j-1}\left[\mathbb{E}_j\left(\overline{\beta}_{j,t_1}(z_1)\hat{\gamma}_{j,t_1}(z_1)\right)\mathbb{E}_j\left(\overline{\beta}_{j,t_2}(z_2)\hat{\gamma}_{j,t_2}(z_2)\right)\right]. \tag{6.19}$$

We use the dominated convergence theorem in combination with (6.17) to get

$$\frac{\partial^2}{\partial z_1 \partial z_2} V_n^{(0)}(z_1,z_2,t_1,t_2) = V_n(z_1,z_2,t_1,t_2). \tag{6.20}$$

In the monograph of Bai and Silverstein (2010), it is shown that it suffices to show that $V_n^{(0)}(z_1,z_2,t_1,t_2)$ given in (6.19) converges in probability to a constant and in this case, the mixed partial derivative of its limit will give the limit of $V_n(z_1,z_2,t_1,t_2)$. It holds

$$\mathbb{E}|\overline{\beta}_{j,t}(z) - b_{j,t}(z)|^2 \leq Kn^{-1},$$

which implies together with (7.3) that

$$\begin{aligned}
&\mathbb{E}\Big|\mathbb{E}_{j-1}\left[\mathbb{E}_j\left(\overline{\beta}_{j,t_1}(z_1)\hat{\gamma}_{j,t_1}(z_1)\right)\mathbb{E}_j\left(\overline{\beta}_{j,t_2}(z_2)\hat{\gamma}_{j,t_2}(z_2)\right)\right] \\
&\quad - \mathbb{E}_{j-1}\left[\mathbb{E}_j\left(b_{j,t_1}(z_1)\hat{\gamma}_{j,t_1}(z_1)\right)\mathbb{E}_j\left(b_{j,t_2}(z_2)\hat{\gamma}_{j,t_2}(z_2)\right)\right]\Big| \\
=&\mathbb{E}\Big|\mathbb{E}_{j-1}\left[\mathbb{E}_j\left(\left(\overline{\beta}_{j,t_1}(z_1) - b_{j,t_1}(z_1)\right)\hat{\gamma}_{j,t_1}(z_1)\right)\mathbb{E}_j\left(\overline{\beta}_{j,t_2}(z_2)\hat{\gamma}_{j,t_2}(z_2)\right)\right] \\
&\quad + \mathbb{E}_{j-1}\left[\mathbb{E}_j\left(b_{j,t_1}(z_1)\hat{\gamma}_{j,t_1}(z_1)\right)\mathbb{E}_j\left(\left(\overline{\beta}_{j,t_2}(z_2) - b_{j,t_2}(z_2)\right)\hat{\gamma}_{j,t_2}(z_2)\right)\right]\Big| \\
\leq& K\mathbb{E}^{\frac{1}{2}}\left|\overline{\beta}_{j,t_1}(z_1) - b_{j,t_1}(z_1)\right|^2 \mathbb{E}^{\frac{1}{4}}\left|\hat{\gamma}_{j,t_1}(z_1)\right|^4 \mathbb{E}^{\frac{1}{4}}\left|\hat{\gamma}_{j,t_2}(z_2)\right|^4 \\
&\quad + K\mathbb{E}^{\frac{1}{4}}\left|\hat{\gamma}_{j,t_1}(z_1)\right|^4 \mathbb{E}^{\frac{1}{2}}\left|\overline{\beta}_{j,t_2}(z_2) - b_{j,t_2}(z_2)\right|^2 \mathbb{E}^{\frac{1}{4}}\left|\hat{\gamma}_{j,t_2}(z_2)\right|^4 \\
=&o\left(n^{-1}\right).
\end{aligned}$$

Since $b_{j,t}(z)$ is nonrandom, we obtain

$$\begin{aligned}
&\sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} \mathbb{E}_{j-1}\left[\mathbb{E}_j\left[\overline{\beta}_{j,t_1}(z_1)\hat{\gamma}_{j,t_1}(z_1)\right]\mathbb{E}_j\left[\overline{\beta}_{j,t_2}(z_2)\hat{\gamma}_{j,t_2}(z_2)\right]\right] \\
&- \sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} b_{j,t_1}(z_1)b_{j,t_2}(z_2)\mathbb{E}_{j-1}\left[\mathbb{E}_j\left[\hat{\gamma}_{j,t_1}(z_1)\right]\mathbb{E}_j\left[\hat{\gamma}_{j,t_2}(z_2)\right]\right] \\
&\overset{\mathbb{P}}{\to} 0.
\end{aligned}$$

Consequently, we have

$$V_n^{(0)}(z_1,z_2,t_1,t_2) = V_n^{(1)}(z_1,z_2,t_1,t_2) + o_{\mathbb{P}}(1),$$

where

$$V_n^{(1)}(z_1,z_2,t_1,t_2) = \sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} b_{j,t_1}(z_1)b_{j,t_2}(z_2)\mathbb{E}_{j-1}\left[\mathbb{E}_j\left[\hat{\gamma}_{j,t_1}(z_1)\right]\mathbb{E}_j\left[\hat{\gamma}_{j,t_2}(z_2)\right]\right].$$

Observing (6.20), the mixed partial derivative of $V_n^{(1)}(z_1,z_2,t_1,t_2)$ is asymptotically equivalent to $V_n(z_1,z_2,t_1,t_2)$.
For the complex case, we have $\mathbb{E}x_{ij}^2 = o(1/n)$, $\mathbb{E}|x_{ij}|^4 = 2 + o(1)$ and it follows that

$$\begin{aligned}
&V_n^{(1)}(z_1,z_2,t_1,t_2)\\
&= \sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} b_{j,t_1}(z_1)b_{j,t_2}(z_2)\mathbb{E}_{j-1}\left[\prod_{k=1}^{2}\left(\mathbb{E}_j\left(\mathbf{r}_j^{\star}\mathbf{D}_{j,t_k}^{-1}(z_k)\mathbf{r}_j - n^{-1}\operatorname{tr}\left(\mathbf{D}_{j,t_k}^{-1}(z_k)\mathbf{T}_n\right)\right)\right)\right]\\
&= \sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} b_{j,t_1}(z_1)b_{j,t_2}(z_2)\mathbb{E}_{j-1}\Big[\prod_{k=1}^{2}\Big(n^{-1}\mathbf{x}_j^{\star}\mathbf{T}_n^{\frac{1}{2}}\mathbb{E}_j\left(\mathbf{D}_{j,t_k}^{-1}(z_k)\right)\mathbf{T}_n^{\frac{1}{2}}\mathbf{x}_j\\
&\qquad\qquad - n^{-2}\operatorname{tr}\mathbb{E}_j\left(\mathbf{D}_{j,t_k}^{-1}(z_k)\mathbf{T}_n\right)\Big)\Big]\\
&= \frac{1}{n^2}\sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} b_{j,t_1}(z_1)b_{j,t_2}(z_2)\left(\operatorname{tr}\left(\mathbf{T}_n^{\frac{1}{2}}\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}_{j,t_2}^{-1}(z_2)\right]\mathbf{T}_n^{\frac{1}{2}}\right) + o(1)A_n\right),
\end{aligned} \tag{6.21}$$

where

$$A_n = O(n).$$

Hence, it suffices to study the limit of

$$V_n^{(2)}(z_1,z_2,t_1,t_2) = \frac{1}{n^2}\sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} b_{j,t_1}(z_1)b_{j,t_2}(z_2)\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}_{j,t_2}^{-1}(z_2)\right]\mathbf{T}_n\right). \tag{6.22}$$

For (6.21), we used the following identity

$$\begin{aligned}
\mathbb{E}_{j-1}\left[\prod_{k=1}^{2}\left(\mathbf{x}_j^{\star}\mathbb{E}_j\left[\mathbf{B}^{(k)}\right]\mathbf{x}_j - \operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{B}^{(k)}\right]\right)\right)\right] &= \sum_{i=1}^{p}\left(\mathbb{E}|x_{ij}|^4 - \left(\mathbb{E}x_{ij}^2\right)^2 - 2\right)b_{ii}^{(1)}b_{ii}^{(2)}\\
&\quad + \operatorname{tr}(\mathbf{B}_x^{(1)}(\mathbf{B}_x^{(2)})^T) + \operatorname{tr}(\mathbf{B}^{(1)}\mathbf{B}^{(2)}),
\end{aligned} \tag{6.23}$$

where

$$\begin{aligned}
\mathbf{B}^{(k)} &= \left(b_{il}^{(k)}\right)_{il} = \mathbf{T}_n^{\frac{1}{2}}\mathbb{E}_j\left[\mathbf{D}_{j,t_k}^{-1}(z_k)\right]\mathbf{T}_n^{\frac{1}{2}},\\
\mathbf{B}_x^{(k)} &= \left(b_{il}^{(k)}\mathbb{E}x_{ij}^2\right)_{il}.
\end{aligned}$$

One can observe that under the assumptions for the complex case, only the last term on the right remains in (6.23) whereas those for the real case leave the last two. That is, in the real case, we have to consider two times the limit of $V_n^{(2)}(z_1,z_2,t_1,t_2)$ in (6.22). (For instance, we have for the real case $\mathbb{E}[x_{ij}^2] = 1$ and $\mathbb{E}[x_{ij}^4] = 3 + o(1)$

uniformly in $i \in \{1, \ldots, n\}, j \in \{1, \ldots, p\}$ (due to the truncation steps). Thus, in the real case, the first term

$$\frac{1}{n^2} \sum_{j=1}^{\min(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \sum_{i=1}^{p} \left(\mathbb{E}|x_{ij}|^4 - \left(\mathbb{E}x_{ij}^2\right)^2 - 2 \right) b_{ii}^{(1)} b_{ii}^{(2)}$$

is asymptotically negligible, and we have $\mathbf{B}_x^{(k)} = \mathbf{B}^{(k)}$ for $k = 1, 2$. The complex case can be handled similarly.)

Step 2.1: *Decomposition of* $\mathbf{D}_{j,t}^{-1}(z)$
As suggested by the structure of the random variable $V_n^{(2)}(z_1, z_2, t_1, t_1)$ given in (6.22), we need to study the random matrices $\mathbf{D}_{j,t}^{-1}(z)$ further. For this aim, we will introduce a decomposition of $\mathbf{D}_{j,t}^{-1}(z)$.
We recall the definitions

$$\begin{aligned} \mathbf{D}_{i,j,t}(z) &= \mathbf{D}_t(z) - \mathbf{r}_i \mathbf{r}_i^\star - \mathbf{r}_j \mathbf{r}_j^\star, \\ \beta_{i,j,t}(z) &= \frac{1}{1 + \mathbf{r}_i^\star \mathbf{D}_{i,j,t}^{-1}(z) \mathbf{r}_i}, \\ b_{i,j,t}(z) &= \frac{1}{1 + n^{-1} \mathbb{E} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{i,j,t}^{-1}(z)}, \end{aligned}$$

and note that

$$\mathbf{D}_{j,t}(z) + z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z) \mathbf{T}_n = \sum_{i \neq j, 1 \leq i \leq \lfloor nt \rfloor} \mathbf{r}_i \mathbf{r}_i^\star - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z) \mathbf{T}_n.$$

Multiplying by $(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z) \mathbf{T}_n)^{-1}$ on the left, $\mathbf{D}_{j,t}^{-1}(z)$ on the right, and using

$$\mathbf{r}_i^\star \mathbf{D}_{j,t}^{-1}(z) = \beta_{i,j,t}(z) \mathbf{r}_i^\star \mathbf{D}_{i,j,t}^{-1}(z)$$

(this is a consequence of the Sherman-Morrison-Woodbury formula; see also formula (6.1.11) in Bai and Silverstein (2010)), we have

$$\begin{aligned} \mathbf{D}_{j,t}^{-1}(z) = &- \left(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z) \mathbf{T}_n \right)^{-1} \qquad (6.24) \\ &+ \sum_{i \neq j, 1 \leq i \leq \lfloor nt \rfloor} \beta_{i,j,t}(z) \left(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z) \mathbf{T}_n \right)^{-1} \mathbf{r}_i \mathbf{r}_i^\star \mathbf{D}_{i,j,t}^{-1}(z) \\ &- \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z_1) \left(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z) \\ = &- \left(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z) \mathbf{T}_n \right)^{-1} + b_{j,t}(z) \mathbf{A}_t(z) + \mathbf{B}_t(z) + \mathbf{C}_t(z), \qquad (6.25) \end{aligned}$$

where for fixed j

$$\mathbf{A}_t(z) = \sum_{\substack{i=1\\i\neq j}}^{\lfloor nt\rfloor} \left(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\right)^{-1} \left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right) \mathbf{D}_{i,j,t}^{-1}(z), \tag{6.26}$$

$$\mathbf{B}_t(z) = \sum_{\substack{i=1\\i\neq j}}^{\lfloor nt\rfloor} \left(\beta_{i,j,t}(z) - b_{j,t}(z)\right) \left(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\right)^{-1} \mathbf{r}_i\mathbf{r}_i^\star \mathbf{D}_{i,j,t}^{-1}(z), \tag{6.27}$$

$$\mathbf{C}_t(z) = b_{j,t}(z) \left(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\right)^{-1} \mathbf{T}_n n^{-1} \sum_{\substack{i=1\\i\neq j}}^{\lfloor nt\rfloor} \left(\mathbf{D}_{i,j,t}^{-1}(z) - \mathbf{D}_{j,t}^{-1}(z)\right). \tag{6.28}$$

(Here, we do not reflect the dependence on index j in our notation.) Considering the term $\mathbf{A}_t(z)$ more carefully, we see that $\mathbf{r}_i$ is independent of $\mathbf{r}_1, ..., \mathbf{r}_j$ and of $\mathbf{D}_{i,j,t}(z)$ for $i > j$, which implies

$$\begin{aligned}&\mathbb{E}_j \left[\left(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\right)^{-1} \left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right) \mathbf{D}_{i,j,t}^{-1}(z)\right] \\ &= \left(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\right)^{-1} \left(\mathbb{E}\left[\mathbf{r}_i\mathbf{r}_i^\star\right] - n^{-1}\mathbf{T}_n\right) \mathbb{E}_j\left[\mathbf{D}_{i,j,t}^{-1}(z)\right] = 0.\end{aligned}$$

This means, that in the definition of $\mathbf{A}_t$ we only have to consider summands with $i < j$.
From formula (9.9.13) in Bai and Silverstein (2010), we conclude that

$$\left|\left|\left(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\right)^{-1}\right|\right| \leq K. \tag{6.29}$$

Let $\mathbf{M}$ be a $p \times p$ (random) matrix and let $||\mathbf{M}||$ denote a nonrandom bound on the spectral norm of $\mathbf{M}$ for all parameters governing $\mathbf{M}$ and all realizations of $\mathbf{M}$.
From formula (9.9.5) in Bai and Silverstein (2010), we get

$$|b_{i,j,t}(z) - b_{j,t}(z)| \leq Kn^{-1}, \tag{6.30}$$

$$\mathbb{E}|\beta_{i,j,t}(z) - b_{j,t}(z)|^2 \leq Kn^{-1}. \tag{6.31}$$

From formula (9.9.6) in Bai and Silverstein (2010), (6.29), Hölder's inequality and the bound (7.2) on $\mathbb{E}||\mathbf{D}_{i,j,t}^{-1}(z)||^2$, we conclude

$$\begin{aligned}&\mathbb{E}|\operatorname{tr}(\mathbf{B}_t(z)\mathbf{M})| \\ \leq &\sum_{\substack{i=1\\i\neq j}}^{\lfloor nt\rfloor} \mathbb{E}^{\frac{1}{2}} \left|\beta_{i,j,t}(z) - b_{j,t}(z)\right|^2 \mathbb{E}^{\frac{1}{2}} \left|\mathbf{r}_i^\star \mathbf{D}_{i,j,t}^{-1}(z)\mathbf{M}\left(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\right)^{-1} \mathbf{r}_i\right|^2 \\ \leq &K||\mathbf{M}||n^{\frac{1}{2}}.\end{aligned} \tag{6.32}$$

Using formula (9.9.6) of Bai and Silverstein (2010), (6.29) and the bounds on $b_{j,t}(z)$ and $||\mathbf{T}_n||$ yields

$$|\mathrm{tr}(\mathbf{C}_t(z)\mathbf{M})| \leq K||\mathbf{M}||. \tag{6.33}$$

Moreover, we have for nonrandom $\mathbf{M}$ and any j by using (9.9.6) in Bai and Silverstein (2010), (7.2), and (6.29)

$$\mathbb{E}\,|\mathrm{tr}\,\mathbf{A}_t(z)\mathbf{M}| \leq K||\mathbf{M}||. \tag{6.34}$$

Step 2.2: *Application of Step 2.1 to* $V_n^{(2)}$
In this final step, we use the decomposition for $\mathbf{D}_{j,t}^{-1}(z)$ derived in Step 2.1 to determine the limit of the random variable $V_n^{(2)}$ given in (6.22). We remind the reader that the main part of (6.22) is

$$\mathrm{tr}\left(\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}_{j,t_2}^{-1}(z_2)\right]\mathbf{T}_n\right).$$

Thus, we need to apply the decomposition twice, namely for $\mathbf{D}_{j,t_1}^{-1}(z_1)$ and $\mathbf{D}_{j,t_2}^{-1}(z_2)$. Observing (6.20), we finally obtain the covariance $\sigma_{t_1,t_2}^2(z_1,z_2)$ by differentiating the limit of (6.22).
Using (6.14), we decompose

$$\mathrm{tr}\left(\mathbb{E}_j[\mathbf{A}_{t_1}(z_1)]\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right) = A_1(z_1,z_2) + A_2(z_1,z_2) + A_3(z_1,z_2),$$

where

$$\begin{aligned}
A_1(z_1,z_2) =& -\mathrm{tr}\left(\sum_{i<j,1\leq i\leq\lfloor nt_1\rfloor}\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{r}_i\mathbf{r}_i^{\star}\mathbb{E}_j[\mathbf{D}_{i,j,t_1}^{-1}(z_1)]\right.\\
&\left.\times\mathbf{T}_n\beta_{i,j,t_2}(z_2)\mathbf{D}_{i,j,t_2}^{-1}(z_2)\mathbf{r}_i\mathbf{r}_i^{\star}\mathbf{D}_{i,j,t_2}^{-1}(z_2)\mathbf{T}_n\right)\\
=& -\sum_{i<j,1\leq i\leq\lfloor nt_1\rfloor}\beta_{i,j,t_2}(z_2)\mathbf{r}_i^{\star}\mathbb{E}_j\left[\mathbf{D}_{i,j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbf{D}_{i,j,t_2}^{-1}(z_2)\mathbf{r}_i\mathbf{r}_i^{\star}\mathbf{D}_{i,j,t_2}^{-1}(z_2)\mathbf{T}_n\\
&\times\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{r}_i,\\
A_2(z_1,z_2) =& -\mathrm{tr}\left(\sum_{i<j,1\leq i\leq\lfloor nt_1\rfloor}\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}n^{-1}\mathbf{T}_n\mathbb{E}_j[\mathbf{D}_{i,j,t_1}^{-1}(z_1)]\mathbf{T}_n\right.\\
&\left.\times\left(\mathbf{D}_{j,t_2}^{-1}(z_2)-\mathbf{D}_{i,j,t_2}^{-1}(z_2)\right)\mathbf{T}_n\right),\\
A_3(z_1,z_2) =& \,\mathrm{tr}\left(\sum_{i<j,1\leq i\leq\lfloor nt_1\rfloor}\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\left(\mathbf{r}_i\mathbf{r}_i^{\star}-n^{-1}\mathbf{T}_n\right)\right.\\
&\left.\times\mathbb{E}_j[\mathbf{D}_{i,j,t_1}^{-1}(z_1)]\mathbf{T}_n\mathbf{D}_{i,j,t_2}^{-1}(z_2)\mathbf{T}_n\right).
\end{aligned}$$

As in Bai and Silverstein (2010), it can be shown that

$$|A_2(z_1,z_2)| \leq K, \tag{6.35}$$
$$\mathbb{E}|A_3(z_1,z_2)| \leq Kn^{\frac{1}{2}}, \tag{6.36}$$

and it remains to investigate the term $A_1(z_1,z_2)$. For this purpose, we define

$$\mathbf{E} = \mathbb{E}_j \left[\mathbf{D}^{-1}_{i,j,t_1}(z_1)\right] \mathbf{T}_n \mathbf{D}^{-1}_{i,j,t_2}(z_2),$$
$$\mathbf{F} = \mathbf{D}^{-1}_{i,j,t_2}(z_2)\mathbf{T}_n \left(z_1\mathbf{I} - \frac{\lfloor nt_1 \rfloor - 1}{n} b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}$$

and note that for $q \geq 1$

$$\mathbb{E}^{\frac{1}{q}} |\operatorname{tr}(\mathbf{E}\mathbf{T}_n)|^q \leq K_q n, \ \mathbb{E}^{\frac{1}{q}} |\operatorname{tr}(\mathbf{F}\mathbf{T}_n)|^q \leq K_q n.$$

This gives together with formula (9.9.6) in Bai and Silverstein (2010)

$$\begin{aligned}
&\mathbb{E}\Big|\beta_{i,j,t_2}(z_2)\mathbf{r}_i^\star \mathbf{E}\mathbf{r}_i \mathbf{r}_i^\star \mathbf{F}\mathbf{r}_i - b_{j,t_2}(z_2) n^{-2} \operatorname{tr}(\mathbf{E}\mathbf{T}_n) \operatorname{tr}(\mathbf{F}\mathbf{T}_n)\Big| \\
\leq & n^{-2}\mathbb{E}\Big|\beta_{i,j,t_2}(z_2) x_i^\star \mathbf{T}_n^{\frac{1}{2}} \mathbf{E} \mathbf{T}_n^{\frac{1}{2}} x_i \left(x_i^\star \mathbf{T}_n^{\frac{1}{2}} \mathbf{F} \mathbf{T}_n^{\frac{1}{2}} x_i - \operatorname{tr}(\mathbf{F}\mathbf{T}_n)\right)\Big| \\
& + n^{-2}\mathbb{E}\Big| \left(\beta_{i,j,t_2}(z_2) - b_{j,t_2}(z_2)\right) \operatorname{tr}(\mathbf{E}\mathbf{T}_n) \operatorname{tr}(\mathbf{F}\mathbf{T}_n)\Big| \\
& + n^{-2}\mathbb{E}\Big|\beta_{i,j,t_2}(z_2) \left(x_i^\star \mathbf{T}_n^{\frac{1}{2}} \mathbf{E} \mathbf{T}_n^{\frac{1}{2}} x_i - \operatorname{tr}(\mathbf{E}\mathbf{T}_n)\right) \operatorname{tr}(\mathbf{F}\mathbf{T}_n)\Big| \\
\leq & Kn^{-\frac{1}{2}}
\end{aligned} \tag{6.37}$$

Moreover, letting

$$\mathbf{G} = \mathbf{T}_n \left(z_1\mathbf{I} - \frac{\lfloor nt_1 \rfloor - 1}{n} b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1} \mathbf{T}_n,$$

we have by using (9.9.5)

$$\begin{aligned}
&\Big|\operatorname{tr}(\mathbf{E}\mathbf{T}_n) \operatorname{tr}(\mathbf{F}\mathbf{T}_n) - \operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right] \mathbf{T}_n \mathbf{D}^{-1}_{j,t_2}(z_2)\mathbf{T}_n\right) \operatorname{tr}\left(\mathbf{D}^{-1}_{j,t_2}(z_2)\mathbf{G}\right)\Big| \\
=&\Big|\operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}^{-1}_{i,j,t_1}(z_1)\right] \mathbf{T}_n \mathbf{D}^{-1}_{i,j,t_2}(z_2)\mathbf{T}_n\right) \operatorname{tr}\left(\mathbf{D}^{-1}_{i,j,t_2}(z_2)\mathbf{G}\right) \\
&- \operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right] \mathbf{T}_n \mathbf{D}^{-1}_{j,t_2}(z_2)\mathbf{T}_n\right) \operatorname{tr}\left(\mathbf{D}^{-1}_{j,t_2}(z_2)\mathbf{G}\right)\Big| \\
\leq&\Big|\operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}^{-1}_{i,j,t_1}(z_1) - \mathbf{D}^{-1}_{j,t_1}(z_1)\right] \mathbf{T}_n \mathbf{D}^{-1}_{i,j,t_2}(z_2)\mathbf{T}_n\right) \operatorname{tr}\left(\mathbf{D}^{-1}_{i,j,t_2}(z_2)\mathbf{G}\right)\Big| \\
&+ \Big|\operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right] \mathbf{T}_n \left(\mathbf{D}^{-1}_{i,j,t_2}(z_2) - \mathbf{D}^{-1}_{j,t_2}(z_2)\right)\mathbf{T}_n\right) \operatorname{tr}\left(\mathbf{D}^{-1}_{i,j,t_2}(z_2)\mathbf{G}\right)\Big| \\
&+ \Big|\operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right] \mathbf{T}_n \mathbf{D}^{-1}_{j,t_2}(z_2)\mathbf{T}_n\right) \operatorname{tr}\left(\left(\mathbf{D}^{-1}_{i,j,t_2}(z_2) - \mathbf{D}^{-1}_{j,t_2}(z_2)\right)\mathbf{G}\right)\Big| \\
\leq& Kn.
\end{aligned} \tag{6.38}$$

The considerations in (6.37) and (6.38) imply

$$\mathbb{E}\Big|A_1(z_1,z_2) + \frac{j-1}{n^2} b_{j,t_2}(z_2) \operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right] \mathbf{T}_n \mathbf{D}^{-1}_{j,t_2}(z_2)\mathbf{T}_n\right)$$

$$\times \operatorname{tr}\left(\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\Bigg|$$
$$\le Kn^{\frac{1}{2}}.$$

Recall that we aim to derive a representation for

$$\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right).$$

For this purpose, we use the formula for $\mathbf{D}_{j,t_1}^{-1}(z_1)$ given in (6.25), we have (using also (6.32), (6.33), (6.35), (6.36))

$$\begin{aligned}
&\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right)\Bigg[1+\frac{j-1}{n^2}b_{j,t_1}(z_1)b_{j,t_2}(z_2)\\
&\operatorname{tr}\left(\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\Bigg]\\
=&\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right)-b_{j,t_1}(z_1)A_1(z_1,z_2)+A_4(z_1,z_2)\\
=&-\operatorname{tr}\left(\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right)\\
&+\operatorname{tr}\left(\mathbb{E}_j[b_{j,t_1}(z_1)\mathbf{A}_{t_1}(z_1)+\mathbf{B}_{t_1}(z_1)+\mathbf{C}_{t_1}(z_1)]\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\right)\\
&-b_{j,t_1}(z_1)A_1(z_1,z_2)+A_4(z_1,z_2)\\
=&-\operatorname{tr}\left(\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right)\\
&+b_{j,t_1}(z_1)\left(A_2(z_1,z_2)+A_3(z_1,z_2)\right)+A_4(z_1,z_2)\\
=&-\operatorname{tr}\left(\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right)+A_4(z_1,z_2),
\end{aligned}$$

where we have used the representation in (6.25), the estimates (6.32), (6.33), (6.36) and the term $A_4(z_1,z_2)$ may change from line to line with the universal property

$$\mathbb{E}|A_4(z_1,z_2)|\le Kn^{\frac{1}{2}}.$$

Now using again the representation for $\mathbf{D}_{j,t_2}^{-1}(z_2)$ in (6.25) yields

$$\begin{aligned}
&\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbf{D}_{j,t_2}^{-1}(z_2)\mathbf{T}_n\right)\\
&\times\Bigg[1-\frac{j-1}{n^2}b_{j,t_1}(z_1)b_{j,t_2}(z_2)\operatorname{tr}\Bigg(\Bigg\{\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}\\
&-b_{j,t_2}(z_2)\mathbf{A}_{t_2}(z_2)-\mathbf{B}_{t_2}(z_2)-\mathbf{C}_{t_2}(z_2)\Bigg\}\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\Bigg)\Bigg]\\
=&\operatorname{tr}\Bigg(\Bigg\{\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}-b_{j,t_2}(z_2)\mathbf{A}_{t_2}(z_2)-\mathbf{B}_{t_2}(z_2)-\mathbf{C}_{t_2}(z_2)\Bigg\}
\end{aligned}$$

$$\times \mathbf{T}_n \left(z_1 \mathbf{I} - \frac{\lfloor nt_1 \rfloor - 1}{n} b_{j,t_1}(z_1) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \Bigg) + A_4(z_1, z_2),$$

Invoking (6.32), (6.33) and (6.34), we conclude

$$\begin{aligned}
&\operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right] \mathbf{T}_n \mathbf{D}_{j,t_2}^{-1}(z_2) \mathbf{T}_n\right) \\
&\times \Bigg[1 - \frac{j-1}{n^2} b_{j,t_1}(z_1) b_{j,t_2}(z_2) \operatorname{tr}\Bigg(\left(z_2 \mathbf{I} - \frac{\lfloor nt_2 \rfloor - 1}{n} b_{j,t_2}(z_2) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \\
&\times \left(z_1 \mathbf{I} - \frac{\lfloor nt_1 \rfloor - 1}{n} b_{j,t_1}(z_1) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \Bigg) \Bigg] \\
=& \operatorname{tr}\left(\left(z_2 \mathbf{I} - \frac{\lfloor nt_2 \rfloor - 1}{n} b_{j,t_2}(z_2) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \left(z_1 \mathbf{I} - \frac{\lfloor nt_1 \rfloor - 1}{n} b_{j,t_1}(z_1) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \right) \\
&+ A_5(z_1, z_2),
\end{aligned}$$

where the remainder A_5 satisfies

$$\mathbb{E}|A_5(z_1, z_2)| \leq K n^{\frac{1}{2}}.$$

This implies for the conditional expectation $\mathbb{E}_j$ with respect to $\mathbf{r}_1, \ldots, \mathbf{r}_j$

$$\begin{aligned}
&\operatorname{tr}\left(\mathbb{E}_j \left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right] \mathbf{T}_n \mathbb{E}_j \left[\mathbf{D}_{j,t_2}^{-1}(z_2)\right] \mathbf{T}_n\right) \\
&\times \Bigg[1 - \frac{j-1}{n^2} b_{j,t_1}(z_1) b_{j,t_2}(z_2) \operatorname{tr}\Bigg(\left(z_2 \mathbf{I} - \frac{\lfloor nt_2 \rfloor - 1}{n} b_{j,t_2}(z_2) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \\
&\times \left(z_1 \mathbf{I} - \frac{\lfloor nt_1 \rfloor - 1}{n} b_{j,t_1}(z_1) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \Bigg) \Bigg] \\
=& \operatorname{tr}\left(\left(z_2 \mathbf{I} - \frac{\lfloor nt_2 \rfloor - 1}{n} b_{j,t_2}(z_2) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \times \left(z_1 \mathbf{I} - \frac{\lfloor nt_1 \rfloor - 1}{n} b_{j,t_1}(z_1) \mathbf{T}_n \right)^{-1} \mathbf{T}_n \right) \\
&+ \tilde{A}_5(z_1, z_2),
\end{aligned}$$

where $\tilde{A}_5(z_1, z_2) = \mathbb{E}_j[A_5(z_1, z_2)]$. Hence,

$$\mathbb{E}|\tilde{A}_5(z_1, z_2)| \leq \mathbb{E}\left[\mathbb{E}_j |A_5(z_1, z_2)|\right] \leq K n^{\frac{1}{2}}.$$

Let $\tilde{\underline{s}}_{n,t}$ be the Stieltjes transform of $F^{\underline{\mathbf{B}}_{n,t}}$, where $\underline{\mathbf{B}}_{n,t}$ is the companion matrix of $\mathbf{B}_{n,t}$ defined in (2.2) and let $\tilde{\underline{s}}_{n,t}^0$ be the Stieltjes transform of $\underline{\tilde{F}}^{y_{\lfloor nt \rfloor}, H_n}$, that is

$$\begin{aligned}
\tilde{\underline{s}}_{n,t} &= s_{F^{\underline{\mathbf{B}}_{n,t}}}, \\
\tilde{\underline{s}}_{n,t}^0 &= s_{\underline{\tilde{F}}^{y_{\lfloor nt \rfloor}, H_n}}.
\end{aligned}$$

By Lemma 7.1.3, we get

$$\frac{1}{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}\left[\beta_{j,t}(z)\right] = -z \mathbb{E}\left[\tilde{\underline{s}}_{n,t}(z)\right].$$

We use Theorem 6.3.8 given in Section 6.3.4 and conclude

$$|\mathbb{E}\underline{\tilde{s}}_{n,t}(z) - \underline{\tilde{s}}^0_{n,t}(z)| \leq Kn^{-1}.$$

Combining this with the following bounds

$$|b_{j,t}(z) - b_t(z)| \leq Kn^{-1},$$
$$|b_{j,t}(z) - \mathbb{E}\beta_{j,t}(z)| \leq Kn^{-\frac{1}{2}},$$

we have for all $j \in \{1, \dots, \lfloor nt \rfloor\}$

$$\begin{aligned}
&\left|b_{j,t}(z) + z\underline{\tilde{s}}^0_{n,t}(z)\right| \\
\leq &\left|b_{j,t}(z) - b_t(z)\right| + \left|b_t(z) - \frac{1}{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}[\beta_{j,t}(z)]\right| + \left|z\underline{\tilde{s}}^0_{n,t}(z) - z\mathbb{E}\left[\underline{\tilde{s}}_{n,t}(z)\right]\right| \\
\leq &Kn^{-1} + \frac{1}{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor}\left(\left|b_t(z) - b_{j,t}(z)\right| + \left|b_{j,t}(z) - \mathbb{E}\beta_{j,t}(z)\right|\right) \\
\leq &Kn^{-\frac{1}{2}}.
\end{aligned} \tag{6.39}$$

This yields

$$\max_j |b_{j,t}(z) + z\underline{\tilde{s}}^0_{n,t}(z)| \leq Kn^{-\frac{1}{2}}, \tag{6.40}$$

and implies

$$\begin{aligned}
&\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_2}(z_2)\right]\mathbf{T}_n\right) \\
&\times\Bigg\{1 - \frac{j-1}{n^2}\underline{\tilde{s}}^0_{n,t_1}(z_1)\underline{\tilde{s}}^0_{n,t_2}(z_2)\operatorname{tr}\Bigg(\left(\mathbf{I} + \frac{\lfloor nt_2 \rfloor}{n}\underline{\tilde{s}}^0_{n,t_2}(z_2)\mathbf{T}_n\right)^{-1} \qquad (6.41) \\
&\times\mathbf{T}_n\left(\mathbf{I} + \frac{\lfloor nt_1 \rfloor}{n}\underline{\tilde{s}}^0_{n,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\Bigg)\Bigg\} \\
=&\frac{1}{z_1z_2}\operatorname{tr}\left(\left(\mathbf{I} + \frac{\lfloor nt_2 \rfloor}{n}\underline{\tilde{s}}^0_{n,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I} + \frac{\lfloor nt_1 \rfloor}{n}\underline{\tilde{s}}^0_{n,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right) + A_6(z_1, z_2),
\end{aligned} \tag{6.42}$$

where the remainder A_6 may change from line to line and satisfies

$$\mathbb{E}|A_6(z_1, z_2)| \leq Kn^{\frac{1}{2}}$$

(the details for this estimate are given in Lemma 7.1.4 and 7.1.5). Recalling that $H_n = F^{\mathbf{T}_n}$ is the empirical spectral distribution of $\mathbf{T}_n$, we have

$$\operatorname{tr}\left(\left(\mathbf{I} + \frac{\lfloor nt_2 \rfloor}{n}\underline{\tilde{s}}^0_{n,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I} + \frac{\lfloor nt_1 \rfloor}{n}\underline{\tilde{s}}^0_{n,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)$$

$$=p\int \frac{\lambda^2}{\left(1+\lambda\frac{\lfloor nt_1\rfloor}{n}\underline{\tilde{s}}^0_{n,t_1}(z_1)\right)\left(1+\frac{\lfloor nt_2\rfloor}{n}\lambda\underline{\tilde{s}}^0_{n,t_2}(z_2)\right)}dH_n(\lambda),$$

and consequently, we can rewrite equation (6.42) as

$$\begin{aligned}&\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_2}(z_2)\right]\mathbf{T}_n\right)\left\{1-\frac{j-1}{n}a_n(z_1,z_2,t_1,t_2)\right\}\\ =&\frac{n}{z_1z_2\underline{\tilde{s}}^0_{n,t_1}(z_1)\underline{\tilde{s}}^0_{n,t_2}(z_2)}a_n(z_1,z_2,t_1,t_2)+A_6(z_1,z_2),\end{aligned}$$

where

$$a_n(z_1,z_2,t_1,t_2)=y_n\underline{\tilde{s}}^0_{n,t_1}(z_1)\underline{\tilde{s}}^0_{n,t_2}(z_2)\int\frac{\lambda^2}{\left(1+\lambda\frac{\lfloor nt_1\rfloor}{n}\underline{\tilde{s}}^0_{n,t_1}(z_1)\right)\left(1+\lambda\frac{\lfloor nt_2\rfloor}{n}\underline{\tilde{s}}^0_{n,t_2}(z_2)\right)}dH_n(\lambda). \tag{6.43}$$

Applying Lemma 7.1.7, we have

$$\begin{aligned}&\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_2}(z_2)\right]\mathbf{T}_n\right)\\ =&\frac{n}{z_1z_2\underline{\tilde{s}}^0_{n,t_1}(z_1)\underline{\tilde{s}}^0_{n,t_2}(z_2)}a_n(z_1,z_2,t_1,t_2)\frac{1}{1-\frac{j-1}{n}a_n(z_1,z_2,t_1,t_2)}\\ &+A_6(z_1,z_2)\frac{1}{1-\frac{j-1}{n}a_n(z_1,z_2,t_1,t_2)}\\ \le&\frac{n}{z_1z_2\underline{\tilde{s}}^0_{n,t_1}(z_1)\underline{\tilde{s}}^0_{n,t_2}(z_2)}a_n(z_1,z_2,t_1,t_2)\frac{1}{1-\frac{j-1}{n}a_n(z_1,z_2,t_1,t_2)}+A_6(z_1,z_2).\end{aligned}$$

Consequently, the random variable $V_n^{(2)}$ in (6.22) can be written as

$$\begin{aligned}&V_n^{(2)}(z_1,z_2,t_1,t_2)\\ =&\frac{1}{n}a_n(z_1,z_2,t_1,t_2)\sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)}\frac{b_{j,t_1}(z_1)b_{j,t_2}(z_2)}{z_1\underline{s}^0_{n,t_1}(z_1)z_2\underline{s}^0_{n,t_2}(z_2)}\frac{1}{1-\frac{j-1}{n}a_n(z_1,z_2,t_1,t_2)}+A_7(z_1,z_2)\\ =&\frac{a_n(z_1,z_2,t_1,t_2)}{n}\sum_{j=1}^{\min(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)}\frac{1}{1-\frac{j-1}{n}a_n(z_1,z_2,t_1,t_2)}+A_7(z_1,z_2),\end{aligned}$$

where the remainder $A_7(z_1,z_2)$ may change from line to line and satisfies

$$\mathbb{E}|A_7(z_1,z_2)|\le Kn^{-\frac{1}{2}}.$$

Then, $V_n^{(2)}(z_1,z_2,t_1,t_2)$ in (6.22) converges in probability to

$$a(z_1,z_2,t_1,t_2)\int_0^{\min(t_1,t_2)}\frac{1}{1-\lambda a(z_1,z_2,t_1,t_2)}d\lambda,$$

where

$$
\begin{aligned}
&a(z_1, z_2, t_1, t_2)\\
=&y\underline{\tilde{s}}_{t_1}(z_1)\underline{\tilde{s}}_{t_2}(z_2)\int \frac{\lambda^2}{\left(1+\lambda t_1\underline{\tilde{s}}_{t_1}(z_1)\right)\left(1+\lambda t_2\underline{\tilde{s}}_{t_2}(z_2)\right)}dH(\lambda)\\
=&\frac{\underline{\tilde{s}}_{t_1}(z_1)\underline{\tilde{s}}_{t_2}(z_2)}{t_2\underline{\tilde{s}}_{t_2}(z_2)-t_1\underline{\tilde{s}}_{t_1}(z_1)}\left(y\int\frac{\lambda}{1+\lambda t_1\underline{\tilde{s}}_{t_1}(z_1)}dH(\lambda)-y\int\frac{\lambda}{1+\lambda t_2\underline{\tilde{s}}_{t_2}(z_2)}dH(\lambda)\right)\\
=&\frac{\underline{\tilde{s}}_{t_1}(z_1)\underline{\tilde{s}}_{t_2}(z_2)}{t_2\underline{\tilde{s}}_{t_2}(z_2)-t_1\underline{\tilde{s}}_{t_1}(z_1)}\left(\frac{1}{\underline{\tilde{s}}_{t_1}(z_1)}+z_1-\frac{1}{\underline{\tilde{s}}_{t_2}(z_2)}-z_2\right)\\
=&\frac{1}{t_2\underline{\tilde{s}}_{t_2}(z_2)-t_1\underline{\tilde{s}}_{t_1}(z_1)}\left(\underline{\tilde{s}}_{t_2}(z_2)+(z_1-z_2)\underline{\tilde{s}}_{t_1}(z_1)\underline{\tilde{s}}_{t_2}(z_2)-\underline{\tilde{s}}_{t_1}(z_1)\right)\\
=&\frac{\underline{\tilde{s}}_{t_2}(z_2)-\underline{\tilde{s}}_{t_1}(z_1)+(z_1-z_2)\underline{\tilde{s}}_{t_1}(z_1)\underline{\tilde{s}}_{t_2}(z_2)}{t_2\underline{\tilde{s}}_{t_2}(z_2)-t_1\underline{\tilde{s}}_{t_1}(z_1)},
\end{aligned}
$$

and we have used Lemma 7.1.6, which says

$$
y\int\frac{\lambda}{1+\lambda t\underline{\tilde{s}}_t(z)}dH(\lambda)=\frac{1}{\underline{\tilde{s}}_t(z)}+z.
$$

Finally, the limit of (6.18) is given by

$$
\sigma^2_{z_1,z_2,t_1,t_2}=\frac{\partial^2}{\partial z_1\partial z_2}\int_0^{\min(t_1,t_2)a(z_1,z_2,t_1,t_2)}\frac{1}{1-\lambda}d\lambda=\frac{\partial}{\partial z_2}\left(\frac{\min(t_1,t_2)\frac{\partial}{\partial z_1}a(z_1,z_2,t_1,t_2)}{1-\min(t_1,t_2)a(z_1,z_2,t_1,t_2)}\right)
$$

$$
=\frac{\text{numerator}}{\text{denominator}}, \tag{6.44}
$$

where

$$
\begin{aligned}
\text{numerator}=&\min(t_1,t_2)\Big\{-t_2(t_2-\min(t_1,t_2))\underline{\tilde{s}}^2_{t_2}(z_2)\underline{\tilde{s}}'_{t_1}(z_1)\Big[t_2\underline{\tilde{s}}^2_{t_2}(z_2)+(t_1-t_2)\underline{\tilde{s}}'_{t_2}(z_2)\Big]\Big\}\\
&-t_1^2\underline{\tilde{s}}^4_{t_1}(z_1)\Big\{\min(t_1,t_2)\underline{\tilde{s}}^2_{t_2}(z_2)+(t_1-\min(t_1,t_2)\underline{\tilde{s}}'_{t_2}(z_2))\Big\}\\
&+2t_1t_2\underline{\tilde{s}}^3_{t_1}(z_1)\underline{\tilde{s}}_{t_2}(z_2)\Big\{\min(t_1,t_2)\underline{\tilde{s}}^2_{t_2}(z_2)+(t_1-\min(t_1,t_2))\underline{\tilde{s}}'_{t_2}(z_2)\Big\}\\
&+2t_1t_2(t_2-\min(t_1,t_2))\underline{\tilde{s}}_{t_1}(z_1)\underline{\tilde{s}}^2_{t_2}(z_2)\underline{\tilde{s}}'_{t_1}(z_1)\Big\{\underline{\tilde{s}}_{t_2}(z_2)+(-z_1+z_2)\underline{\tilde{s}}'_{t_2}(z_2)\Big\}\\
&+\underline{\tilde{s}}^2_{t_1}(z_1)\Big\{-t_2^2\min(t_1,t_2)\underline{\tilde{s}}^4_{t_2}(z_2)+t_1(t_1-t_2)(t_1-\min(t_1,t_2))\underline{\tilde{s}}'_{t_1}(z_1)\underline{\tilde{s}}'_{t_2}(z_2)\\
&+2t_1t_2(t_1-\min(t_1,t_2))(z_1-z_2)\underline{\tilde{s}}_{t_2}(z_2)\underline{\tilde{s}}'_{t_1}(z_1)\underline{\tilde{s}}'_{t_2}(z_2)\\
&+\underline{\tilde{s}}^2_{t_2}(z_2)\Big[t_2^2(-t_1+\min(t_1,t_2))\underline{\tilde{s}}'_{t_2}(z_2)\\
&+t_1\underline{\tilde{s}}'_{t_1}(z_1)\Big(t_1(-t_2+\min(t_1,t_2))+t_2\min(t_1,t_2)(z_1-z_2)^2\underline{\tilde{s}}'_{t_2}(z_2)\Big)\Big]\Big\}\\
\text{denominator}=&\left(t_1\underline{\tilde{s}}_{t_1}(z_1)-t_2\underline{s}_{t_2}(z_2)\right)^2\Big\{(-t_2+\min(t_1,t_2))\underline{\tilde{s}}_{t_2}(z_2)\\
&+\underline{\tilde{s}}_{t_1}(z_1)(t_1-\min(t_1,t_2)+\min(t_1,t_2)(z_1-z_2)\underline{\tilde{s}}_{t_2}(z_2))\Big\}^2.
\end{aligned}
$$

Note that for the special case $t_1=t_2=1$, this covariance structure coincides with the one given in formula (9.8.4) in Bai and Silverstein (2010).

6.3.3 Proof of Theorem 6.3.2 and continuity of the limiting process

Proof of Theorem 6.3.2. We will show that the assumptions of Corollary A.4 in Dette and Tomecki (2019) are satisfied, where we identify the curve $\mathcal{C}^+$ with the compact interval $[0,1]$. For this purpose, we define the increments for the first and second coordinate of $\hat{M}_n^1$ by

$$m^1(z,t,z',z'') = \min\{|\hat{M}_n^1(z,t) - \hat{M}_n^1(z',t)|, |\hat{M}_n^1(z,t) - \hat{M}_n^1(z'',t)|\}, \tag{6.45}$$
$$m^2(z,t,t',t'') = \min\{|\hat{M}_n^1(z,t) - \hat{M}_n^1(z,t')|, |\hat{M}_n^1(z,t) - \hat{M}_n^1(z,t'')|\}, \tag{6.46}$$

where $t,t',t'' \in [t_0,1]$ and $z,z',z'' \in \mathcal{C}^+$. In order to find estimates for the tails of (6.45) and (6.46), we establish in the following lemma estimates on the moments of the increments of $\hat{M}_n^1(z,t)$, which are proved in Section 7.2. For this purpose, note that it follows from (6.15) that

$$\hat{M}_n^1(z,t_1) - \hat{M}_n^1(z,t_2) = \hat{Z}_n^1(z,t_1,t_2) + \hat{Z}_n^2(z,t_1,t_2),\ z \in \mathcal{C}^+,\ t_1,t_2 \in [t_0,1],\ t_1 \le t_2,$$

where $\hat{Z}_n^1$ and $\hat{Z}_n^2$ are the processes obtained from

$$Z_n^1(z,t_1,t_2) = \sum_{j=1}^{\lfloor nt_1 \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) \left(\beta_{j,t_2}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j - \beta_{j,t_1}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t_1}^{-2}(z)\mathbf{r}_j\right), \tag{6.47}$$
$$Z_n^2(z,t_1,t_2) = \sum_{j=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j \tag{6.48}$$

using the definition (6.4).

Lemma 6.3.4 *For $t \in [t_0,1], z_1,z_2 \in \mathcal{C}^+$, it holds for sufficiently large $n \in \mathbb{N}$ under the assumptions of Theorem 6.3.2*

$$\mathbb{E}|\hat{M}_n^1(z_1,t) - \hat{M}_n^1(z_2,t)|^{2+\delta} \le K|z_1 - z_2|^{2+\delta}, \tag{6.49}$$

where $K > 0$ is some universal constant independent of n,t,z_1,z_2. We also have for $t_1,t_2 \in [t_0,1], z \in \mathcal{C}^+$

$$\mathbb{E}|\hat{Z}_n^1(z,t_1,t_2)|^4 \le K\Big(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\Big)^4, \tag{6.50}$$
$$\mathbb{E}|\hat{Z}_n^2(z,t_1,t_2)|^{4+\delta} \le K\Big(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\Big)^{2+\delta/2}. \tag{6.51}$$

In order to simplify notation, we write $a \lesssim b$ for $a \le Kb$, where $a,b \ge 0$ and $K > 0$ denote some universal constant independent of n,t,t_1,t_2,z,z_1,z_2. We continue with the proof of Theorem 6.3.2 by using results from Lemma 6.3.4.

We observe that for $t' \le t \le t''$ and $\lambda > 0$

$$\mathbb{P}\left(m^2(z,t,t',t'') > \lambda\right)$$

$$
\begin{aligned}
&\leq \mathbb{P}\left(|\hat{M}_n^1(z,t)-\hat{M}_n^1(z,t')||\hat{M}_n^1(z,t)-\hat{M}_n^1(z,t'')|>\lambda^2\right)\\
&=\mathbb{P}\left(|\hat{Z}_n^1(z,t',t)+\hat{Z}_n^2(z,t',t)||\hat{Z}_n^1(z,t,t'')+\hat{Z}_n^2(z,t,t'')|>\lambda^2\right)\\
&\leq \mathbb{P}\left(|\hat{Z}_n^1(z,t',t)+\hat{Z}_n^2(z,t',t)|>\lambda\right)+\mathbb{P}\left(|\hat{Z}_n^1(z,t,t'')+\hat{Z}_n^2(z,t,t'')|>\lambda\right)\\
&\leq \sum_{k=1}^{2}\left\{\mathbb{P}\left(|\hat{Z}_n^k(z,t',t)|>\lambda/2\right)+\mathbb{P}\left(|\hat{Z}_n^k(z,t,t'')|>\lambda/2\right)\right\}\\
&\leq \Big(\frac{2}{\lambda}\Big)^4\mathbb{E}|\hat{Z}_n^1(z,t',t)|^4+\Big(\frac{2}{\lambda}\Big)^{4+\delta}\mathbb{E}|\hat{Z}_n^2(z,t',t)|^{4+\delta}+\Big(\frac{2}{\lambda}\Big)^4\mathbb{E}|\hat{Z}_n^1(z,t,t'')|^4\\
&\quad+\Big(\frac{2}{\lambda}\Big)^{4+\delta}\mathbb{E}|\hat{Z}_n^2(z,t,t'')|^{4+\delta}.
\end{aligned}
$$

In the case $t''-t'\geq 1/n$, we use Lemma 6.3.4 and obtain

$$
\begin{aligned}
\mathbb{E}|\hat{Z}_n^1(z,t',t)|^4 &\lesssim\Big(\frac{\lfloor nt\rfloor-\lfloor nt'\rfloor}{n}\Big)^4\lesssim\Big(t-t'+\frac{1}{n}\Big)^4\leq\Big(t''-t'+\frac{1}{n}\Big)^4\leq 2^4(t''-t')^4\\
&\lesssim(t''-t')^4,\\
\mathbb{E}|\hat{Z}_n^2(z,t,t'')|^{4+\delta} &\lesssim\Big(\frac{\lfloor nt''\rfloor-\lfloor nt\rfloor}{n}\Big)^{2+\delta/2}\lesssim\Big(t''-t+\frac{1}{n}\Big)^{2+\delta/2}\leq\Big(t''-t'+\frac{1}{n}\Big)^{2+\delta/2}\\
&\leq 2^{2+\delta/2}(t''-t')^{2+\delta/2}\lesssim(t''-t')^{2+\delta/2}.
\end{aligned}
$$

The remaining terms can be treated similarly in this case, which gives

$$
\mathbb{P}\left(m^2(z,t,t',t'')>\lambda\right)\lesssim\max(\lambda^{-4},\lambda^{-(4+\delta)})(t''-t')^{2+\delta/2}
$$

for $t''-t'\geq 1/n$. In the other case $t''-t'<1/n$, we have $\lfloor nt\rfloor=\lfloor nt''\rfloor$ or $\lfloor nt\rfloor=\lfloor nt'\rfloor$ and consequently,

$$
\hat{M}_n^1(z,t)-\hat{M}_n^1(z,t')=0 \text{ or } \hat{M}_n^1(z,t'')-\hat{M}_n^1(z,t)=0.
$$

Therefore we obtain for $t'\leq t\leq t''\leq 1$

$$
\mathbb{P}\left(m^2(z,t,t',t'')>\lambda\right)\lesssim\max(\lambda^{-4},\lambda^{-(4+\delta)})(t''-t')^{2+\delta/2}.
$$

For the following analysis, we identify $\mathcal{C}^+$ with $[0,1]$. In order to derive a similar estimate for the term m^1, we note that it follows for $z,z',z''\in[0,1]$ such that $z'\leq z\leq z''$

$$
\begin{aligned}
\mathbb{P}\left(m^1(z,t,z',z'')>\lambda\right)&\leq\mathbb{P}\left(|\hat{M}_n^1(z,t)-\hat{M}_n^1(z',t)||\hat{M}_n^1(z,t)-\hat{M}_n^1(z'',t)|>\lambda^2\right)\\
&\leq\lambda^{-(2+\delta)}\mathbb{E}[|\hat{M}_n^1(z,t)-\hat{M}_n^1(z',t)||\hat{M}_n^1(z,t)-\hat{M}_n^1(z'',t)|]^{1+\delta/2}\\
&\leq\lambda^{-(2+\delta)}\left(\mathbb{E}|\hat{M}_n^1(z,t)-\hat{M}_n^1(z',t)|^{2+\delta}\mathbb{E}|\hat{M}_n^1(z,t)-\hat{M}_n^1(z'',t)|^{2+\delta}\right)^{1/2}\\
&\lesssim\lambda^{-(2+\delta)}\left(|z-z'|^{2+\delta}|z-z''|^{2+\delta}\right)^{1/2}\leq\lambda^{-(2+\delta)}|z'-z''|^{2+\delta},
\end{aligned}
$$

where we used Lemma 6.3.4 in the last line. Moreover, we have

$$
\mathbb{P}\left(|\hat{M}_n^1(z_1,t_1)-\hat{M}_n^1(z_2,t_2)|>\lambda\right)
$$

$$
\begin{aligned}
&\leq \mathbb{P}\left(|\hat{M}_n^1(z_1,t_1)-\hat{M}_n^1(z_2,t_1)|>\frac{\lambda}{2}\right)+\mathbb{P}\left(|\hat{M}_n^1(z_2,t_1)-\hat{M}_n^1(z_2,t_2)|>\frac{\lambda}{2}\right)\\
&\leq \mathbb{P}\Big(|\hat{M}_n^1(z_1,t_1)-\hat{M}_n^1(z_2,t_1)|>\frac{\lambda}{2}\Big)+\sum_{k=1}^{2}\mathbb{P}\Big(|\hat{Z}_n^k(z_2,t_1,t_2)|>\frac{\lambda}{4}\Big)\\
&\leq \Big(\frac{2}{\lambda}\Big)^{2+\delta}\mathbb{E}|\hat{M}_n^1(z_1,t_1)-\hat{M}_n^1(z_2,t_1)|^{2+\delta}+\Big(\frac{4}{\lambda}\Big)^{4}\mathbb{E}|\hat{Z}_n^1(z_2,t_1,t_2)|^4+\Big(\frac{4}{\lambda}\Big)^{4+\delta}\mathbb{E}|\hat{Z}_n^2(z_2,t_1,t_2)|^{4+\delta}\\
&\lesssim \Big(\frac{2}{\lambda}\Big)^{2+\delta}|z_1-z_2|^{2+\delta}+\Big(\frac{4}{\lambda}\Big)^{4}\Big(\frac{\lfloor nt_2\rfloor-\lfloor nt_1\rfloor}{n}\Big)^4+\Big(\frac{4}{\lambda}\Big)^{4+\delta}\Big(\frac{\lfloor nt_2\rfloor-\lfloor nt_1\rfloor}{n}\Big)^{2+\delta/2}\\
&\lesssim C_{1,\lambda}\Big[\Big|\frac{\lfloor nt_2\rfloor-\lfloor nt_1\rfloor}{n}\Big|^{2+\delta/2}+|z_1-z_2|^{2+\delta}\Big]\\
&\leq C_{1,\lambda}\Big[\Big(|t_2-t_1|+\frac{1}{n}\Big)^{2+\delta/2}+|z_1-z_2|^{2+\delta}\Big]\\
&\leq C_{1,\lambda}\Big(\Big\|(z_1,t_1)^\top-(z_2,t_2)^\top\Big\|_\infty+\frac{1}{n}\Big)^{2+\delta/2},
\end{aligned}
$$

where

$$C_{1,\lambda}=\max(\lambda^{-4},\lambda^{-(2+\delta)},\lambda^{-(4+\delta)}).$$

Let $m\in\mathbb{N}$ and define for $j=(j_1,j_2)\in\{1,\dots,m\}^2$ the set

$$K_j=\Big[\frac{j_1-1}{m},\frac{j_1}{m}\Big]\times\Big[\frac{j_2-1}{m}\wedge t_0,\frac{j_2}{m}\wedge t_0\Big].$$

Combining the three inequalities above, we are able to apply Corollary A.4 in Dette and Tomecki (2019) with the parameters $\varepsilon=1/m,\delta'=2+\delta/2$ and get

$$\mathbb{P}\Big(\sup_{(z_1,t_1),(z_2,t_2)\in K_j}|\hat{M}_n^1(z_1,t_1)-\hat{M}_n^2(z_2,t_2)|>\lambda\Big)\lesssim C_{2,\lambda}\Big(\frac{1}{m}\Big)^{2+\delta/2}+C_{1,\lambda}\Big(\frac{1}{m}+\frac{1}{n}\Big)^{2+\delta/2},$$

where $C_{2,\lambda}=\max(\lambda^{-4},\lambda^{-(4+\delta)},\lambda^{-(2+\delta)})$. This implies

$$
\begin{aligned}
&\limsup_{n\to\infty}\mathbb{P}\Big(\sup_{j\in\{1,\dots,m\}^2}\sup_{(z_1,t_1),(z_2,t_2)\in K_j}|\hat{M}_n^1(z_1,t_1)-\hat{M}_n^2(z_2,t_2)|>\lambda\Big)\\
\leq&\limsup_{n\to\infty}\sum_{j\in\{1,\dots,m\}^2}\mathbb{P}\Big(\sup_{(z_1,t_1),(z_2,t_2)\in K_j}|\hat{M}_n^1(z_1,t_1)-\hat{M}_n^2(z_2,t_2)|>\lambda\Big)\\
\lesssim&\limsup_{n\to\infty}m^2\Big[C_{2,\lambda}\Big(\frac{1}{m}\Big)^{2+\delta/2}+C_{1,\lambda}\Big(\frac{1}{m}+\frac{1}{n}\Big)^{2+\delta/2}\Big]\\
\lesssim&\, m^2\frac{1}{m^{2+\delta/2}}\to 0,\text{ as } m\to\infty.
\end{aligned}
$$

Theorem 1.5.7 in Van Der Vaart and Wellner (1996) finally implies the asymptotic tightness of the sequence $(\hat{M}_n^1)_{n\in\mathbb{N}}$, which completes the proof of Theorem 6.3.2. □

Corollary 6.3.5 *There exists a version of the process $(M^1(z,t))_{z\in\mathcal{C}^+,t\in[t_0,1]}$ with continuous sample paths.*

Proof. By Addendum 1.5.8 in Van Der Vaart and Wellner (1996), almost all paths $(z,t,\omega)\in(\mathcal{C}^+\setminus\{x_l,x_r\})\times[t_0,1]\times\Omega\mapsto\hat{M}^1(z,t)(\omega)$ are continuous. Since $(\mathcal{C}^+\setminus\{x_l,x_r\})\times[t_0,1]\subset\mathcal{C}^+\times[t_0,1]$ is a dense set, we conclude that almost all paths $(z,t,\omega)\in\mathcal{C}^+\times[t_0,1]\times\Omega\mapsto\hat{M}^1(z,t)(\omega)$ are continuous. □

6.3.4 Proof of Theorem 6.3.3

Recall that

$$\underline{\tilde{s}}_{n,t}(z) = s_{F^{\underline{\mathbf{B}}_{n,t}}}(z) = -\frac{1-y_{\lfloor nt \rfloor}}{z} + y_{\lfloor nt \rfloor}\tilde{s}_{n,t}(z)$$

and

$$\underline{\tilde{s}}^0_{n,t}(z) = s_{\underline{\tilde{F}}^{y_{\lfloor nt \rfloor},H_n}(z)} = -\frac{1-y_{\lfloor nt \rfloor}}{z} + y_{\lfloor nt \rfloor}\tilde{s}^0_{n,t}(z),$$

where $F^{\underline{\mathbf{B}}_{n,t}}$ denotes the empirical spectral distribution of the matrix $\underline{\mathbf{B}}_{n,t}$ defined in (2.2) and the distribution $\underline{\tilde{F}}^{y_{\lfloor nt \rfloor},H_n}$ is defined in (3.4).

Recalling the definition (6.11) we have

$$M_n^2(z,t) = p\left(\mathbb{E}[\tilde{s}_{n,t}(z)] - \tilde{s}^0_{n,t}(z)\right) = \lfloor nt \rfloor \left(\mathbb{E}\left[\underline{\tilde{s}}_{n,t}(z)\right] - \underline{\tilde{s}}^0_{n,t}(z)\right). \tag{6.52}$$

We begin with a lemma, which is used to derive an alternative representation of $M_n^2(z,t)$. Note that this Lemma corrects an error in formula (9.11.1) in Bai and Silverstein (2010) and is proved in Section 7.3.

Lemma 6.3.6 *It holds*

$$\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] - \underline{\tilde{s}}^0_{n,t}(z)\right)\left(1 - \frac{y_n\frac{\lfloor nt \rfloor}{n}\int\frac{\lambda^2\underline{\tilde{s}}^0_{n,t}(z)dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt \rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z))}}{-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]} - \frac{\lfloor nt \rfloor}{n}R_{n,t}(z)}\right)$$
$$=\frac{\lfloor nt \rfloor}{n}R_{n,t}(z)\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\underline{\tilde{s}}^0_{n,t}(z),$$

where

$$\begin{aligned}
R_{n,t}(z) &= y_{\lfloor nt \rfloor}\lfloor nt \rfloor^{-1}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)^{-1} \\
&= y_{\lfloor nt \rfloor}n^{-1}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\Big(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\Big)^{-1}, \\
d_{j,t}(z) &= -\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j \\
&\quad + \frac{1}{p}\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_t^{-1}(z)\Big], \\
\mathbf{q}_j &= \frac{1}{\sqrt{p}}\mathbf{x}_j.
\end{aligned}$$

The next main step is the following result, which is proved in Section 7.4.

Theorem 6.3.7 *Under the assumptions of Theorem 3.2.1, we have*

$$\lim_{n\to\infty}\sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}}\left|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] - \underline{\tilde{s}}_t(z)\right| = 0,$$

where $\underline{\tilde{s}}_t$ is defined in (3.1).

The third step in the proof of Theorem 6.3.3 is the following result, which is proved in Section 7.5.

Theorem 6.3.8 *Under the assumptions of Theorem 3.2.1, we have*

$$\sup_{\substack{n\in\mathbb{N},\\ z\in\mathcal{C}_n,\\ t\in[t_0,1]}} |M_n^2(z,t)| \leq K \ , \qquad \lim_{n\to\infty} \sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} |\underline{\tilde{s}}_{n,t}^0(z) - \underline{\tilde{s}}_t(z)| = 0.$$

Using Lemma 6.3.6 and Theorem 6.3.7, we show in Section 7.6 that

$$\lfloor nt \rfloor R_{n,t}(z)\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \to \begin{cases} \frac{y\int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^3}dH(\lambda)}{1-ty\int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^2}dH(\lambda)} & \text{for the real case,} \\ 0 & \text{for the complex case,} \end{cases} \tag{6.53}$$

uniformly with respect to $z \in \mathcal{C}_n, t \in [t_0, 1]$. Combining this result with Theorem 6.3.7 and Lemma 7.7.6 yields

$$\lim_{n\to\infty} \sup_{\substack{z\in\mathcal{C}_n\\ t\in[t_0,1]}} |R_{n,t}(z)| = 0 \tag{6.54}$$

This result and Lemma 7.7.6, Theorem 6.3.7, Theorem 6.3.8, Lemma 7.7.2 and the equation (3.2) show that

$$\frac{y_n \frac{\lfloor nt\rfloor}{n} \int \frac{\lambda^2 \underline{\tilde{s}}_{n,t}^0(z)dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z))}}{-z + y_n\int \frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]} - R_{n,t}(z)} \to ty\int \frac{\lambda^2\underline{\tilde{s}}_t^2(z)dH(\lambda)}{(1+\lambda t\underline{\tilde{s}}_t(z))^2} \ .$$

Observing the representation in (6.52), Lemma 6.3.6 and Theorem 6.3.8, this implies

$$M_n^2(z,t) \to \begin{cases} \frac{ty\int \frac{\underline{\tilde{s}}_t^3(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^3}dH(\lambda)}{\left(1-ty\int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^2}dH(\lambda)\right)^2} & \text{for the real case,} \\ 0 & \text{for the complex case.} \end{cases}$$

uniformly with respect $z \in \mathcal{C}_n, t \in [t_0, 1]$, which completes the proof of Theorem 6.3.3.

6.4 Details on the arguments in Section 6.2

Lemma 6.4.1 *Let* Γ_F *denote the support of a c.d.f.* F. *Then it holds*

$$\Gamma_{\tilde{F}^{y_{\lfloor nt\rfloor},H_n}} \subset \left[\frac{\lfloor nt_0\rfloor}{n}\lambda_p(\mathbf{T}_n)I_{(0,1)}(y_{\lfloor nt_0\rfloor})(1-\sqrt{y_{\lfloor nt_0\rfloor}})^2, \lambda_1(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt_0\rfloor}})^2\right],$$

where $\tilde{F}^{y_{\lfloor nt\rfloor},H_n}$ *is defined in* (3.6).

Proof of Lemma 6.4.1. By definition, we have

$$\tilde{F}^{y_{\lfloor nt \rfloor}, H_n}(\cdot) = F^{y_{\lfloor nt \rfloor}, H_n}\left(\frac{n}{\lfloor nt \rfloor}\cdot\right),$$

and thus,

$$\Gamma_{\tilde{F}^{y_{\lfloor nt \rfloor}, H_n}} = \frac{\lfloor nt \rfloor}{n}\Gamma_{F^{y_{\lfloor nt \rfloor}, H_n}}.$$

Since

$$\underline{F}^{y_{\lfloor nt \rfloor}, H_n} - y_{\lfloor nt \rfloor}F^{y_{\lfloor nt \rfloor}, H_n} = (1 - y_{\lfloor nt \rfloor})I_{[0,\infty)},$$

it is sufficient to investigate the support of $\underline{F}^{y_{\lfloor nt \rfloor}, H_n}$ in order to study the support of $F^{y_{\lfloor nt \rfloor}, H_n}$. Define for all $\alpha \notin \Gamma_{H_n}$

$$\psi_{y_{\lfloor nt \rfloor}, H_n}(\alpha) = \psi(\alpha) = \alpha + y_{\lfloor nt \rfloor}\alpha \int \frac{\lambda}{\alpha - \lambda} dH_n(\lambda) = \alpha + \alpha \frac{y_{\lfloor nt \rfloor}}{p} \sum_{i=1}^{p} \frac{\lambda_i(\mathbf{T}_n)}{\alpha - \lambda_i(\mathbf{T}_n)}.$$

We need the following result [see, e.g., Lemma 6.1, Bai and Silverstein (2010) or Proposition 2.17, Yao et al. (2015)].

Proposition 6.4.2 *If $\lambda \notin \Gamma_{\underline{F}^{y_{\lfloor nt \rfloor}, H_n}}$, then $\underline{s}_{n,t}(\lambda) \neq 0$ and $\alpha = -1/\underline{s}_{n,t}(\lambda)$ satisfies*

1. *$\alpha \notin \Gamma_{H_n}$ and $\alpha \neq 0$,*
2. *$\psi'(\alpha) > 0$.*

Conversely, if α satisfies 1-2, then $\lambda = \psi(\alpha) \notin \Gamma_{\underline{F}^{y_{\lfloor nt \rfloor}, H_n}}$.

That is, we can determine a superset of the support of $\underline{F}^{y_{\lfloor nt \rfloor}, H_n}$ by considering the complement of intervals where ψ is monotonously increasing. The derivative of ψ is given by

$$\psi'(\alpha) = 1 - \frac{y_{\lfloor nt \rfloor}}{p} \sum_{i=1}^{p} \frac{(\lambda_i(\mathbf{T}_n))^2}{(\alpha - \lambda_i(\mathbf{T}_n))^2}.$$

First, consider the case $\alpha < \lambda_p(\mathbf{T}_n)$, which implies $\alpha \notin \Gamma_{H_n}$. Noting that $f_\alpha(x) = f(x) = x^2/(x-\alpha)^2$ is a monotone decreasing function for $x > \alpha > 0$, we have to solve the following inequality for $0 < \alpha < \lambda_p(\mathbf{T}_n)$

$$p\frac{(\lambda_p(\mathbf{T}_n))^2}{(\alpha - \lambda_p(\mathbf{T}_n))^2} < \frac{p}{y_{\lfloor nt \rfloor}},$$

which is in this case a sufficient condition for $\psi'(\alpha) > 0$. This gives for $y_{\lfloor nt \rfloor} \in (0,1)$

$$0 < \alpha < \lambda_p(\mathbf{T}_n)(1 - \sqrt{y_{\lfloor nt \rfloor}}).$$

If $\alpha < 0 < \lambda_p(\mathbf{T}_n)$, we conclude

$$\psi'(\alpha) > 1 - y_{\lfloor nt \rfloor},$$

which is non-negative if and only if $y_{\lfloor nt \rfloor} \geq 1$. In this case, we conclude $\psi(\alpha) \leq 0$.

Now we calculate for $y_{\lfloor nt \rfloor} \in (0,1)$

$$\begin{aligned}\psi(\alpha) \leq & \psi\left(\lambda_p(\mathbf{T}_n)\left(1 - \sqrt{y_{\lfloor nt \rfloor}}\right)\right) \\ & \leq \lambda_p(\mathbf{T}_n)\left(1 - \sqrt{y_{\lfloor nt \rfloor}}\right) + y_{\lfloor nt \rfloor}\lambda_p(\mathbf{T}_n)\left(1 - \sqrt{y_{\lfloor nt \rfloor}}\right)\frac{\lambda_p(\mathbf{T}_n)}{-\sqrt{y_{\lfloor nt \rfloor}}\lambda_p(\mathbf{T}_n)} \\ & = \lambda_p(\mathbf{T}_n)\left(1 - \sqrt{y_{\lfloor nt \rfloor}}\right)^2.\end{aligned}$$

Using Proposition 6.4.2, this gives

$$\Gamma_{F^{y_{\lfloor nt \rfloor}, H_n}} \subset \left[\lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt_0 \rfloor})(1 - \sqrt{y_{\lfloor nt_0 \rfloor}})^2, \infty\right).$$

Next, we specify the right endpoint of this interval. Let $\alpha > \lambda_1(\mathbf{T}_n)$ which implies $\alpha \notin \Gamma_{H_n}$. Note that the function f is monotone increasing for $0 < x < \alpha$. Solving the inequality

$$p\frac{(\lambda_1(\mathbf{T}_n))^2}{(\alpha - \lambda_1(\mathbf{T}_n))^2} < \frac{p}{y_{\lfloor nt \rfloor}},$$

which is in this case a sufficient condition for $\psi'(\alpha) > 0$, gives

$$\alpha > \lambda_1(\mathbf{T}_n)\left(1 + \sqrt{y_{\lfloor nt \rfloor}}\right).$$

Similar consideration as in the other case yield

$$\psi\left(\lambda_1(\mathbf{T}_n)\left(1 + \sqrt{y_{\lfloor nt \rfloor}}\right)\right) \geq \lambda_1(\mathbf{T}_n)\left(1 + \sqrt{y_{\lfloor nt \rfloor}}\right)^2.$$

This finishes the proof of Lemma 6.4.1. □

The following lemma ensures that the process $(\hat{M}_n(z,t))_{z \in \mathcal{C}^+, t \in [t_0,1]}$ defined in (6.4) provides an appropriate approximation for the process $(M_n(z,t))_{z \in \mathcal{C}^+, t \in [t_0,1]}$.

Lemma 6.4.3 *Let $i \in \{1,2\}$. It holds with probability 1 (uniformly in $t \in [t_0,1]$)*

$$\left| \int_{\mathcal{C}} f_i(z)\left(M_n(z,t) - \hat{M}_n(z,t)\right) dz \right| = o(1), \text{ as } n \to \infty, \ i = 1,2.$$

Proof of Lemma 6.4.3. For convenience, we write $f_i = f$. Since $\mathcal{C} = \mathcal{C}^+ \cup \overline{\mathcal{C}^+}$ and $M_n(\overline{z},t) = \overline{M_n(z,t)}$ for all $z = x + iv \in \mathcal{C}^+$, we have (using also the definition of $\hat{M}_n$ in (6.4))

$$\left| \int_{\mathcal{C}} f(z)\left(M_n(z,t) - \hat{M}_n(z,t)\right) dz \right| \leq K \int_{[0,n^{-1}\varepsilon_n]} \left\{ |M_n(x_r + iv, t) - M_n(x_r + in^{-1}\varepsilon_n, t)| \right.$$

$$+ |M_n(x_l + iv, t) - M_n(x_l + in^{-1}\varepsilon_n, t)|\Big\} dv.$$

Let Γ_F denote the support of a c.d.f. F, then it follows by Proposition 2.4 in Yao et al. (2015) that

$$|s_F(z)| \leq \frac{1}{\operatorname{dist}(z, \Gamma_F)}, \tag{6.55}$$

where $z \in \mathbb{C} \setminus \Gamma_F$ and s_F is the Stieltjes transform of F. Using (6.8) and Lemma 6.4.1, we have for $v \in [0, n^{-1}\varepsilon_n]$ and sufficiently large n

$$\begin{aligned}
\operatorname{dist}(x_r + iv, \Gamma_{F^{\mathbf{B}_{n,t}}}) &\geq \big|x_r - \lambda_1(\mathbf{B}_{n,t})\big| \\
&\geq \big|x_r - \max\big(\lambda_1(\mathbf{B}_{n,t}), \lambda_1(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt \rfloor}})^2\big)\big|, \\
\operatorname{dist}\big(x_l + iv, \Gamma_{\tilde{F}^{y_{\lfloor nt \rfloor}, H_n}}\big) &\geq \Big|x_l - \frac{\lfloor nt_0 \rfloor}{n} \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt \rfloor})(1-\sqrt{y_{\lfloor nt \rfloor}})^2\Big| \\
&\geq \Big|x_l - \min\Big(\lambda_p(\mathbf{B}_{n,t}), \frac{\lfloor nt_0 \rfloor}{n} \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt \rfloor})(1-\sqrt{y_{\lfloor nt \rfloor}})^2\Big)\Big|.
\end{aligned}$$

Similarly, one can show that for sufficiently large n

$$\begin{aligned}
\operatorname{dist}\big(x_r + iv, \Gamma_{\tilde{F}^{y_{\lfloor nt \rfloor}, H_n}}\big) &\geq \big|x_r - \max\big(\lambda_1(\mathbf{B}_{n,t}), \lambda_1(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt \rfloor}})^2\big)\big|, \\
\operatorname{dist}(x_l + iv, \Gamma_{F^{\mathbf{B}_{n,t}}}) &\geq \Big|x_l - \min\Big(\lambda_p(\mathbf{B}_{n,t}), \frac{\lfloor nt_0 \rfloor}{n} \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt \rfloor})(1-\sqrt{y_{\lfloor nt \rfloor}})^2\Big)\Big|.
\end{aligned}$$

Recall the definition of M_n, then (6.55) implies

$$\begin{aligned}
&\Big|\int_{\mathcal{C}} f(z)\Big(M_n(z,t) - \hat{M}_n(z,t)\Big)\, dz\Big| \\
\leq & 4K\varepsilon_n \Big\{ \big|x_r - \max\big(\lambda_1(\mathbf{B}_{n,t}), \lambda_1(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt \rfloor}})^2\big)\big|^{-1} \\
&+ \Big|x_l - \min\Big(\lambda_p(\mathbf{B}_{n,t}), \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt \rfloor}) \frac{\lfloor nt_0 \rfloor}{n}(1-\sqrt{y_{\lfloor nt \rfloor}})^2\Big)\Big|^{-1}\Big\}.
\end{aligned}$$

Due to (6.8), for every $t \in [t_0, 1]$, the denominators are bounded away from 0 for sufficiently large n with probability 1 (nullset may depend on t). Note that for every $n \in \mathbb{N}$, there are only finitely many $t_1, t_2 \in [t_0, 1]$ such that $\lfloor nt_1 \rfloor \neq \lfloor nt_2 \rfloor$. That is, since the countable union of nullsets is again a nullset, we find that with probability 1 (uniformly in t)

$$\begin{aligned}
&\limsup_{n\to\infty} \Big|\int_{\mathcal{C}} f(z)\Big(M_n(z,t) - \hat{M}_n(z,t)\Big)\, dz\Big| \\
&\leq 4K \lim_{n\to\infty} \varepsilon_n \Big\{ \Big(x_r - \limsup_{n\to\infty} \max\big(\lambda_1(\mathbf{B}_{n,t}), \lambda_1(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt \rfloor}})^2\big)\Big)^{-1} \\
&+ \Big(\liminf_{n\to\infty} \min\Big(\lambda_p(\mathbf{B}_{n,t}), \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt \rfloor}) \frac{\lfloor nt_0 \rfloor}{n}(1-\sqrt{y_{\lfloor nt \rfloor}})^2\Big) - x_l\Big)^{-1}\Big\} \\
&\leq 4K \lim_{n\to\infty} \varepsilon_n \Big\{ \Big(x_r - \limsup_{n\to\infty} \lambda_1(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt_0 \rfloor}})^2\Big)^{-1} \\
&+ \Big(\liminf_{n\to\infty} \lambda_p(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt_0 \rfloor}) \frac{\lfloor nt_0 \rfloor}{n}(1-\sqrt{y_{\lfloor nt_0 \rfloor}})^2 - x_l\Big)^{-1}\Big\} = 0.
\end{aligned}$$

□

Chapter 7

More details on the proof of Theorem 6.1.1

In this chapter, we provide the remaining arguments in the proof of Theorem 6.1.1 in Section 6.3. Several further very technical results are given in Section 7.7.

7.1 Auxiliary results for the proof of Theorem 6.3.1 in Section 6.3.2

Lemma 7.1.1 *$Z_{nj}^{t_1,t_2}$ defined in* (6.16) *forms a martingale difference scheme with respect to the filtration $\mathcal{F}_{nj} = \sigma(\{\mathbf{r}_1, ..., \mathbf{r}_j\})$ $(1 \leq j \leq \max(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor))$.*

Proof of Lemma 7.1.1. Obviously, $Z_{nj}^{t_1,t_2}$ is $\sigma(\mathbf{r}_1, ..., \mathbf{r}_j)$-measurable by the definition of $\mathbb{E}_j$. It remains to show that

$$\mathbb{E}_{j-1}[Z_{nj}^{t_1,t_2}] = 0.$$

By the tower property of the expected value, it suffices to prove

$$\mathbb{E}_{j-1}\left(\overline{\beta}_{j,t}(z)\alpha_{j,t}(z) - \overline{\beta}_{j,t}^2(z)\hat{\gamma}_{j,t}(z)\frac{1}{n}\operatorname{tr}(\mathbf{T}_n\mathbf{D}_{j,t}^{-2}(z))\right) = 0. \tag{7.1}$$

Considering the first summand, we have

$$\begin{aligned}
&\mathbb{E}_{j-1}\left[\frac{1}{1+n^{-1}\operatorname{tr}(\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z))}\left(\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-2}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}(\mathbf{D}_{j,t}^{-2}(z)\mathbf{T}_n)\right)\right] \\
&\quad = \mathbb{E}_{j-1}\left[\frac{1}{1+n^{-1}\operatorname{tr}(\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z))}\left(n^{-1}\operatorname{tr}(\mathbf{D}_{j,t}^{-2}(z)\mathbf{T}_n) - n^{-1}\operatorname{tr}(\mathbf{D}_{j,t}^{-2}(z)\mathbf{T}_n)\right)\right] = 0,
\end{aligned}$$

where we have used the fact that x_j is independent of $\mathbf{r}_1, ..., \mathbf{r}_{j-1}$, $\mathbf{D}_{j,t}$ and $\mathbb{E}[x_j x_j^\star] = \mathbf{I}$. The assertion for the second summand follows in a similar fashion, which yields the desired MDS property. □

Lemma 7.1.2 *The scheme $Z_{nj}^{t_1,t_2}$ satisfies condition (5.31) given in Lemma 5.6 of Najim and Yao (2016), that is,*

$$\lim_{n\to\infty} \sum_{j=1}^{\max(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} \mathbb{E}\left(\left|Z_{nj}^{t_1,t_2}\right|^2 I\left(\left|Z_{nj}^{t_1,t_2}\right| > \varepsilon\right)\right) = 0.$$

Proof of Lemma 7.1.2. Note that we have for any $q>0$ and any $z\in\mathbb{C}$ with $\operatorname{Im}(z)\neq 0$

$$\max\left(\left|\left|\mathbf{D}_t^{-1}(z)\right|\right|^q, \left|\left|\mathbf{D}_{j,t}^{-1}(z)\right|\right|^q, \left|\left|\mathbf{D}_{i,j,t}^{-1}(z)\right|\right|^q\right) \le K_{q,z}. \tag{7.2}$$

where $K_{q,z}>0$ is allowed to depend on q and z. Later on, we will strengthen this result and find a bound for the moments of these random variables uniformly in $z\in\mathcal{C}_n$ and $t\in[t_0,1]$ (see Lemma 7.7.3). It holds

$$\mathbb{E}|Y_{j,t}(z)|^4 \le K\left(\mathbb{E}|\alpha_{j,t}(z)|^4 + \mathbb{E}|\hat{\gamma}_{j,t}(z)|^4\right) = o(n^{-1}).$$

where we used (9.9.6) in Bai and Silverstein (2010) in the following way (exemplarily for the second term):

$$\begin{aligned}
\mathbb{E}|\hat{\gamma}_{j,t}(z)|^4 &= \mathbb{E}\left|\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}(\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n)\right|^4\\
&= \mathbb{E}^{(-j)}\left[\mathbb{E}^{(j)}\left|\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}(\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n)\right|^4\right]\\
&\le Kn^{-1}\eta_n^4\mathbb{E}^{(-j)}\left[||\mathbf{D}_{j,t}^{-1}(z)||^4\right] = Kn^{-1}\eta_n^4\mathbb{E}\left[||\mathbf{D}_{j,t}^{-1}(z)||^4\right] = o(n^{-1}). \quad (7.3)
\end{aligned}$$

Here, $\mathbb{E}^{(-j)}$ denotes the expected value with respect to $\{\mathbf{r}_1,...,\mathbf{r}_n\}\setminus\{\mathbf{r}_j\}$ and $\mathbb{E}^{(j)}$ the expected value with respect to $\mathbf{r}_j$ and we used the independence of $\mathbf{r}_1,...,\mathbf{r}_n$. The fourth absolute moment of $\alpha_{j,t}(z)$ can be controlled similarly. This implies the Lindeberg-type condition

$$\begin{aligned}
\sum_{j=1}^{\max(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} \mathbb{E}\left(\left|Z_{nj}^{t_1,t_2}\right|^2 I\left(\left|Z_{nj}^{t_1,t_2}\right| > \varepsilon\right)\right) &\le \frac{1}{\varepsilon^2}\sum_{j=1}^{\max(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} \mathbb{E}\left|Z_{nj}^{t_1,t_2}\right|^4\\
&= \frac{1}{\varepsilon^2}\sum_{j=1}^{\max(\lfloor nt_1\rfloor,\lfloor nt_2\rfloor)} \mathbb{E}\left|\sum_{i=1}^k \alpha_{i,1}Y_{j,t_1}(z_i) + \alpha_{i,2}Y_{j,t_2}(z_i)\right|^4 \to 0,
\end{aligned}$$

as $n\to\infty$. □

Lemma 7.1.3 *It holds for $t\in[0,1]$, $z\in\mathbb{C}$ such that Im$(z)>0$*

$$\underline{\tilde{s}}_{n,t}(z) = -\frac{1}{z\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\beta_{j,t}(z).$$

Proof (similar to formula (6.2.4) in Bai and Silverstein (2010)). Note that

$$\mathbf{D}_t(z) + z\mathbf{I} = \sum_{j=1}^{\lfloor nt\rfloor}\mathbf{r}_j\mathbf{r}_j^\star.$$

Multiplying by $\mathbf{D}_t^{-1}(z)$ and using the identity (6.1.11) from Bai and Silverstein (2010)

$$\mathbf{r}_j^\star(\mathbf{C}+\mathbf{r}_j\mathbf{r}_j^\star)^{-1} = \frac{1}{1+\mathbf{r}_j^\star\mathbf{C}^{-1}\mathbf{r}_j}\mathbf{r}_j^\star\mathbf{C}^{-1} \tag{7.4}$$

for any $p \times p$ matrix $\mathbf{C}$ such that $\mathbf{C}+\mathbf{r}_j^\star\mathbf{r}_j$ is invertible, we conclude that

$$\begin{aligned}\mathbf{I}+z\mathbf{D}_t^{-1}(z) &= \sum_{j=1}^{\lfloor nt\rfloor}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_t^{-1}(z) = \sum_{j=1}^{\lfloor nt\rfloor}\mathbf{r}_j\mathbf{r}_j^\star\left(\mathbf{D}_{j,t}(z)+\mathbf{r}_j\mathbf{r}_j^\star\right)^{-1}\\ &= \sum_{j=1}^{\lfloor nt\rfloor}\frac{1}{1+\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z).\end{aligned}$$

Applying the trace on both sides, dividing by $\lfloor nt\rfloor$ and noting that

$$\tilde{s}_{n,t}(z) = \frac{1}{p}\operatorname{tr}(\mathbf{D}_t^{-1}(z)),$$

we get

$$\begin{aligned}y_{\lfloor nt\rfloor}+zy_{\lfloor nt\rfloor}\tilde{s}_{n,t}(z) &= \sum_{j=1}^{\lfloor nt\rfloor}\frac{1}{\lfloor nt\rfloor}\frac{\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j}{1+\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j}\\ &= 1-\frac{1}{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\frac{1}{1+\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j}\\ &= 1-\frac{1}{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\beta_{j,t}(z).\end{aligned}$$

Recalling (2.3), we see that

$$\tilde{s}_{n,t}(z) = \frac{1}{zy_{\lfloor nt\rfloor}}-\frac{1}{z}+\frac{1}{y_{\lfloor nt\rfloor}}\tilde{\underline{s}}_{n,t}(z),$$

which finally implies

$$\tilde{\underline{s}}_{n,t}(z) = -\frac{1}{\lfloor nt\rfloor z}\sum_{j=1}^{\lfloor nt\rfloor}\beta_{j,t}(z).$$

□

Lemma 7.1.4

$$\begin{aligned}&\left|\frac{1}{z_1z_2}\operatorname{tr}\left(\left(\mathbf{I}+\frac{\lfloor nt_2\rfloor}{n}\underline{s}_{n,t_2}^0(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I}+\frac{\lfloor nt_1\rfloor}{n}\underline{s}_{n,t_1}^0(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\right.\\ &\left.\quad-\operatorname{tr}\left(\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\right|\\ &\leq Kn^{\frac{1}{2}}.\end{aligned}$$

Proof. We use (6.40) to get

$$\left|\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)+z\frac{\lfloor nt\rfloor}{n}\underline{s}^0_{n,t}(z)\right|\leq\left|\frac{1}{n}b_{j,t}(z)\right|+\left|b_{j,t}(z)+z\underline{s}^0_{n,t}(z)\right|\leq Kn^{-\frac{1}{2}}.$$

and make the following consideration for invertible $p\times p$ matrices $\mathbf{A}$ and $\mathbf{B}$:

$$||\mathbf{A}^{-1}-\mathbf{B}^{-1}||=||\mathbf{B}^{-1}(\mathbf{B}-\mathbf{A})\mathbf{A}^{-1}||\leq||\mathbf{B}^{-1}||\;||\mathbf{A}^{-1}||\;||\mathbf{B}-\mathbf{A}||.$$

Combining these two results yields

$$\left|\operatorname{tr}\left\{\left(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\right)^{-1}-\left(z\mathbf{I}+z\frac{\lfloor nt\rfloor}{n}\underline{s}^0_{n,t}(z)\mathbf{T}_n\right)^{-1}\right\}\right|\leq Kn^{\frac{1}{2}}.$$

Hence,

$$\begin{aligned}
&\left|\frac{1}{z_1z_2}\operatorname{tr}\left(\left(\mathbf{I}+\frac{\lfloor nt_2\rfloor}{n}\underline{s}^0_{n,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I}+\frac{\lfloor nt_1\rfloor}{n}\underline{s}^0_{n,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\right.\\
&\left.\quad-\operatorname{tr}\left(\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\right|\\
\leq&\left|\operatorname{tr}\left(\left\{\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}-\left(z_2\mathbf{I}+z_2\frac{\lfloor nt_2\rfloor}{n}\underline{s}^0_{n,t_2}(z_2)\mathbf{T}_n\right)^{-1}\right\}\right.\right.\\
&\left.\left.\quad\times\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\right|\\
+&\left|\operatorname{tr}\left(\frac{1}{z_2}\left(\mathbf{I}+\frac{\lfloor nt_2\rfloor}{n}\underline{s}^0_{n,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right.\right.\\
&\left.\left.\quad\times\left\{\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}-\left(z_1\mathbf{I}+z_1\frac{\lfloor nt_1\rfloor}{n}\underline{s}^0_{n,t_1}(z_1)\mathbf{T}_n\right)^{-1}\right\}\mathbf{T}_n\right)\right|\\
\leq&Kn^{\frac{1}{2}}.
\end{aligned}$$

□

Lemma 7.1.5

$$\begin{aligned}
&\mathbb{E}\left|\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}^{-1}_{j,t_2}(z_2)\right]\mathbf{T}_n\right)\right.\\
&\left\{\frac{j-1}{n^2}b_{j,t_1}(z_1)b_{j,t_2}(z_2)\operatorname{tr}\left(\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{\lfloor nt_1\rfloor}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\right.\\
&\left.\left.-\frac{j-1}{n^2}\underline{s}^0_{n,t_1}(z_1)\underline{s}^0_{n,t_2}(z_2)\operatorname{tr}\left(\left(\mathbf{I}+\underline{s}^0_{n,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I}+\underline{s}^0_{n,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\right\}\right|\\
&\leq Kn^{\frac{1}{2}}
\end{aligned}$$

Proof. By (7.2), we have

$$\mathbb{E}\left|\operatorname{tr}\left(\mathbb{E}_j\left[\mathbf{D}_{j,t_1}^{-1}(z_1)\right]\mathbf{T}_n\mathbb{E}_j\left[\mathbf{D}_{j,t_2}^{-1}(z_2)\right]\mathbf{T}_n\right)\right| \le Kn.$$

Thus, it is left to show that

$$\begin{aligned}&\Bigg|b_{j,t_1}(z_1)b_{j,t_2}(z_2)\operatorname{tr}\left(\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\\&-\underline{s}_{n,t_1}^0(z_1)\underline{s}_{n,t_2}^0(z_2)\operatorname{tr}\left(\left(\mathbf{I}+\frac{\lfloor nt_2\rfloor}{n}\underline{s}_{n,t_2}^0(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I}+\frac{\lfloor nt_1\rfloor}{n}\underline{s}_{n,t_1}^0(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\Bigg| \qquad (7.5)\\&\le Kn^{\frac{1}{2}}.\end{aligned}$$

Invoking Lemma 7.1.4 and (6.40), we see that

$$\begin{aligned}&(7.5)\\&\le\Bigg|b_{j,t_1}(z_1)b_{j,t_2}(z_2)\operatorname{tr}\left(\left(z_2\mathbf{I}-\frac{\lfloor nt_2\rfloor-1}{n}b_{j,t_2}(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(z_1\mathbf{I}-\frac{\lfloor nt_1\rfloor-1}{n}b_{j,t_1}(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\\&-b_{j,t_1}(z_1)b_{j,t_2}(z_2)\frac{1}{z_1z_2}\operatorname{tr}\left(\left(\mathbf{I}+\frac{\lfloor nt_2\rfloor}{n}\underline{s}_{n,t_2}^0(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I}+\frac{\lfloor nt_1\rfloor}{n}\underline{s}_{n,t_1}^0(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\Bigg|\\&+\Bigg|\left(b_{j,t_1}(z_1)b_{j,t_2}(z_2)\frac{1}{z_1z_2}-\underline{s}_{n,t_1}^0(z_1)\underline{s}_{n,t_2}^0(z_2)\right)\\&\times\operatorname{tr}\left(\left(\mathbf{I}+\frac{\lfloor nt_2\rfloor}{n}\underline{s}_{n,t_2}^0(z_2)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\left(\mathbf{I}+\frac{\lfloor nt_1\rfloor}{n}\underline{s}_{n,t_1}^0(z_1)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\right)\Bigg|\\&\le Kn^{\frac{1}{2}}.\end{aligned}$$

□

Lemma 7.1.6 *We have for $z \in \mathbb{C}^+$ and $t \in (0,1]$*

$$z = -\frac{1}{\underline{\tilde{s}}_t(z)} + y\int\frac{\lambda}{1+\lambda t\underline{\tilde{s}}_t(z)}dH(\lambda).$$

and

$$z = -\frac{1}{\underline{\tilde{s}}_{n,t}^0(z)} + y_n\int\frac{\lambda}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)}dH_n(\lambda).$$

Proof. We begin by proving the first assertion. Similarly to (3.3), we have

$$\underline{\tilde{s}}_t(z) = -\frac{1-y_t}{z} + y_t\tilde{s}_t(z), \qquad (7.6)$$

which is equivalent to

$$\tilde{s}_t(z) = \frac{1}{y_t}\underline{\tilde{s}}_t(z) + \frac{1}{zy_t} - \frac{1}{z}.$$

Using

$$s_t\left(\frac{z}{t}\right) = t\tilde{s}_t(z)$$

and substituting this in (2.1) gives

$$\tilde{s}_t(z) = \int \frac{1}{\lambda t\left(1 - y_t - y_t z\tilde{s}_t(z)\right) - z} dH(\lambda).$$

We apply (7.6) to get

$$\tilde{\underline{s}}_t(z) = \frac{y_t - 1}{z} - \frac{y_t}{z}\int \frac{1}{\lambda t\tilde{\underline{s}}_t(z) + 1} dH(\lambda)$$

and solving this equation in z yields

$$\begin{aligned} z =& \frac{y_t - 1}{\tilde{\underline{s}}_t(z)} - \frac{y_t}{\tilde{\underline{s}}_t(z)}\int \frac{1}{1 + \lambda t\tilde{\underline{s}}_t(z)} dH(\lambda) \\ =& \frac{y_t - 1}{\tilde{\underline{s}}_t(z)} + \frac{y_t}{\tilde{\underline{s}}_t(z)}\int \frac{\lambda t\tilde{\underline{s}}_t(z)}{1 + \lambda t\tilde{\underline{s}}_t(z)} dH(\lambda) - \frac{y_t}{\tilde{\underline{s}}_t(z)} \\ =& -\frac{1}{\tilde{\underline{s}}_t(z)} + y\int \frac{\lambda}{1 + \lambda t\tilde{\underline{s}}_t(z)} dH(\lambda). \end{aligned}$$

Next, we will prove the second assertion. By using

$$\tilde{s}^0_{n,t}(z) = \frac{1}{y_{\lfloor nt\rfloor}}\tilde{\underline{s}}^0_{n,t}(z) + \frac{1}{zy_{\lfloor nt\rfloor}} - \frac{1}{z}$$

in equation (6.3), we get

$$\tilde{\underline{s}}^0_{n,t}(z) = -\frac{y_{\lfloor nt\rfloor}}{z}\int \frac{1}{1 + \lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z)} dH_n(\lambda) + \frac{y_{\lfloor nt\rfloor}}{z} - \frac{1}{z}.$$

Solving this equation in z,

$$\begin{aligned} z =& -\frac{y_{\lfloor nt\rfloor}}{\tilde{\underline{s}}^0_{n,t}(z)}\int \frac{1}{1 + \lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z)} dH_n(\lambda) + \frac{y_{\lfloor nt\rfloor}}{\tilde{\underline{s}}^0_{n,t}(z)} - \frac{1}{\tilde{\underline{s}}^0_{n,t}(z)} \\ =& y_n\int \frac{\lambda}{1 + \lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z)} dH_n(\lambda) - \frac{1}{\tilde{\underline{s}}^0_{n,t}(z)}. \end{aligned}$$

□

Lemma 7.1.7 *It holds*

$$\left|\frac{\min(\lfloor nt_1\rfloor, \lfloor nt_2\rfloor)}{n} a_n(z_1, z_2, t_1, t_2)\right| < 1,$$

where $a_n(z_1, z_2, t_1, t_2)$ *is defined in* (6.43).

Proof. For the sake of simplicity, we write $a_n(z_1, z_2) = a_n(z_1, z_2, t_1, t_2)$. Assume w.l.o.g. that $t_2 \leq t_1$. As a consequence of Lemma 7.1.6, we see

$$\tilde{\underline{s}}^0_{n,t}(z) = -\frac{1}{z + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)}}, \tag{7.7}$$

and conclude that

$$\begin{aligned}
\left| \frac{\lfloor nt_2 \rfloor}{n} a_n(z_1, z_2) \right| &= \left| \frac{\lfloor nt_2 \rfloor}{n} \frac{y_n \int \frac{\lambda^2 dH_n(\lambda)}{(1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z_1))(1+\frac{\lfloor nt_2 \rfloor}{n} \lambda \tilde{\underline{s}}^0_{n,t_2}(z_2))}}{\left(-z_1 + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z_1)}\right)\left(-z_2 + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt_2 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_2}(z_2)}\right)} \right| \\
&\leq \left(\frac{y_n \frac{\lfloor nt_2 \rfloor}{n} \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z_1)|^2}}{\left| -z_1 + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z_1)} \right|^2} \right)^{\frac{1}{2}} \left(\frac{y_n \frac{\lfloor nt_2 \rfloor}{n} \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_2 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_2}(z_2)|^2}}{\left| -z_2 + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt_2 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_2}(z_2)} \right|^2} \right)^{\frac{1}{2}} \\
&\leq \left(\frac{y_n \frac{\lfloor nt_1 \rfloor}{n} \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z_1)|^2}}{\left| -z_1 + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z_1)} \right|^2} \right)^{\frac{1}{2}} \left(\frac{y_n \frac{\lfloor nt_2 \rfloor}{n} \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_2 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_2}(z_2)|^2}}{\left| -z_2 + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt_2 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_2}(z_2)} \right|^2} \right)^{\frac{1}{2}} \\
&= \left(\frac{\frac{\lfloor nt_1 \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{n,t_1}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z)|^2}}{\operatorname{Im}(z) + \frac{\lfloor nt_1 \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{n,t_1}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z)|^2}} \right)^{\frac{1}{2}} \\
&\times \left(\frac{\frac{\lfloor nt_2 \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{n,t_2}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_2 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_2}(z)|^2}}{\operatorname{Im}(z) + \frac{\lfloor nt_2 \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{n,t_2}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_2 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_2}(z)|^2}} \right)^{\frac{1}{2}}. \\
&< 1.
\end{aligned}$$

The second equality follows from Lemma 7.7.1. Moreover,

$$\frac{\operatorname{Im}(z)}{\frac{\lfloor nt_1 \rfloor}{n} \operatorname{Im} \underline{s}^0_{n,t_1}(z) y_{\lfloor nt_1 \rfloor} \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z)|^2}}$$

is bounded away from zero, since by Lemma 6.10 (a) in Bai and Silverstein (2010), we have

$$\begin{aligned}
\frac{\lfloor nt_1 \rfloor}{n} \operatorname{Im} \tilde{\underline{s}}^0_{n,t_1}(z) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1 + \lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z)|^2} &= \left| y_n \operatorname{Im} \left(\int \frac{\lambda dH_n(\lambda)}{1 + \lambda \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z)} \right) \right| \\
&\leq y_n \left\| \mathbf{T}_n \left(\mathbf{I} + \frac{\lfloor nt_1 \rfloor}{n} \tilde{\underline{s}}^0_{n,t_1}(z) \mathbf{T}_n \right)^{-1} \right\| \leq \frac{4 y_n}{\operatorname{Im}(z)}.
\end{aligned}$$

Here, we used

$$\operatorname{Im}\left(\frac{1}{z}\right) = \frac{-\operatorname{Im}(z)}{|z|^2}$$

for the first equality. This finishes the proof. □

7.2 Proof of Lemma 6.3.4

To be precise, recall the definition of Z_n^1 and Z_n^2 in (6.47) and (6.48) and define $\hat{Z}_n^1$ and $\hat{Z}_n^2$ by Z_n^1 and Z_n^2, respectively, in the same way as $\hat{M}_n^1$ is defined by M_n^1 in equation (6.4). The bounds for the increments of $M_n(z,t)$, $z \in \mathcal{C}_n, t \in [t_0, 1]$ are given in the following lemma, which will be proven later in Section 7.7.

Lemma 7.2.1 *For $t \in [t_0, 1], z_1, z_2 \in \mathcal{C}_n$, it holds for sufficiently large $n \in \mathbb{N}$ under the assumptions of Theorem 6.3.2*

$$\mathbb{E}|M_n^1(z_1,t) - M_n^1(z_2,t)|^{2+\delta} \lesssim |z_1 - z_2|^{2+\delta}. \tag{7.8}$$

We also have for $t_1, t_2 \in [t_0, 1], z \in \mathcal{C}_n$

$$\mathbb{E}|Z_n^1(z,t_1,t_2)|^4 \lesssim \left(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\right)^4, \tag{7.9}$$

$$\mathbb{E}|Z_n^2(z,t_1,t_2)|^{4+\delta} \lesssim \left(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\right)^{2+\delta/2}, \tag{7.10}$$

where

$$M_n^1(z,t_1) - M_n^1(z,t_2) = Z_n^1(z,t_1,t_2) + Z_n^2(z,t_1,t_2), \tag{7.11}$$

and Z_n^1 and Z_n^2 are defined in (6.47) *and* (6.48)*, respectively.*

The bounds (6.50) and (6.51) for the moments of $\hat{Z}_n^1$ and $\hat{Z}_n^2$ follow directly from corresponding bounds (7.9) and (7.10) in Lemma 7.2.1.

We continue by proving the first assertion (6.49). If z_1 and z_2 are both contained in $\mathcal{C}_n$, the assertion directly follows from (7.8). Otherwise, we assume that $N \in \mathbb{N}$ is sufficiently large so that for all $n \geq N$

$$v_0 > \varepsilon_n n^{-1}.$$

Let $z_1 \in \mathcal{C}_n$ and $z_2 \notin \mathcal{C}_n$, that is, $0 \leq \operatorname{Im}(z_2) \leq \varepsilon_n n^{-1} \leq \operatorname{Im}(z_1)$. With the notation $\operatorname{Re}(z_2) = x \in \{x_l, x_r\}$ we have from (7.8)

$$\begin{aligned}
\mathbb{E}|\hat{M}_n^1(z_1,t) - \hat{M}_n^1(z_2,t)|^{2+\delta} =& \mathbb{E}|M_n^1(z_1,t) - M_n^1(x + i\varepsilon_n n^{-1},t)|^{2+\delta} \lesssim |z_1 - (x + i\varepsilon_n n^{-1})|^{2+\delta} \\
\leq& \left[(\operatorname{Re}(z_1) - x)^2 + (\operatorname{Im}(z_1) - \varepsilon_n n^{-1})^2\right]^{(2+\delta)/2} \\
\leq& \left[(\operatorname{Re}(z_1) - x)^2 + (\operatorname{Im}(z_1) - \operatorname{Im}(z_2))^2\right]^{(2+\delta)/2} \\
=& |z_1 - z_2|^{2+\delta}.
\end{aligned}$$

Finally, if both $z_1, z_2 \in \mathcal{C}^+ \setminus \mathcal{C}_n$, it follows from (7.8) that

$$\begin{aligned}
\mathbb{E}|\hat{M}_n^1(z_1,t) - \hat{M}_n^1(z_2,t)|^{2+\delta} =& \mathbb{E}|M_n^1(\operatorname{Re}(z_1) + i\varepsilon_n n^{-1}) - M_n^1(\operatorname{Re}(z_2) + i\varepsilon_n n^{-1})|^{2+\delta} \\
\lesssim& |\operatorname{Re}(z_1) - \operatorname{Re}(z_2)|^{2+\delta} \leq |z_1 - z_2|^{2+\delta},
\end{aligned}$$

which completes the proof of Lemma 6.3.4.

7.3 Proof of Lemma 6.3.6

As a preparation, we need the following auxiliary result.

Lemma 7.3.1 *It holds for all $n \in \mathbb{N}, z \in \mathbb{C}^+, t \in (0,1]$*

$$y_n \int \frac{dH_n(\lambda)}{1 + \lambda \frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]} + z y_n \mathbb{E}[\tilde{s}_{n,t}(z)]$$
$$= \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}\Bigg\{ \beta_{j,t}(z) \Bigg[\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z) \left(\frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \mathbf{T}_n - \mathbf{I} \right)^{-1} \mathbf{r}_j$$
$$- \frac{1}{n} \operatorname{tr} \left(\frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \mathbf{T}_n - \mathbf{I} \right)^{-1} \mathbf{T}_n \mathbb{E}[\mathbf{D}_t^{-1}(z)] \Bigg] \Bigg\}.$$

Proof of Lemma 7.3.1. This proof is inspired by (5.2) in Bai and Silverstein (1998). We write

$$\mathbf{D}_t(z) - \left(-z \frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \mathbf{T}_n - z\mathbf{I} \right) = \sum_{j=1}^{\lfloor nt \rfloor} \mathbf{r}_j \mathbf{r}_j^\star - \left(-z \frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \right) \mathbf{T}_n.$$

Next, we use (see (7.4))

$$\mathbf{r}_j^\star \mathbf{D}_t^{-1}(z) = \beta_{j,t}(z) \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z)$$

and from Lemma 7.1.3

$$-z \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] = \frac{1}{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)]$$

to conclude

$$\left(-z \frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \mathbf{T}_n - z\mathbf{I} \right)^{-1} - \mathbb{E}[\mathbf{D}_t^{-1}(z)]$$
$$= \left(-z \frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \mathbf{T}_n - z\mathbf{I} \right)^{-1} \mathbb{E} \left[\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbf{r}_j \mathbf{r}_j^\star - \left(-z \frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \mathbf{T}_n \right) \right) \mathbf{D}_t^{-1}(z) \right]$$
$$= -z^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}\Bigg\{ \beta_{j,t}(z) \Bigg[\left(\frac{\lfloor nt \rfloor}{n} \mathbb{E}\underline{\tilde{s}}_{n,t}(z) \mathbf{T}_n - \mathbf{I} \right)^{-1} \mathbf{r}_j \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z)$$
$$- \frac{1}{n} \left(\frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z) \mathbf{T}_n - \mathbf{I} \right)^{-1} \mathbf{T}_n \mathbb{E}[\mathbf{D}_t^{-1}(z)] \Bigg] \Bigg\}.$$

Taking traces on both sides and multiplying by $-z/n$, we get

$$y_n \int \frac{dH_n(\lambda)}{1 + \lambda \frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]} + z y_n \mathbb{E}[\tilde{s}_{n,t}(z)] = \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}\Bigg\{ \beta_{j,t}(z) \Bigg[\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z) \left(\frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \mathbf{T}_n - \mathbf{I} \right)^{-1} \mathbf{r}_j$$

$$-\frac{1}{n}\operatorname{tr}\left(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n-\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbb{E}[\mathbf{D}_t^{-1}(z)]\Bigg]\Bigg\}$$

□

Proof of Lemma 6.3.6 . We begin by deriving an alternative form for $R_{n,t}(z)$. By

$$\mathbb{E}[\tilde{s}_{n,t}(z)]=\frac{1}{y_{\lfloor nt\rfloor}}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]+\frac{1}{zy_{\lfloor nt\rfloor}}-\frac{1}{z}$$

and Lemma 7.3.1, we have

$$\begin{aligned}
-\frac{\lfloor nt\rfloor}{n}\mathbb{E}\underline{\tilde{s}}_{n,t}(z)\Big(-z-\frac{1}{\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}\Big)&\\
=&y_n\int\frac{dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}+zy_n\mathbb{E}[\tilde{s}_{n,t}(z)]\\
=&-n^{-1}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\beta_{j,t}(z)\Big(\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}\underline{\tilde{s}}_{n,t}(z)\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{r}_j\\
&-\frac{1}{n}\mathbb{E}[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}\underline{\tilde{s}}_{n,t}(z)\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_t^{-1}(z)]\Big)\Big]\\
=&-y_nn^{-1}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\beta_{j,t}(z)\Big(\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}\underline{\tilde{s}}_{n,t}(z)\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j\\
&-\frac{1}{p}\mathbb{E}[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}\underline{\tilde{s}}_{n,t}(z)\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_t^{-1}(z)]\Big)\Big]\\
=&-y_nn^{-1}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\beta_{j,t}(z)d_{j,t}(z)\right]=-\frac{\lfloor nt\rfloor}{n}y_{\lfloor nt\rfloor}n^{-1}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\beta_{j,t}(z)d_{j,t}(z)\right].
\end{aligned}$$

This implies

$$\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)=-z-\frac{1}{\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]},$$

and we can conclude

$$\begin{aligned}
\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]-\underline{\tilde{s}}_{n,t}^0(z)&\\
=&\frac{1}{-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}-\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)}-\frac{1}{-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)}}\\
=&\frac{y_n\Big(\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)}-\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}\Big)+\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)}{\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}-\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)\Big)\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)}\Big)}
\end{aligned}$$

$$=\frac{y_n\frac{\lfloor nt\rfloor}{n}\left(\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]-\tilde{\underline{s}}^0_{n,t}(z)\right)\int\frac{\lambda^2 dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z))}+\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)}{\left(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]}-\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)\right)\left(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z)}\right)} \tag{7.12}$$

$$=\frac{y_n\frac{\lfloor nt\rfloor}{n}\left(\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]-\tilde{\underline{s}}^0_{n,t}(z)\right)\int\frac{\lambda^2 \tilde{\underline{s}}^0_{n,t}(z) dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z))}}{-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]}-\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)}$$
$$+\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\tilde{\underline{s}}^0_{n,t}(z).$$

□

7.4 Proof of Theorem 6.3.7

In this section, $D[0,1]^2$ denotes the Skorokhod space on $[0,1]^2$ (see Bickel and Wichura, 1971; Neuhaus, 1971, for a formal definition). We will identify the set $\mathcal{C}^+\times[0,1]$ with the square $[0,1]^2$ and proceed in several steps. First, we will show a uniqueness condition, second we prove the existence of a Skorokhod-limit of $(\mathbb{E}[\tilde{s}_{n,\cdot}(\cdot)])_{n\in\mathbb{N}}$. We conclude by proving that the Skorokhod-limit is in fact a uniform limit.

Lemma 7.4.1 *Let* $(\mathbb{E}[\tilde{s}_{k(n),t}(z)])_{n\in\mathbb{N}}$ *and* $(\mathbb{E}[\tilde{s}_{l(n),t}(z)])_{n\in\mathbb{N}}$ *be two subsequences of* $(\mathbb{E}[\tilde{s}_{n,t}(z)])_{n\in\mathbb{N}}$ *and* m_1 *and* m_2 *be functions on* $\mathcal{C}^+\times[t_0,1]$. *If for* $z\in\mathcal{C}^+, t\in[t_0,1]$,

$$\lim_{n\to\infty}\mathbb{E}[\tilde{s}_{k(n),t}(z)]=m_1(z,t)\quad\text{and}\quad\lim_{n\to\infty}\mathbb{E}[\tilde{s}_{j(n),t}(z)]=m_2(z,t),$$

then we have for $z\in\mathcal{C}^+, t\in[t_0,1]$

$$m_1(z,t)=m_2(z,t)=\tilde{s}_t(z),$$

where $\tilde{s}_t$ *denotes the Stieltjes transform of* $\tilde{F}^{y_t,H}$ *given in* (2.4)

Proof of Lemma 7.4.1. We show that a potential limit of the sequence $(\mathbb{E}[\tilde{s}_{n,\cdot}(\cdot)])_{n\in\mathbb{N}}$ satisfies an equation which admits a unique solution. For this purpose, we will adapt ideas from Bai and Zhou (2008) and also correct some arguments in step 2 in the proof of their Theorem 1.1. To be precise, define for $z\in\mathcal{C}^+$ and $t\in[t_0,1]$

$$\mathbf{K}=b_t(z)\mathbf{T}_n,$$

and note that

$$\mathbf{D}_t(z)-\left(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\right)=\sum_{k=1}^{\lfloor nt\rfloor}\mathbf{r}_k\mathbf{r}_k^\star-\frac{\lfloor nt\rfloor}{n}\mathbf{K}.$$

Multiplying with $((\lfloor nt\rfloor/n)\mathbf{K}-z\mathbf{I})^{-1}$ and $\mathbf{D}_t^{-1}(z)$ from the left and from the right, respectively, and using identity (6.1.11) from Bai and Silverstein (2010) yields

$$\left(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\right)^{-1}-\mathbf{D}_t^{-1}(z)$$

$$
\begin{aligned}
&=\sum_{k=1}^{\lfloor nt\rfloor}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_t^{-1}(z)-\frac{\lfloor nt\rfloor}{n}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{K}\mathbf{D}_t^{-1}(z)\\
&=\sum_{k=1}^{\lfloor nt\rfloor}\beta_{k,t}(z)\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)-\frac{\lfloor nt\rfloor}{n}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{K}\mathbf{D}_t^{-1}(z).
\end{aligned}
$$

This implies for $l\in\{0,1\}$

$$
\begin{aligned}
&\frac{1}{p}\operatorname{tr}\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}-\frac{1}{p}\operatorname{tr}\mathbf{T}_n^l\mathbf{D}_t^{-1}(z)\\
=&\frac{1}{p}\sum_{k=1}^{\lfloor nt\rfloor}\beta_{k,t}(z)\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{r}_k-\frac{1}{p}\operatorname{tr}\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{K}\mathbf{D}_t^{-1}(z)\\
=&\frac{1}{p}\sum_{k=1}^{\lfloor nt\rfloor}\beta_{k,t}(z)\varepsilon_k,
\end{aligned}
$$

where

$$
\begin{aligned}
\varepsilon_k=&\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{r}_k-n^{-1}\beta_{k,t}^{-1}(z)\operatorname{tr}\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{K}\mathbf{D}_t^{-1}(z)\\
=&\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{r}_k-n^{-1}\operatorname{tr}\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{K}\mathbf{D}_t^{-1}(z)\left(1+\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)\mathbf{r}_k\right).
\end{aligned}
$$

We decompose $\varepsilon_k=\varepsilon_{k1}+\varepsilon_{k2}+\varepsilon_{k3}$, where

$$
\begin{aligned}
\varepsilon_{k1}=&n^{-1}\operatorname{tr}\mathbf{T}_n^{l+1}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{D}_{k,t}^{-1}(z)-n^{-1}\operatorname{tr}\mathbf{T}_n^{l+1}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{D}_t^{-1}(z)\\
\varepsilon_{k2}=&\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{r}_k-n^{-1}\operatorname{tr}\mathbf{T}_n^{l+1}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{D}_{k,t}^{-1}(z)\\
\varepsilon_{k3}=&-n^{-1}\operatorname{tr}\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{K}\mathbf{D}_t^{-1}(z)\left((1+\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)\mathbf{r}_k)\right)\\
&+n^{-1}\operatorname{tr}\mathbf{T}_n^{l+1}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{D}_t^{-1}(z)\\
=&-n^{-1}\operatorname{tr}\mathbf{T}_n^{l+1}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\mathbf{D}_t^{-1}(z)\left\{b_t(z)\left(\mathbf{r}_k^\star\mathbf{D}_{k,t}^{-1}(z)\mathbf{r}_k+1\right)-1\right\},
\end{aligned}
$$

and we have used the fact that the matrices $\mathbf{T}_n$ and $((\lfloor nt\rfloor/n)\mathbf{K}-z\mathbf{I})^{-1}$ commute. Similar arguments as given by Bai and Silverstein (2010) for their estimate (9.9.13) yield

$$
\Big|\Big|\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}\Big|\Big|\le K,
$$

and this estimate can be used to show

$$
\mathbb{E}|\varepsilon_{ki}|^2\to 0,\ n\to\infty\quad,i\in\{1,2,3\}.
$$

This implies for $l \in \{0,1\}$

$$\frac{1}{p}\Big(\mathbb{E}\operatorname{tr}\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}-\mathbb{E}\operatorname{tr}\mathbf{T}_n^l\mathbf{D}_t^{-1}(z)\Big)\to 0,\ n\to\infty. \tag{7.13}$$

Using (7.13) with $l=0$ for the first line and $l=1$ for the second one, we have

$$\frac{1}{p}\mathbb{E}\operatorname{tr}\Big(\frac{\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}-z\mathbf{I}\Big)^{-1}-\mathbb{E}\tilde{s}_{n,t}(z)\to 0, \tag{7.14}$$

$$\frac{1}{p}\mathbb{E}\operatorname{tr}\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n\Big(\frac{\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}-z\mathbf{I}\Big)^{-1}-a_{n,t}(z)\to 0, \tag{7.15}$$

where $a_{n,t}(z)=(\lfloor nt\rfloor/n)p^{-1}\mathbb{E}\operatorname{tr}\mathbf{T}_n\mathbf{D}_t^{-1}(z)$, so that $1+y_{\lfloor nt\rfloor}a_{n,t}(z)=b_t(z)$. We use

$$\Big|\frac{1}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}\Big|\le\frac{|z|}{v}$$

to conclude from (7.15)

$$1+\frac{z}{p}\mathbb{E}\operatorname{tr}\Big(\frac{\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}-z\mathbf{I}\Big)^{-1}-\frac{a_{n,t}(z)}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}\to 0.$$

Combining this with (7.14) yields

$$1+z\mathbb{E}\tilde{s}_{n,t}(z)-\frac{a_{n,t}(z)}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}\to 0$$

and, by rearranging terms and multiplying with $y_{\lfloor nt\rfloor}$,

$$\frac{1}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}=1-y_{\lfloor nt\rfloor}(1+z\mathbb{E}\tilde{s}_{n,t}(z))+o(1).$$

Substituting this in (7.14), we get

$$\frac{1}{p}\mathbb{E}\operatorname{tr}\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n\left(1-y_{\lfloor nt\rfloor}(1+z\mathbb{E}\tilde{s}_{n,t}(z))\right)-z\mathbf{I}\Big)^{-1}-\mathbb{E}\tilde{s}_{n,t}(z)\to 0. \tag{7.16}$$

Due to (7.16), any potential limit $\tilde{s}_\cdot(\cdot)$ of $(\mathbb{E}\tilde{s}_{n,\cdot}(\cdot))_{n\in\mathbb{N}}$ satisfies

$$\tilde{s}_t(z)=\int\frac{1}{\lambda t(1-y_t(1+z\tilde{s}_t(z)))-z}dH(\lambda).$$

It follows from Theorem 1.1 in Bai and Zhou (2008), that this equation admits a unique solution $\tilde{s}_\cdot(\cdot)$.

□

In the following lemma, we consider for technical reasons the functions $\hat{s}_{n,\cdot}(\cdot):\mathcal{C}^+\times[0,1]\to\mathbb{C}$ with $\hat{s}_{n,t}(z)=0$ for $t<t_0$ and for $t\in[t_0,1], z=x+iv\in\mathcal{C}^+$

$$\hat{s}_{n,t}(z)=\begin{cases}\tilde{s}_{n,t}(z) & : z\in\mathcal{C}_n\\ \tilde{s}_{n,t}(x_r+in^{-1}\varepsilon_n) & : x=x_r,\ v\in[0,n^{-1}\varepsilon_n]\\ \tilde{s}_{n,t}(x_l+in^{-1}\varepsilon_n) & : x=x_l,\ v\in[0,n^{-1}\varepsilon_n]\end{cases}$$

and for $t \in [0,1], z \in \mathcal{C}^+$

$$\hat{m}_{n,t}(z) = \begin{cases} \lim_{t \to 1} \hat{s}_{n,t}(z) = \hat{s}_{n,\frac{n-1}{n}}(z) & : t = 1 \\ \hat{s}_{n,t}(z) & : t \in [0,1). \end{cases}$$

Note that for $t \in [0,1]$, the functions $\hat{s}_{n,t}(\cdot)$ and $\tilde{s}_{n,t}(\cdot)$ coincide on $\mathcal{C}_n$, $n \in \mathbb{N}$ and that for $z \in \mathcal{C}^+$, the functions $\hat{s}_{n,t}(z)$ and $\hat{m}_{n,t}(z)$ differ only in the point $t = 1$.

Lemma 7.4.2 *The set $\{\mathbb{E}[\hat{m}_{n,\cdot}(\cdot)] : n \in \mathbb{N}\}$ has a compact closure in the Skorokhod space $D[0,1]^2$.*

Proof of Lemma 7.4.2. The sequence $(\mathbb{E}\hat{m}_{n,\cdot}(\cdot))_{n\in\mathbb{N}}$ is bounded, since by Lemma 7.7.3, we get uniformly with respect to $t \in [t_0,1], z \in \mathcal{C}_n, n \in \mathbb{N}$

$$|\mathbb{E}\tilde{s}_{n,t}(z)| = \frac{1}{p} \left|\mathbb{E} \operatorname{tr} \mathbf{D}_t^{-1}(z)\right| \le \mathbb{E}||\mathbf{D}_t^{-1}(z)|| \le K.$$

We aim to show

$$\lim_{\delta\to 0} \sup_{n\in\mathbb{N}} \sup_{\substack{(t_1,t_2,t)\in\mathcal{A}_\delta,\\ z\in\mathcal{C}^+}} \min\left(|\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_1}(z)]|, |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_2}(z)]|\right) = 0, \tag{7.17}$$

where

$$\mathcal{A}_\delta = \{(t_1,t_2,t) \ : \ t_1 \le t \le t_2, \ t_2 - t_1 \le \delta\}.$$

Let $\varepsilon > 0$ be given. We choose $N \in \mathbb{N}$ sufficiently large such that $\frac{1}{N} < \varepsilon$ and $\delta > 0$ sufficiently small such that $\delta < \varepsilon$ and for all $n \in \{1,\dots,N\}$

$$\lfloor nt \rfloor - \lfloor nt_2 \rfloor = 0 \text{ or } \lfloor nt \rfloor - \lfloor nt_1 \rfloor = 0,$$

where $(t_1,t_2,t) \in \mathcal{A}_\delta$. Then, it holds

$$\sup_{n\le N} \sup_{\substack{(t_1,t_2,t)\in\mathcal{A}_\delta,\\ z\in\mathcal{C}^+}} \min\left(|\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_1}(z)]|, |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_2}(z)]|\right) = 0.$$

For $n \ge N$ we conclude

$$\begin{aligned} \left|\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_1}(z)]\right| \le& K\frac{\lfloor nt \rfloor - \lfloor nt_1 \rfloor}{n} \\ \le& K\left(\left|\frac{\lfloor nt \rfloor - nt}{n}\right| + |t_1 - t| + \left|\frac{\lfloor nt_1 \rfloor - nt_1}{n}\right|\right) \le 3\varepsilon K \end{aligned}$$

and obtain

$$\sup_{n\ge N} \sup_{\substack{(t_1,t_2,t)\in\mathcal{A}_\delta,\\ z\in\mathcal{C}^+}} \min\left\{|\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_1}(z)]|, |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_2}(z)]|\right\} \le 3\varepsilon K.$$

Thus, (7.17) holds true. Similarly, one can show

$$\lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\sup_{\substack{t_1,t_2\in[1-\delta,1),\\ z\in\mathcal{C}^+}}|\mathbb{E}[\hat{s}_{n,t_1}(z)]-\mathbb{E}[\hat{s}_{n,t_2}(z)]|=0.$$

By definition, this implies

$$\lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\sup_{\substack{t_1,t_2\in[1-\delta,1],\\ z\in\mathcal{C}^+}}|\mathbb{E}[\hat{m}_{n,t_1}(z)]-\mathbb{E}[\hat{m}_{n,t_2}(z)]|=0.$$

Since $\hat{s}_{n,t}(z)=0$ for $t<t_0$, we also have for $\delta<t_0$

$$\sup_{n\in\mathbb{N}}\sup_{\substack{t_1,t_2\in[0,\delta),\\ z\in\mathcal{C}^+}}|\mathbb{E}[\hat{s}_{n,t_1}(z)]-\mathbb{E}[\hat{s}_{n,t_2}(z)]|=0.$$

Therefore, it follows from the proof of Theorem 14.4 in Billingsley (1968) that

$$\begin{aligned}0=&\lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\sup_{z\in\mathcal{C}^+}\inf_{\substack{(t_0,\dots,t_r)\in\mathcal{B}_{\delta,r},\\ r\in\mathbb{N}}}\max_{0\le i\le r}\sup_{t,t'\in[t_{i-1},t_i\rangle}|\mathbb{E}[\hat{s}_{n,t}(z)]-\mathbb{E}[\hat{s}_{n,t'}(z)]|\\ =&\lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\sup_{z\in\mathcal{C}^+}\inf_{\substack{(t_0,\dots,t_r)\in\mathcal{B}_{\delta,r},\\ r\in\mathbb{N}}}\max_{0\le i\le r}\sup_{t,t'\in[t_{i-1},t_i\rangle}|\mathbb{E}[\hat{m}_{n,t}(z)]-\mathbb{E}[\hat{m}_{n,t'}(z)]|,\end{aligned}\tag{7.18}$$

where $[t_{i-1},t_i\rangle$ is defined as $[t_{i-1},t_i]$ if $t_i=1$ and as $[t_{i-1},t_i)$ otherwise, and we set

$$\mathcal{B}_{\delta,r}=\{(t_0,\dots,t_r):0=t_0<t_1<\dots<t_r=1,\ t_i-t_{i-1}>\delta\text{ for }i\in\{1,\dots,r\}\}.$$

For the next step, we have for $z_1,z_2\in\mathcal{C}_n$

$$\begin{aligned}|\mathbb{E}[\tilde{s}_{n,t}(z_1)]-\mathbb{E}[\tilde{s}_{n,t}(z_2)]|=&\frac{1}{p}|z_1-z_2||\mathbb{E}\operatorname{tr}\mathbf{D}_t^{-1}(z_1)\mathbf{D}_t^{-1}(z_2)|\\ \le&K|z_1-z_2|\mathbb{E}||\mathbf{D}_t^{-1}(z_1)\mathbf{D}_t^{-1}(z_2)||\le K|z_1-z_2|\end{aligned}$$

uniformly in $t\in[t_0,1]$, which implies for $z_1,z_2\in\mathcal{C}_n$ or $z_1,z_2\notin\mathcal{C}_n$ that

$$|\mathbb{E}[\hat{s}_{n,t}(z_1)]-\mathbb{E}[\hat{s}_{n,t}(z_2)]|\le K|z_1-z_2|.$$

In the case $z_1=x_1+iv_1\in\mathcal{C}_n$ and $z_2=x_2+iv_2\notin\mathcal{C}_n$, we conclude

$$\begin{aligned}|\mathbb{E}[\hat{s}_{n,t}(z_1)]-\mathbb{E}[\hat{s}_{n,t}(z_2)]|=&|\mathbb{E}[\tilde{s}_{n,t}(z_1)]-\mathbb{E}[\tilde{s}_{n,t}(x_2+in^{-1}\varepsilon_n)]|\le K|z_1-(x_2+in^{-1}\varepsilon_n)|\\ =&\left\{(x_1-x_2)^2+(v_1-n^{-1}\varepsilon_n)^2\right\}^{\frac12}\le\left\{(x_1-x_2)^2+(v_1-v_2)^2\right\}^{\frac12}\\ \le&K|z_1-z_2|,\end{aligned}$$

since $v_2\le n^{-1}\varepsilon_n\le v_1$. Thus, we have

$$\lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\sup_{\substack{t\in[0,1],\\ z_1,z_2\in\mathcal{C}^+,\\ |z_1-z_2|<\delta}}|\mathbb{E}[\hat{s}_{n,t}(z_1)]-\mathbb{E}[\hat{s}_{n,t}(z_2)]|=0.\tag{7.19}$$

Since for $A \times B \subset (\mathcal{C}^+)^2, C \times D \subset [0,1]^2$

$$\begin{aligned}&\sup_{\substack{(z,z')\in A\times B,\\(t,t')\in C\times D}} |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t'}(z')]|\\ \leq &\sup_{\substack{(z,z')\in A\times B,\\ t\in C}} |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t}(z')]| + \sup_{\substack{z'\in B,\\(t,t')\in C\times D}} |\mathbb{E}[\hat{s}_{n,t}(z')] - \mathbb{E}[\hat{s}_{n,t'}(z')]|,\end{aligned}$$

we conclude from (7.18) and (7.19)

$$\begin{aligned}&\lim_{\delta\to 0}\sup_{n\in\mathbb{N}} \inf_{\substack{((t_0,z_0),\ldots,(t_r,z_r))\in\mathcal{B}^{(2)}_{\delta,r},\\ r\in\mathbb{N}}} \max_{0\leq i\leq r} \sup_{\substack{t,t'\in[t_{i-1},t_i)\\ z,z'\in[z_{i-1},z_i\rangle}} |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t'}(z')]|\\ =&\lim_{\delta\to 0}\sup_{n\in\mathbb{N}} \inf_{\substack{((t_0,z_0),\ldots,(t_r,z_r))\in\mathcal{B}^{(2)}_{\delta,r},\\ r\in\mathbb{N}}} \max_{0\leq i\leq r} \sup_{\substack{t,t'\in[t_{i-1},t_i\rangle\\ z,z'\in[z_{i-1},z_i\rangle}} |\mathbb{E}[\hat{m}_{n,t}(z)] - \mathbb{E}[\hat{m}_{n,t'}(z')]|\\ =&0, \end{aligned} \tag{7.20}$$

where

$$\begin{aligned}B^{(2)}_{\delta,r} = \{((t_0,z_0),\ldots,(t_r,z_r)) \ : &0 = t_0 < t_1 < \ldots < t_r = 1,\ 0 = z_0 < z_1 < \ldots < z_r = 1,\\ &t_i - t_{i-1} > \delta,\ z_i - z_{i-1} > \delta \text{ for } i \in \{1,\ldots,r\}\}.\end{aligned}$$

Note that in this definition, an element $z \in \mathcal{C}^+$ is identified with its representative in $[0,1]$.
One can observe that (7.20) is equivalent to

$$\lim_{\delta\to 0}\sup_{n\in\mathbb{N}} \omega'_{\mathbb{E}[\hat{m}_{n,\cdot}(\cdot)]}(\delta) = 0,$$

where the modulus ω' is defined in Neuhaus (1971). Applying Theorem 2.1 in this reference, we conclude that $\{\mathbb{E}[\hat{m}_{n,\cdot}(\cdot)] : n \in \mathbb{N}\}$ has a compact closure in $D[0,1]^2$. □

Proof of Theorem 6.3.7. From Lemma 7.4.1 and Lemma 7.4.2, we conclude that

$$\lim_{n\to\infty} d_2|_{\mathcal{C}_n\times[t_0,1)}(\mathbb{E}[\tilde{s}_{n,\cdot}(\cdot)], \tilde{s}_\cdot(\cdot)) = \lim_{n\to\infty} d_2|_{\mathcal{C}_n\times[t_0,1)}(\mathbb{E}[\hat{m}_{n,\cdot}(\cdot)], \tilde{s}_\cdot(\cdot)) = 0,$$

where $d_2|_A$ for some set $A \subset \mathcal{C}^+ \times [0,1]$ denotes the Skorokhod metric restricted to functions on A. Observe that for $t = 1$

$$\lim_{n\to\infty}\sup_{z\in\mathcal{C}_n} |\mathbb{E}[\tilde{s}_{n,1}(z)] - \tilde{s}_1(z)| = 0.$$

Then it is straightforward to show that

$$\lim_{n\to\infty} d_2|_{\mathcal{C}_n\times[t_0,1]}(\mathbb{E}[\tilde{s}_{n,\cdot}(\cdot)], \tilde{s}_\cdot(\cdot)) = 0. \tag{7.21}$$

The considerations in the proof of Lemma 7.4.2 reveal that $\tilde{s}_\cdot(\cdot) \in C(\mathcal{C}^+ \times [t_0,1])$. In this case, the convergence in the Skorokhod space in (7.21) implies the uniform convergence

$$\lim_{n\to\infty}\sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} |\mathbb{E}[\tilde{s}_{n,t}(z)] - \tilde{s}_t(z)| = 0.$$

A similar convergence result with respect to the sup-norm can be shown for the Stieltjes transform $\tilde{\underline{s}}_{n,t}(z)$. More precisely, since

$$\tilde{\underline{s}}_t(z) = -\frac{1-y_t}{z} + y_t \tilde{s}_t(z),$$
$$\tilde{\underline{s}}_{n,t}(z) = -\frac{1-y_{\lfloor nt \rfloor}}{z} + y_{\lfloor nt \rfloor} \tilde{s}_{n,t}(z),$$

we also have

$$\lim_{n\to\infty} \sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} |\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)] - \tilde{\underline{s}}_t(z)| = 0.$$

□

7.5 Proof of Theorem 6.3.8

The second assertion directly follows from the first one combined with Theorem 6.3.7. Therefore, it is sufficient to show that $(M_n^2)_{n\in\mathbb{N}}$ is uniformly bounded. For this purpose, we use the following lemma.

Lemma 7.5.1 *We have*

$$\sup_{\substack{n\in\mathbb{N},\\ z\in\mathcal{C}_n,\\ t\in[t_0,1]}} \Big| \frac{\operatorname{Im}(\tilde{\underline{s}}^0_{n,t}(z))}{\operatorname{Im}(z)} \Big| \leq K.$$

Proof of Lemma 7.5.1. We have for sufficiently large n

$$\operatorname{Im}(\tilde{\underline{s}}^0_{n,t}(z)) = \int \operatorname{Im}\Big(\frac{1}{\lambda - z}\Big) d\tilde{\underline{F}}^{y_{\lfloor nt\rfloor},H_n}(\lambda) = \int \frac{-\operatorname{Im}(\lambda - z)}{|\lambda - z|^2} d\tilde{\underline{F}}^{y_{\lfloor nt\rfloor},H_n}(\lambda)$$
$$= \int \frac{\operatorname{Im}(z)}{(\lambda - \operatorname{Re}(z))^2 + \operatorname{Im}^2(z)} d\tilde{\underline{F}}^{y_{\lfloor nt\rfloor},H_n}(\lambda) \leq K \operatorname{Im}(z),$$

since for $z \in \mathcal{C}_l \cup \mathcal{C}_r$, $\operatorname{Re}(z) \in \{x_l, x_r\}$ is uniformly bounded away from the support of $\tilde{\underline{F}}^{y_{\lfloor nt\rfloor},H_n}$ for sufficiently large n (Lemma 6.4.1). If $z \in \mathcal{C}_u$, then $\operatorname{Im}(z) = v_0$ is constant and hence, the denominator is also uniformly bounded away from 0.

□

To continue with the proof of Theorem 6.3.8, we note that it follows from (7.12) in the proof of Lemma 6.3.6 in Section 7.3 that

$$\lfloor nt\rfloor(\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)] - \tilde{\underline{s}}^0_{n,t}(z)) - \lfloor nt\rfloor \frac{\lfloor nt\rfloor}{n} R_{n,t}(z)\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\tilde{\underline{s}}^0_{n,t}(z)$$
$$= \frac{y_n \frac{\lfloor nt\rfloor}{n} \lfloor nt\rfloor \left(\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)] - \tilde{\underline{s}}^0_{n,t}(z)\right) \int \frac{\lambda^2 dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z))}}{\left(-z + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]} - R_{n,t}(z)\right)\left(-z + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\tilde{\underline{s}}^0_{n,t}(z)}\right)},$$

which is equivalent to

$$\lfloor nt\rfloor\big(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]-\underline{\tilde{s}}^0_{n,t}(z)\big)=\frac{\lfloor nt\rfloor\frac{\lfloor nt\rfloor}{n}R_{n,t}(z)\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\underline{\tilde{s}}^0_{n,t}(z)}{1-\frac{y_n\frac{\lfloor nt\rfloor}{n}\int\frac{\lambda^2dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z))}}{\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}-R_{n,t}(z)\Big)\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)}\Big)}}.$$

Note that $\underline{\tilde{s}}^0_{n,t}(z)$ is uniformly bounded which follows by a similar argument as given in the proof of Lemma 7.5.1. In order to show that the sequence $(M_n^2)_{n\in\mathbb{N}}$ is uniformly bounded, by using (6.53), it is sufficient to show that the denominator is uniformly bounded away from 0 for sufficiently large n. For this aim, it is sufficient to prove that

$$\left|\frac{y_n\frac{\lfloor nt\rfloor}{n}\int\frac{\lambda^2\underline{\tilde{s}}^0_{n,t}(z)\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z))}}{\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}-R_{n,t}(z)\Big)\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)}\Big)}\right|<1$$

holds uniformly. Similarly to the proof of Lemma 7.7.6, we conclude that for any bounded subset $S\subset\mathbb{C}^+$

$$\inf_{\substack{n\in\mathbb{N},\\ z\in S,\\ t\in[t_0,1]}}|\underline{\tilde{s}}^0_{n,t}(z)|>0.$$

Using this, Hölder's inequality, Lemma 7.5.1 and Lemma 7.7.1, we obtain for sufficiently large n

$$\begin{aligned}
&\left|y_n\frac{\int\frac{\lambda^2\frac{\lfloor nt\rfloor}{n}dH_n(\lambda)}{(1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])(1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z))}}{\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}-R_{n,t}(z)\Big)\Big(-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)}\Big)}\right|^2\\
\le&\frac{y_n\frac{\lfloor nt\rfloor}{n}\int\frac{\lambda^2dH_n(\lambda)}{\left|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)\right|^2}}{\left|-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)}\right|^2}\frac{y_n\frac{\lfloor nt\rfloor}{n}\int\frac{\lambda^2dH_n(\lambda)}{\left|1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right|^2}}{\left|-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}-R_{n,t}(z)\right|^2}\\
=&\frac{\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\underline{\tilde{s}}^0_{n,t}(z))y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)|^2}}{\operatorname{Im}(z)+\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\underline{\tilde{s}}^0_{n,t}(z))y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)|^2}}\\
&\times\frac{\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}}{\operatorname{Im}(z)+\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}+\operatorname{Im}(R_{n,t}(z))}\\
\le&1-\frac{\operatorname{Im}(z)}{\operatorname{Im}(z)+\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\underline{\tilde{s}}^0_{n,t}(z))y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)|^2}}\\
\le&1-\frac{\operatorname{Im}(z)}{\operatorname{Im}(z)+\frac{\lfloor nt\rfloor}{n}K\operatorname{Im}(z)y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}^0_{n,t}(z)|^2}}\le1-\frac{1}{1+K}<1,
\end{aligned}$$

where we used the fact that $\operatorname{Im}(R_{n,t}(z))+\operatorname{Im}(z)\ge0$ for sufficiently large n, which follows from Lemma 7.7.8. This finishes the proof of Theorem 6.3.8.

7.6 Proof of the statement (6.53)

Using (6.14) and the representation

$$\beta_{j,t}(z) = \overline{\beta}_{j,t}(z) - \overline{\beta}_{j,t}^2(z)\hat{\gamma}_{j,t}(z) + \overline{\beta}_{j,t}^2(z)\beta_{j,t}(z)\hat{\gamma}_{j,t}^2(z), \tag{7.22}$$

we obtain

$$\begin{aligned}
&\lfloor nt\rfloor R_{n,t}(z)\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)] = y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\\
=&-y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\beta_{j,t}(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac12}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac12}\mathbf{q}_j\\
&-\frac1p\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_t^{-1}(z)\Big]\Big\}\Big]\\
=&-y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\beta_{j,t}(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac12}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac12}\mathbf{q}_j\\
&-\frac1p\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big]\Big\}\Big]\\
&+\frac{1}{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\beta_{j,t}(z)\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbb{E}\left[\mathbf{D}_t^{-1}(z)-\mathbf{D}_{j,t}^{-1}(z)\right]\Big]\\
=&T_{n,1}(z,t)+T_{n,2}(z,t)+o(1)
\end{aligned}$$

uniformly with respect to $z\in\mathcal{C}_n, t\in[t_0,1]$, where the terms $T_{n,1}$ and $T_{n,2}$ are defined by

$$\begin{aligned}
T_{n,1}(z,t) =& y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\overline{\beta}_{j,t}^2(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac12}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac12}\mathbf{q}_j\\
&-\frac1p\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big]\Big\}\hat{\gamma}_{j,t}(z)\Big],
\end{aligned}\tag{7.23}$$

$$T_{n,2}(z,t) = -\frac{1}{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\beta_{j,t}(z)\right]\mathbb{E}\Big[\beta_{j,t}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\Big(\frac{\lfloor nt\rfloor}{n}\mathbb{E}\tilde{\underline{s}}_{n,t}(z)\mathbf{T}_n+\mathbf{I}\Big)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j\Big], \tag{7.24}$$

For this argument we used the fact

$$\begin{aligned}
\mathbb{E}\Big[\overline{\beta}_{j,t}(z)\Big\{&\mathbf{q}_j^\star\mathbf{T}_n^{\frac12}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac12}\mathbf{q}_j\\
&-\frac1p\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big]\Big\}\Big]=0
\end{aligned}$$

and that by the estimate (9.10.2) in Bai and Silverstein (2010)

$$\mathbb{E}\Big|\overline{\beta}_{j,t}^2(z)\beta_{j,t}(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac12}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac12}\mathbf{q}_j$$

$$
\begin{aligned}
&- \frac{1}{p}\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big]\Big\}\hat{\gamma}_{j,t}(z)^2\Big| \\
\leq & \mathbb{E}^{\frac{1}{2}}\Big|\overline{\beta}_{j,t}^2(z)\beta_{j,t}(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j \\
&- \frac{1}{p}\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big]\Big|^2\mathbb{E}^{\frac{1}{2}}|\hat{\gamma}_{j,t}(z)|^4\Big| \\
\leq & Kn^{-1}\eta_n^2 = o\left(n^{-1}\right).
\end{aligned}
$$

For the term in (7.23) we obtain the representation

$$
\begin{aligned}
T_{n,1}(z,t) =& y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}\Big[\overline{\beta}_{j,t}^2(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j \\
&- \frac{1}{p}\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big\}\hat{\gamma}_{j,t}(z)\Big] \\
&- y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}\Big[\overline{\beta}_{j,t}^2(z)\frac{1}{p}\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n\mathbb{E}[\mathbf{D}_{j,t}^{-1}(z)]\hat{\gamma}_{j,t}(z)\Big] \\
&+ y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}\Big[\overline{\beta}_{j,t}^2(z)\frac{1}{p}\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\hat{\gamma}_{j,t}(z)\Big] \\
=& y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}\Big[\overline{\beta}_{j,t}^2(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j \\
&- \frac{1}{p}\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big\}\hat{\gamma}_{j,t}(z)\Big] \\
=& y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor}z^2\underline{\tilde{s}}_t^2(z)\mathbb{E}\Big[\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j \\
&- \frac{1}{p}\operatorname{tr}(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big\}\hat{\gamma}_{j,t}(z)\Big] + o(1),
\end{aligned}
$$

where in the last step we used the inequality (9.10.2) in Bai and Silverstein (2010), to replace all of the terms $\beta_{j,t}(z), \overline{\beta}_{j,t}(z), b_{j,t}(z)$ and similarly defined quantities by $-z\underline{\tilde{s}}_t(z)$. This argument also implies for the term $T_{n,2}$ defined in (7.24)

$$
T_{n,2}(z,t) = -\frac{z^2\underline{\tilde{s}}_t^2(z)}{\lfloor nt \rfloor n}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}\Big[\operatorname{tr}\mathbf{D}_{j,t}^{-1}(z)\Big(\frac{\lfloor nt \rfloor}{n}\mathbb{E}\underline{\tilde{s}}_{n,t}(z)\mathbf{T}_n + \mathbf{I}\Big)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\Big] + o(1).
$$

We now consider the complex case, where we have from equation (9.8.6) in Bai and Silverstein (2010)

$$
T_{n,1}(z,t) = \frac{z^2\underline{\tilde{s}}_t^2(z)}{\lfloor nt \rfloor n}\sum_{j=1}^{\lfloor nt \rfloor}\mathbb{E}\Big[\operatorname{tr}\mathbf{D}_{j,t}^{-1}(z)\Big(\frac{\lfloor nt \rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I}\Big)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\Big] + o(1).
$$

which yields $T_{n,1}(z,t)+T_{n,2}(z,t)=o(1)$, and as a consequence (6.53) in this case.

Next, we consider the real case using again equation (9.8.6) in Bai and Silverstein (2010), which gives

$$\begin{aligned}
&\lfloor nt\rfloor R_{n,t}(z)\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)] = T_{n,1}(z,t)+T_{n,2}(z,t)+o(1)\\
&=\frac{z^2\tilde{\underline{s}}_t^2(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\operatorname{tr}\mathbf{D}_{j,t}^{-1}(z)\Big(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I}\Big)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\Big]+o(1)\\
&=\frac{z^2\tilde{\underline{s}}_t^2(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\big[\operatorname{tr}\mathbf{D}_{j,t}^{-1}(z)\big(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\big)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\big]+o(1). \qquad (7.25)
\end{aligned}$$

For a detailed analysis of the random variable in (7.25), we recall the decomposition of the resolvent $\mathbf{D}_j^{-1}(z)$ given in (7.42). In the following, investigate these terms appearing in this representation in more detail.

Using the decomposition given in (6.25), the estimates (6.32) and (6.33) (which shows that all terms involving $\mathbf{B}_t(z)$ and $\mathbf{C}_t(z)$ are negligible) and the fact

$$\mathbb{E}\big[\operatorname{tr}\big(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\big)^{-1}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{A}_t(z)\mathbf{T}_n\big]=0,$$

we obtain

$$\begin{aligned}
&y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\\
=&\frac{z^2\tilde{\underline{s}}_t^2(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\operatorname{tr}\Big(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\Big)^{-1}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\\
&\times\Big(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\Big)^{-1}\mathbf{T}_n\Big]\\
&+\frac{z^2\tilde{\underline{s}}_t^2(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}b_{j,t}^2(z)\mathbb{E}\left[\operatorname{tr}\mathbf{A}_t(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{A}_t(z)\mathbf{T}_n\right]+o(1)\\
=&\frac{\tilde{\underline{s}}_t^2(z)}{n}\mathbb{E}\left[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-3}\mathbf{T}_n^2\right]\\
&+\frac{z^4\tilde{\underline{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\operatorname{tr}\mathbf{A}_t(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{A}_t(z)\mathbf{T}_n\right]+o(1). \qquad (7.26)
\end{aligned}$$

For the term $\mathbf{A}_t(z)$ in (6.26) (which actually depends on j) we have

$$\begin{aligned}
\mathbf{A}_t(z)&=\sum_{i\neq j,1\leq i\leq\lfloor nt\rfloor}\Big(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\Big)^{-1}\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\mathbf{D}_{i,j,t}^{-1}(z)\\
&=\sum_{i\neq j,1\leq i\leq\lfloor nt\rfloor}\mathbf{D}_{i,j,t}^{-1}(z)\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\Big(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\Big)^{-1},
\end{aligned}$$

which follows from $\mathbf{A}_t(z) = (\mathbf{A}_t(\overline{z}))^\star$. Substituting the first and second expression for the term $\mathbf{A}_t(z)$ on the left and on the right in (7.26), respectively, yields

$$\begin{aligned}
&\frac{z^4\underline{\tilde{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\operatorname{tr}\mathbf{A}_t(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{A}_t(z)\mathbf{T}_n\right]\\
=&\frac{z^4\underline{\tilde{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i,l\neq j}\mathbb{E}\Big[\operatorname{tr}\left(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\right)^{-1}\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\mathbf{D}_{i,j,t}^{-1}(z)\\
&\times\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star-n^{-1}\mathbf{T}_n\right)\left(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{l,t}(z)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\Big]\\
=&\frac{z^2\underline{\tilde{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i,l\neq j}A_{i,l,j}(z,t)+o(1),
\end{aligned}\tag{7.27}$$

where

$$\begin{aligned}
A_{i,l,j}(z,t)=\mathbb{E}\Big[&\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\mathbf{D}_{i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\\
&\times\mathbf{D}_{l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star-n^{-1}\mathbf{T}_n\right)\Big].
\end{aligned}$$

In the following, we will show that the sum of the cross terms $A_{i,l,j}(z,t)$ (i.e., $l\neq i$) in (7.27) vanishes asymptotically. For this purpose we use the following formula for $l\neq i$

$$\mathbf{D}_{i,j,t}^{-1}(z)=\mathbf{D}_{l,i,j,t}^{-1}(z)-\beta_{l,i,j,t}(z)\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z),$$

where

$$\beta_{l,i,j,t}(z)=\frac{1}{1+\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l}.$$

Note that the expectation appearing in the cross term $A_{i,l,j}(z,t)$ will be 0 if $\mathbf{D}_{i,j,t}^{-1}(z)$ or $\mathbf{D}_{l,j,t}^{-1}(z)$ are replaced by $\mathbf{D}_{l,i,j,t}^{-1}(z)$. Hence, it remains to bound for $i\neq l$ (use also (7.22))

$$\begin{aligned}
&|A_{i,l,j}(z,t)|\\
=&\Big|\mathbb{E}\Big[\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\left(\mathbf{D}_{l,i,j,t}^{-1}(z)-\mathbf{D}_{i,j,t}^{-1}(z)\right)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\\
&\times\mathbf{T}_n\left(\mathbf{D}_{i,l,j,t}^{-1}(z)-\mathbf{D}_{l,j,t}^{-1}(z)\right)\left(\mathbf{r}_l\mathbf{r}_l^\star-n^{-1}\mathbf{T}_n\right)\Big]\Big|\\
=&\Big|\mathbb{E}\Big[\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\beta_{l,i,j,t}(z)\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\\
&\times\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\beta_{i,l,j,t}(z)\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star-n^{-1}\mathbf{T}_n\right)\Big]\Big|\\
=&o\left(n^{-1}\right),
\end{aligned}$$

which is shown in Lemma 7.7.7 and corrects a wrong statement on p. 260 in the monograph of Bai and Silverstein (2010).

Summarizing, we have shown that

$$
\begin{aligned}
& y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)] \\
=&\frac{\tilde{\underline{s}}_t^2(z)}{n}\mathbb{E}\left[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-3}\mathbf{T}_n^2\right] \\
&+\frac{z^2\tilde{\underline{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i\neq j}\mathbb{E}\Big[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n \\
&\times\mathbf{D}_{i,j,t}^{-1}(z)\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\Big]+o(1) \\
=&\frac{\tilde{\underline{s}}_t^2(z)}{n}\mathbb{E}\left[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-3}\mathbf{T}_n^2\right] \\
&+\frac{z^2\tilde{\underline{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i\neq j}\mathbb{E}\Big[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\Big] \\
&+o(1) \\
=&\frac{\tilde{\underline{s}}_t^2(z)}{n}\mathbb{E}\left[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-3}\mathbf{T}_n^2\right] \\
&+\frac{z^2\tilde{\underline{s}}_t^4(z)}{\lfloor nt\rfloor n^3}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i\neq j}\mathbb{E}\Big[\operatorname{tr}\left\{\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\right\}\operatorname{tr}\left\{\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{T}_n\right\}\Big] \\
&+o(1) \\
=&\frac{\tilde{\underline{s}}_t^2(z)}{n}\mathbb{E}\left[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-3}\mathbf{T}_n^2\right] \\
&+\frac{z^2\tilde{\underline{s}}_t^4(z)}{n^3}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\operatorname{tr}\left\{\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\right\} \qquad (7.28) \\
&\times\operatorname{tr}\left\{\mathbf{D}_{j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\right\}\Big]+o(1). \qquad (7.29)
\end{aligned}
$$

Here we used for the last estimate the fact

$$
\begin{aligned}
&\Big|\mathbb{E}\Big[\operatorname{tr}\left\{\mathbf{D}_{j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\right\} \\
&-\operatorname{tr}\left\{\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{T}_n\right\}\Big]\Big| \\
\leq&\mathbb{E}\Big|\operatorname{tr}\left(\mathbf{D}_{i,j,t}^{-1}(z)-\mathbf{D}_{j,t}^{-1}(z)\right)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\Big| \\
&+\mathbb{E}\Big|\operatorname{tr}\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\left(\mathbf{D}_{i,j,t}^{-1}(z)-\mathbf{D}_{j,t}^{-1}(z)\right)\mathbf{T}_n\Big| \\
=&\mathbb{E}\Big|\beta_{i,j,t}(z)\mathbf{r}_i^\star\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{r}_i\Big| \\
&+\mathbb{E}\Big|\beta_{i,j,t}(z)\mathbf{r}_i^\star\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{r}_i\Big| \\
\leq&K+\mathbb{E}\Big|\beta_{i,j,t}(z)\mathbf{r}_i^\star\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n
\end{aligned}
$$

$$\times \left(\mathbf{D}_{i,j,t}^{-1}(z) - \beta_{i,j,t}(z)\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,j,t}^{-1}(z)\right)\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{r}_i\Big|$$
$$\leq K.$$

Hence,

$$\Big|\frac{z^2\underline{\tilde{s}}_t^4(z)}{\lfloor nt\rfloor n^3}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i\neq j}\operatorname{tr}\left\{(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I})^{-2}\mathbf{T}_n^2\right\}\mathbb{E}\Big[\operatorname{tr}\left\{\mathbf{D}_{j,t}^{-1}(z)\,(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n\right\}$$
$$-\operatorname{tr}\left\{\mathbf{D}_{i,j,t}^{-1}(z)\,(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{T}_n\right\}\Big]\Big| = o(1).$$

We now apply (7.25) for (7.29) and obtain

$$y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]$$
$$=\frac{\underline{\tilde{s}}_t^2(z)}{n}\mathbb{E}\left[\operatorname{tr}(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I})^{-3}\mathbf{T}_n^2\right]$$
$$+\frac{\underline{\tilde{s}}_t^2(z)\lfloor nt\rfloor}{n^2}\operatorname{tr}\left\{(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I})^{-2}\mathbf{T}_n^2\right\}y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]+o(1).$$

This implies (6.53) for the real case, namely,

$$y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)] = \frac{\frac{\underline{\tilde{s}}_t^2(z)}{n}\operatorname{tr}\left\{(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I})^{-3}\mathbf{T}_n^2\right\}}{1-\frac{\underline{\tilde{s}}_t^2(z)\lfloor nt\rfloor}{n^2}\operatorname{tr}\left\{(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I})^{-2}\mathbf{T}_n^2\right\}}+o(1)$$
$$=\frac{y\int\frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^3}dH(\lambda)}{1-ty\int\frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^2}dH(\lambda)}+o(1).$$

7.7 Further auxiliary results

Lemma 7.7.1 *It holds*

$$\frac{y_n\frac{\lfloor nt\rfloor}{n}\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)|^2}}{\left|-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)}\right|^2} = \frac{\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\underline{\tilde{s}}_{n,t}^0(z))y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)|^2}}{\operatorname{Im}(z)+\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\underline{\tilde{s}}_{n,t}^0(z))y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\underline{\tilde{s}}_{n,t}^0(z)|^2}}$$

and

$$\frac{y_n\frac{\lfloor nt\rfloor}{n}\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}}{\left|-z+y_n\int\frac{\lambda dH_n(\lambda)}{1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}+R_{n,t}\right|^2}$$
$$=\frac{\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}}{\operatorname{Im}(z)+\frac{\lfloor nt\rfloor}{n}\operatorname{Im}(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)])y_n\int\frac{\lambda^2dH_n(\lambda)}{|1+\lambda\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}+\operatorname{Im}(R_{n,t})},$$

where $R_{n,t}$ is defined in Lemma 6.3.6.

Proof. We only show the first assertion, since the second one can be shown in a similar way. Applying

$$\operatorname{Im}\left(\frac{1}{z}\right) = \frac{-\operatorname{Im}(z)}{|z|^2}$$

to (7.7) yields

$$\operatorname{Im}(\tilde{\underline{s}}^0_{n,t}(z)) = \frac{\operatorname{Im}(z) + \frac{\lfloor nt \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{n,t}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)|^2}}{\left| -z + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)} \right|^2}.$$

We conclude

$$\begin{aligned}
\frac{\frac{\lfloor nt \rfloor}{n} y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)|^2}}{\left| -z + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)} \right|^2} &= 1 - \frac{\operatorname{Im}(z)}{\left| -z + y_n \int \frac{\lambda dH_n(\lambda)}{1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)} \right|^2 \operatorname{Im}(\tilde{\underline{s}}^0_{n,t}(z))} \\
&= 1 - \frac{\operatorname{Im}(z)}{\operatorname{Im}(z) + \frac{\lfloor nt \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{\lfloor nt \rfloor}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)|^2}} \\
&= \frac{\frac{\lfloor nt \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{n,t}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)|^2}}{\operatorname{Im}(z) + \frac{\lfloor nt \rfloor}{n} \operatorname{Im}(\tilde{\underline{s}}^0_{n,t}(z)) y_n \int \frac{\lambda^2 dH_n(\lambda)}{|1+\lambda \frac{\lfloor nt \rfloor}{n} \tilde{\underline{s}}^0_{n,t}(z)|^2}}.
\end{aligned}$$

□

Lemma 7.7.2

$$\sup_{n\in\mathbb{N}, z\in\mathcal{C}_n, t\in[t_0,1]} \left\| \left(\frac{\lfloor nt \rfloor}{n} \mathbb{E}[\tilde{\underline{s}}_{n,t}(z)] \mathbf{T}_n + \mathbf{I}\right)^{-1} \right\| \leq K.$$

Proof of Lemma 7.7.2. By Lemma 6.10 (a) in Bai and Silverstein (2010), we have for $z \in \mathcal{C}_u, t \in [t_0, 1]$

$$\left\| \left(\frac{\lfloor nt \rfloor}{n} \mathbb{E}[\tilde{\underline{s}}_{n,t}(z)] \mathbf{T} + \mathbf{I} \right)^{-1} \right\| \leq \max\left(4 \frac{\lfloor nt \rfloor}{n v_0} ||\mathbf{T}||, 2 \right) \leq \max\left(\frac{4}{v_0}, 2 \right).$$

Thus, the assertion holds for $z \in \mathcal{C}_u \times [t_0, 1] \subset \mathcal{C}_n$. Let $x \in \{x_l, x_r\}$. Since $x/t \notin \Gamma_{\underline{F}^{y_t,H}}$ for $t \in [t_0, 1]$, it follows from Lemma 6.1 in Bai and Silverstein (2010),

$$-\left(\underline{s}_t\left(\frac{x}{t}\right)\right)^{-1} \notin \Gamma_H, \text{ so } -(t\tilde{\underline{s}}_t(x))^{-1} \notin \Gamma_H$$

and thus, for any $\lambda \in \Gamma_H$

$$t\tilde{\underline{s}}_t(x)\lambda + 1 \neq 0.$$

Let $\lambda_0 \in \Gamma_H$. Then, there exist $\delta_1, \mu_1 > 0$ such that

$$\inf_{\substack{z\in\mathcal{C}^0,\\ t\in[t_0,1]}} |t\underline{\tilde{s}}_t(z)\lambda_0 + 1| > \delta_1, \tag{7.30}$$

$$\sup_{\substack{z\in\mathcal{C}^0,\\ t\in[t_0,1]}} |t\underline{\tilde{s}}_t(z)| < \mu_1, \tag{7.31}$$

where

$$\mathcal{C}^0 = \{x + iv : v \in [0, v_0]\}.$$

The boundedness condition given in (7.31) is clear, and we continue by proving (7.30). Assume that,

$$\inf_{\substack{z\in\mathcal{C}^0,\\ t\in[t_0,1]}} |t\underline{\tilde{s}}_t(z)\lambda_0 + 1| = 0.$$

Then, there exists a sequence $(z_n, t_n)_{n\in\mathbb{N}}$ in $\mathcal{C}^0 \times [t_0, 1]$ such that

$$t_n\underline{\tilde{s}}_{t_n}(z_n)\lambda_0 + 1 \to 0,\ \ n \to \infty,$$

which contains a further subsequence $(z_{n_k}, t_{n_k})_{n\in\mathbb{N}}$ converging to $(z^\star, t^\star) \in \mathcal{C}^0 \times [t_0, 1]$. By continuity of $(z, t) \mapsto \underline{s}_t(z)$ on $\mathcal{C}^0 \times [t_0, 1]$ (see proof of Theorem 6.3.7), we have

$$t_{n_k}\underline{\tilde{s}}_{t_{n_k}}(z_{n_k})\lambda_0 + 1 \to t^\star\underline{\tilde{s}}_{t^\star}(z^\star)\lambda_0 + 1,\ \ k \to \infty,$$

yielding the contradiction

$$0 = t^\star\underline{\tilde{s}}_{t^\star}(z^\star)\lambda_0 + 1.$$

Thus, (7.30) and (7.31) hold true. Since $H_n \xrightarrow{\mathcal{D}} H$ and hence $\lambda_0 \in \Gamma_{H_n}$ for sufficiently large n, there exists an eigenvalue $\lambda(\mathbf{T}_n)$ of $\mathbf{T}_n$ such that

$$|\lambda(\mathbf{T}_n) - \lambda_0| < \frac{\delta_1}{4\mu_1}.$$

Recalling Theorem 6.3.7, we also have for sufficiently large n

$$\sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} |\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] - \underline{\tilde{s}}_t(z)| < \frac{\delta_1}{4}.$$

Since

$$\begin{aligned} |\underline{\tilde{s}}_t(z)\lambda_0 + 1| &\le |\underline{\tilde{s}}_t(z)||\lambda^{\mathbf{T}} - \lambda_0| + |t\underline{\tilde{s}}_t(z)\lambda^{\mathbf{T}} + 1| \\ &\le |\underline{\tilde{s}}_t(z)||\lambda^{\mathbf{T}} - \lambda_0| + t|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] - \underline{\tilde{s}}(z)||\lambda^{\mathbf{T}}| + |t\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\lambda^{\mathbf{T}} + 1| \end{aligned}$$

we get a contradiction to (7.30) if

$$\inf_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} \left|t\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\lambda^{\mathbf{T}} + 1\right| \le \frac{\delta_1}{4}$$

would hold true. This proves Lemma 7.7.2. $\square$

Lemma 7.7.3 *We have uniformly in $n \in \mathbb{N}, t \in [t_0, 1], z \in \mathcal{C}_n$*

$$\mathbb{E}||\mathbf{D}_t^{-1}(z)||^q \leq K, \tag{7.32}$$

where $K > 0$ is a constant depending on $q \in \mathbb{N}$. Similarly, for pairwise different integers $i, j, k \in \{1, \ldots, \lfloor nt \rfloor\}$,

$$\max\left(\mathbb{E}||\mathbf{D}_t^{-1}(z)||^q, \mathbb{E}||\mathbf{D}_{j,t}^{-1}(z)||^q, \mathbb{E}||\mathbf{D}_{j,k,t}^{-1}(z)||^q\right) \leq K.$$

It also holds that

$$||\mathbf{D}_t^{-1}(z)|| \leq K + n\varepsilon_n^{-1} I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) \leq \eta_{l,t}\}, \tag{7.33}$$

where $\eta_{r,t}$ denotes a fixed number between

$$\limsup_{n\to\infty} ||\mathbf{T}_n||(1 + \sqrt{y_t})^2 t$$

and x_r and $\eta_{l,t}$ between

$$\liminf_{n\to\infty} \lambda_p(\mathbf{T}_n)(1 - \sqrt{y_t})^2 I_{(0,1)}(y_t) t$$

and x_l.

Proof of Lemma 7.7.3. We begin with a proof of (7.33). Let first $z \in \mathcal{C}_u$, that is, $z = x + iv_0$ for some $x \in [x_l, x_r]$. Then,

$$||\mathbf{D}_t^{-1}(z)|| = \frac{1}{\min(|\lambda_p(\mathbf{B}_{n,t}) - z|, |\lambda_1(\mathbf{B}_{n,t}) - z|)} \leq \frac{1}{v_0} = K.$$

This implies $\mathbb{E}||\mathbf{D}_t^{-1}(z)||^q \leq K$. Next, assume $z \in \mathcal{C}_l \cup \mathcal{C}_r$, that is, $z = x_r + iv$ or $z = x_l + iv$ for some $v \in [n^{-1}\varepsilon_n, v_0]$. By formula (9.7.8) and (9.7.9) in Bai and Silverstein (2010) we have for $t \in [t_0, 1]$ and any $m > 0$

$$\mathbb{P}\left(||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) < \eta_{l,t}\right) = o\left(\lfloor nt \rfloor^{-m}\right) = o\left(n^{-m}\right). \tag{7.34}$$

We estimate

$$\begin{aligned}\mathbb{E}||\mathbf{D}_t^{-1}(z)||^q \leq & K\mathbb{E}\left[||\mathbf{D}_t^{-1}(z)|| I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_p(\mathbf{B}_{n,t}) \geq \eta_{l,t}\}\right]^q \\ & + K\mathbb{E}\left[||\mathbf{D}_t^{-1}(z)|| I\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) < \eta_{l,t}\}\right]^q\end{aligned}$$

To derive a bound for the first summand, we distinguish the cases $z \in \mathcal{C}_r$ and $z \in \mathcal{C}_l$. For the sake of brevity, we only consider the first one. It holds

$$\begin{aligned}&||\mathbf{D}_t^{-1}(z)|| I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_p(\mathbf{B}_{n,t}) \geq \eta_{l,t}\} \\ =& \frac{I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_p(\mathbf{B}_{n,t}) \geq \eta_{l,t}\}}{\min(|\lambda_p(\mathbf{B}_{n,t}) - (x_r + iv)|, |\lambda_1(\mathbf{B}_{n,t}) - (x_r + iv)|)} \\ \leq & \frac{1}{x_r - \lambda_1(\mathbf{B}_{n,t})} I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_p(\mathbf{B}_{n,t}) \geq \eta_{l,t}\}\end{aligned}$$

$$\leq \frac{1}{x_r - \eta_{r,t}} \leq \frac{1}{x_r - \limsup_{n\to\infty} ||\mathbf{T}_n||(1+\sqrt{y_{t_0}})^2} = K. \tag{7.35}$$

For the second summand, we conclude

$$\begin{aligned}
&||\mathbf{D}_t^{-1}(z)||I\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) < \eta_{l,t}\} \\
\leq & \frac{1}{\min(|\lambda_p(\mathbf{B}_{n,t}) - z|, |\lambda_1(\mathbf{B}_{n,t}) - z|)} I\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) < \eta_{l,t}\} \\
\leq & n\varepsilon_n^{-1} I\left\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) < \eta_{l,t}\right\}.
\end{aligned} \tag{7.36}$$

The bounds in (7.35) and (7.36) show that (7.33) holds true. The assertion in (7.32) follows by applying (7.34). □

Proof of Lemma 7.2.1

The proof of Lemma 7.2.1 requires some preparations. Note that while a fourth moment condition is sufficient for proving the convergence of the finite-dimensional distribution of $(\hat{M}_n^1)_{n\in\mathbb{N}}$ (Theorem 6.3.1) and the convergence of the non-random part $(\hat{M}_n^2)_{n\in\mathbb{N}}$ (Theorem 6.3.3), we need the stronger moment assumption from Theorem 3.2.1, namely

$$\sup_{i,j,n} \mathbb{E}|x_{ij}|^{12} < \infty, \tag{7.37}$$

exclusively for a proof of the asymptotic tightness of $(\hat{M}_n^1)_{n\in\mathbb{N}}$.
Under this assumption, by Lemma B.26 in Bai and Silverstein (2010), the following estimates for moments of quadratic forms hold true for $q \geq 2$

$$\begin{aligned}
\mathbb{E}|\mathbf{x}_j^\star \mathbf{A}\mathbf{x}_j - \operatorname{tr}\mathbf{A}|^q &\lesssim (\operatorname{tr}\mathbf{A}\mathbf{A}^\star)^{q/2} + \eta_n^{(2q-12)\vee 0} n^{(q-6)\vee 0} \operatorname{tr}(\mathbf{A}\mathbf{A}^\star)^{q/2} \\
&\lesssim \begin{cases} (\operatorname{tr}\mathbf{A}\mathbf{A}^\star)^{q/2}\left(1 + n^{(q-6)\vee 0}\right), \\ n^{q/2}||\mathbf{A}||^q + n n^{(q-6)\vee 0}||\mathbf{A}||^q. \end{cases}
\end{aligned}$$

Thus, we have for $q \geq 2$

$$\mathbb{E}|\mathbf{r}_j^\star \mathbf{A}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}|^q \lesssim \begin{cases} (\operatorname{tr}\mathbf{A}\mathbf{A}^\star)^{q/2} n^{-(q\wedge 6)}, \\ ||\mathbf{A}||^q n^{-((q/2)\wedge 5)}. \end{cases} \tag{7.38}$$

Furthermore, combining (7.38) with arguments given in the proof of (9.9.6) in Bai and Silverstein (2010), we obtain the following lemma.

Lemma 7.7.4 *Let $j, m \in \mathbb{N}_0$, $q \geq 2$ and $\mathbf{A}_l$, $l \in \{1, \dots, m+1\}$ be $p \times p$ (random) matrices independent of $\mathbf{r}_j$ which obey for any $\tilde{q} \geq 2$*

$$\mathbb{E}||\mathbf{A}_l||^{\tilde{q}} < \infty, \ l \in \{1, \dots, m+1\}.$$

Then, it holds

$$\mathbb{E}\Big|\Big(\prod_{k=1}^m \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big)\left(\mathbf{r}_j^\star \mathbf{A}_{m+1}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m+1}\right)\Big|^q \lesssim n^{-((q/2)\wedge 5)}.$$

If even for any $l \in \{1, \dots, m+1\}$, $\tilde{q} \geq 2$

$$\mathbb{E}\,[\operatorname{tr} \mathbf{A}\mathbf{A}_l^\star]^{\tilde{q}} < \infty,$$

holds true, then we have

$$\mathbb{E}\Big|\Big(\prod_{k=1}^{m} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big)\left(\mathbf{r}_j^\star \mathbf{A}_{m+1}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m+1}\right)\Big|^q \lesssim n^{-(q\wedge 6)}.$$

Remark 7.7.5 In fact, as the proof of Lemma 7.7.4 reveals, we could impose a less restrictive condition on the spectral moments of $\mathbf{A}_l$, $l \in \{1, \dots, m+1\}$. For our purpose, it is sufficient to state the previous lemma in this form, since, when applying Lemma 7.7.4, the involved matrices will have bounded spectral moments of any order.
In particular, the second assertion will be useful if $\mathbf{B}_l$ involves a term like $\mathbf{r}_k\mathbf{r}_k^\star$ for some $k \neq j$ among other matrices like $\mathbf{D}_{k,j,t}^{-1}(z)$, while we will make use of the first assertion in the case that $\mathbf{B}_l$ only involves matrices like $\mathbf{D}_{j,t}^{-1}(z)$. In the latter case, contrary to the first one, we are not able to control moments of $\operatorname{tr}\mathbf{B}_l\mathbf{B}_l^\star$ uniformly in n.

Proof of Lemma 7.7.4. For $m = 0$, the assertion of the lemma follows directly from (7.38) for any $q \geq 2$. We continue the proof by an induction over the integer m for some fixed $q \geq 2$.

$$\begin{aligned}
&\mathbb{E}\Big|\Big(\prod_{k=1}^{m} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big)\left(\mathbf{r}_j^\star \mathbf{A}_{m+1}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m+1}\right)\Big|^q \\
\lesssim&\mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big)\left(\mathbf{r}_j^\star \mathbf{A}_{m}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m}\right)\left(\mathbf{r}_j^\star \mathbf{A}_{m+1}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m+1}\right)\Big|^q \\
&+ \mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big)n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_m\left(\mathbf{r}_j^\star \mathbf{A}_{m+1}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m+1}\right)\Big|^q \\
\leq&\Big(\mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big)\left(\mathbf{r}_j^\star \mathbf{A}_{m}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m}\right)\Big|^{2q}\mathbb{E}\left|\left(\mathbf{r}_j^\star \mathbf{A}_{m+1}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m+1}\right)\right|^{2q}\Big)^{\frac{1}{2}} \\
&+ \mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big)n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_m\left(\mathbf{r}_j^\star \mathbf{A}_{m+1}\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_{m+1}\right)\Big|^q.
\end{aligned}$$

By applying the induction hypothesis to these three terms, we get the desired result for each case. □

Adapting the proof of (9.10.5) in Bai and Silverstein (2010), we obtain under the strong moment condition (7.37) for $q \geq 2$

$$\mathbb{E}|\gamma_{j,t}(z)|^q \lesssim n^{-((q/2)\wedge 5)}. \tag{7.39}$$

We need an estimate for moments of complex martingale difference schemes. We refer to Lemma 2.1 in Li (2003), which is a corollary from Burkholder's inequality and can easily be extended to the complex case. We are now in the position to give a proof of Lemma 7.2.1.

Proof of Lemma 7.2.1. In the following, we will often make use of the decompositions

$$\mathbf{D}_t^{-1}(z) = \mathbf{D}_{j,t}^{-1}(z) - \beta_{j,t}(z)\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z), \qquad (7.40)$$
$$\beta_{j,t}(z) = b_{j,t}(z) - \beta_{j,t}(z)b_{j,t}(z)\gamma_{j,t}(z).$$

Observing the decomposition (7.11), our aim is to show the inequalities in (7.9) and (7.10), where we assume $t_2 > t_1$ w.l.o.g.

Step 1: *Analysis of* Z_n^2
Beginning with the proof of (7.10) for Z_n^2, we are able to show that (using Lemma 2.1 in Li (2003) with $q = 4 + \delta$)

$$\begin{aligned}\mathbb{E}|Z_n^2(z,t_1,t_2)|^{4+\delta} &= \mathbb{E}\Big|\sum_{j=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\Big|^{4+\delta}\\
&\lesssim (\lfloor nt_2\rfloor - \lfloor nt_1\rfloor)^{1+\delta/2}\sum_{j=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}\mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\right|^{4+\delta}\\
&\lesssim \Big(\frac{\lfloor nt_2\rfloor - \lfloor nt_1\rfloor}{n}\Big)^{2+\delta/2},\end{aligned}$$

since we can bound

$$\begin{aligned}&\mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\right|^{4+\delta}\\
&\lesssim \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\right|^{4+\delta}\\
&+ \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\gamma_{j,t_2}(z)\right|^{4+\delta}\\
&\lesssim \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\left\{\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{j,t_2}^{-2}(z)\right\}\right|^{4+\delta}\\
&+ \mathbb{E}\left|\beta_{j,t_2}(z)b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\gamma_{j,t_2}(z)\right|^{4+\delta} \qquad (7.41)\\
&\lesssim n^{-(2+\delta/2)}.\end{aligned}$$

We should explain the bound for (7.41) in more detail: First, note that we are able to bound the moments of $||\mathbf{D}_{j,t}^{-1}(z)||$ independent of n, z, t (see Lemma 7.7.3). As a further preparation, we observe for $z \in \mathcal{C}_n, t \in [t_0, 1]$ from Lemma 7.7.3

$$\begin{aligned}||\mathbf{D}_t^{-1}(z)|| &\lesssim 1 + n\varepsilon_n^{-1}I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) \leq \eta_{l,t}\}\\
&\leq 1 + n^2 I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) \leq \eta_{l,t}\}, \qquad (7.42)\end{aligned}$$

where we used the fact that $\varepsilon_n \geq n^{-\alpha}$ for some $\alpha \in (0,1)$. Thus, since $|\mathbf{r}_j|^2 \leq n$, we obtain

$$\begin{aligned}|\beta_{j,t}(z)| &= |1 - \mathbf{r}_j^\star\mathbf{D}_t^{-1}(z)\mathbf{r}_j| \leq 1 + |\mathbf{r}_j|^2||\mathbf{D}_t^{-1}(z)||\\
&\lesssim 1 + |\mathbf{r}_j|^2 + n^3 I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}) \leq \eta_{l,t}\}. \qquad (7.43)\end{aligned}$$

It is easy to see that the inequality (9.10.6) in Bai and Silverstein (2010) also holds for $\beta_{j,t}(z)$ and by the same arguments following (9.10.6), we obtain

$$|b_{j,t}(z)| \lesssim 1. \tag{7.44}$$

Similarly to these bounds, using (7.33) in Lemma 7.7.3 for the matrix $\mathbf{D}_{j,t}^{-1}(z)$, we get for any $m \geq 1$

$$\begin{aligned}
|\gamma_{j,t}(z)| =&|\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j - n^{-1}\mathbb{E}[\operatorname{tr}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)]| \lesssim |\mathbf{r}_j|^2||\mathbf{D}_{j,t}^{-1}(z)|| + \mathbb{E}||\mathbf{D}_{j,t}^{-1}(z)||\\
\lesssim&|\mathbf{r}_j|^2 + |\mathbf{r}_j|^2 n\varepsilon_n^{-1} I\{||\mathbf{B}_{n,t}^{(-j)}|| \geq \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}^{(-j)}) \leq \eta_{l,t}\}\\
&+ |\mathbf{r}_j|^2 n\varepsilon_n^{-1}\mathbb{P}\{||\mathbf{B}_{n,t}^{(-j)}|| \geq \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}^{(-j)}) \leq \eta_{l,t}\}\\
\leq&|\mathbf{r}_j|^2 + n^3 I\{||\mathbf{B}_{n,t}^{(-j)}|| \geq \eta_{r,t} \text{ or } \lambda_p(\mathbf{B}_{n,t}^{(-j)}) \leq \eta_{l,t}\} + o\left(n^{-m}\right),
\end{aligned}$$

where we used the fact that for any $m > 0$

$$\begin{aligned}
&\mathbb{P}\{||\mathbf{B}_{n,t_2}^{(-j)}|| \geq \eta_{r,t_2} \text{ or } \lambda_p(\mathbf{B}_{n,t_2}^{(-j)}) \leq \eta_{l,t_2}\} = o\left(n^{-m}\right),\\
&\mathbb{P}\{||\mathbf{B}_{n,t_2}|| \geq \eta_{r,t_2} \text{ or } \lambda_p(\mathbf{B}_{n,t_2}) \leq \eta_{l,t_2}\} = o\left(n^{-m}\right)
\end{aligned} \tag{7.45}$$

and the notation

$$\mathbf{B}_{n,t}^{(-j)} = \mathbf{B}_{n,t} - \mathbf{r}_j\mathbf{r}_j^\star.$$

Using (7.38) and (7.39), we can also bound

$$\begin{aligned}
\mathbb{E}\Big||\mathbf{r}_j|^2\gamma_{j,t}(z)\Big|^{4+\delta} =&\mathbb{E}|\mathbf{r}_j^\star\mathbf{r}_j\gamma_{j,t}(z)|^{4+\delta} \lesssim \mathbb{E}|\left(\mathbf{r}_j^\star\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\right)\gamma_{j,t}(z)|^{4+\delta} + \mathbb{E}|n^{-1}\operatorname{tr}(\mathbf{T}_n)\gamma_{j,t}(z)|^{4+\delta}\\
\leq& \left(\mathbb{E}|\mathbf{r}_j^\star\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n|^{8+2\delta}\mathbb{E}|\gamma_{j,t}(z)|^{8+2\delta}\right)^{\frac{1}{2}} + \mathbb{E}|\gamma_{j,t}(z)|^{4+\delta} \lesssim n^{-(2+\delta/2)}.
\end{aligned}$$

By induction, one can show for some $q \in \mathbb{N}_0$ and $\delta \geq 0$

$$\mathbb{E}\Big||\mathbf{r}_j|^{2q}\gamma_{j,t}(z)\Big|^{4+\delta} \lesssim n^{-(2+\delta/2)}. \tag{7.46}$$

Combining these inequalities, we conclude

$$\begin{aligned}
&\mathbb{E}\left|\beta_{j,t_2}(z)b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\gamma_{j,t_2}(z)\right|^{4+\delta}\\
\lesssim&\mathbb{E}\Big|\Big(1 + |\mathbf{r}_j|^2 + n^3 I\{||\mathbf{B}_{n,t_2}|| \geq \eta_{r,t_2} \text{ or } \lambda_p(\mathbf{B}_{n,t_2}) \leq \eta_{l,t_2}\}\Big)|\mathbf{r}_j|^2\\
&\times\Big(1 + n^2 I\{||\mathbf{B}_{n,t_2}^{(-j)}|| \geq \eta_{r,t_2} \text{ or } \lambda_p(\mathbf{B}_{n,t_2}^{(-j)}) \leq \eta_{l,t_2}\}\Big)^2\gamma_{j,t_2}(z)\Big|^{4+\delta}.
\end{aligned} \tag{7.47}$$

The expectation in (7.47) can now be estimated by multiplying these terms out and using the inequalities (7.44), (7.45) and (7.46).

Thus, we conclude that

$$\mathbb{E}\left|\beta_{j,t_2}(z)b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\gamma_{j,t_2}(z)\right|^{4+\delta} \lesssim n^{-(2+\delta/2)}.$$

Step 2: *Analysis of* $M_n^1(z_1,t) - M_n^1(z_2,t)$
Before investigating the term Z_n^1 in the decomposition (7.11), we show that (7.8) holds true in a similar fashion to the considerations above. We write for $z_1, z_2 \in \mathcal{C}_n, t \in [t_0, 1]$

$$\begin{aligned} M_n^1(z_1,t) - M_n^1(z_2,t) &= \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\left(\mathbf{D}_t^{-1}(z_1) - \mathbf{D}_t^{-1}(z_2)\right) \\ &= \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1})(z_1 - z_2) \operatorname{tr} \mathbf{D}_t^{-1}(z_1)\mathbf{D}_t^{-1}(z_2) \\ &= G_{n1} + G_{n2} + G_{n3}, \end{aligned}$$

where

$$\begin{aligned} G_{n1} &= (z_1 - z_2) \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_{j,t}(z_1)\beta_{j,t}(z_2) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j\right)^2, \\ G_{n2} &= -(z_1 - z_2) \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_{j,t}(z_1) \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-2}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j, \\ G_{n3} &= -(z_1 - z_2) \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_{j,t}(z_2) \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-2}(z_2)\mathbf{D}_{j,t}^{-1}(z_1)\mathbf{r}_j. \end{aligned}$$

The terms G_{n2} and G_{n3} can be estimated using similar arguments as given in the proof of (7.10). More precisely, we obtain for the second term

$$\mathbb{E}|G_{n2}|^{2+\delta} \lesssim |z_1 - z_2|^{2+\delta},$$

and a similar inequality holds for the third term. For the first summand, we have

$$G_{n1} = G_{n11} + G_{n12} + G_{n13},$$

where

$$\begin{aligned} G_{n11} =& (z_1 - z_2) \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) b_{j,t}(z_1) b_{j,t}(z_2) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j\right)^2, \\ G_{n12} =& -(z_1 - z_2) \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) b_{j,t}(z_2) \beta_{j,t}(z_1)\beta_{j,t}(z_2) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j\right)^2 \gamma_{j,t}(z_2), \\ G_{n13} =& -(z_1 - z_2) \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) b_{j,t}(z_1) b_{j,t}(z_2) \beta_{j,t}(z_1) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j\right)^2 \gamma_{j,t}(z_1). \end{aligned}$$

Here, the terms G_{n12} and G_{n13} can be treated by similar arguments as in the derivation of (7.41) using Lemma 2.1 in Li (2003), which gives for $l \in \{1,2\}$

$$\mathbb{E}|G_{n1l}|^{2+\delta} \lesssim |z_1 - z_2|^{2+\delta}.$$

Therefore, it remains to investigate the term G_{n11}:

$$\mathbb{E}|G_{n11}|^{2+\delta} \lesssim |z_1 - z_2|^{2+\delta} n^{\delta/2} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) b_{j,t}(z_1) b_{j,t}(z_2) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \right)^2 \right|^{2+\delta}.$$

We obtain for the summands in $\mathbb{E}|G_{n11}|^{2+\delta}$ observing (7.44)

$$\begin{aligned}
&\mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) b_{j,t}(z_1) b_{j,t}(z_2) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \right)^2 \right|^{2+\delta} \\
\lesssim& \mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \right)^2 \right|^{2+\delta} \\
=& \mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) \left[\left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \right)^2 - \left(n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right)^2 \right] \right|^{2+\delta} \\
=& \mathbb{E} \Big| (\mathbb{E}_j - \mathbb{E}_{j-1}) \Big[\left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right) \\
&\times \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j + n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right) \Big] \Big|^{2+\delta} \\
\lesssim& \mathbb{E} \Big| (\mathbb{E}_j - \mathbb{E}_{j-1}) \Big[\left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right) \\
&\times \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \Big] \Big|^{2+\delta} \\
+& \mathbb{E} \Big| (\mathbb{E}_j - \mathbb{E}_{j-1}) \Big[\left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right) \\
&\times n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \Big] \Big|^{2+\delta} \\
\lesssim& n^{-(1+\delta/2)},
\end{aligned}$$

where we used Lemma 7.7.4 with $q = 2+\delta$ and $m = 1$ and Lemma 7.7.3 for the last inequality. These considerations show that (7.8) holds true.

Step 3: *Analysis of Z_n^1*
Next, we show the estimate (7.9) for the term Z_n^1. Doing so, we will need condition (7.37) on the moments of x_{ij}. For the following calculation, we will write β_t instead of $\beta_t(z)$, $\mathbf{D}_t^{-1}$ instead of $\mathbf{D}_t^{-1}(z)$ and further omit the z-argument for similar quantities. We have for $j \leq \lfloor nt_1 \rfloor$

$$\begin{aligned}
&\beta_{j,t_2} \mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2} \mathbf{r}_j - \beta_{j,t_1} \mathbf{r}_j^\star \mathbf{D}_{j,t_1}^{-2} \mathbf{r}_j = (\beta_{j,t_2} - \beta_{j,t_1}) \mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2} \mathbf{r}_j + \beta_{j,t_1} \mathbf{r}_j^\star \left(\mathbf{D}_{j,t_2}^{-2} - \mathbf{D}_{j,t_1}^{-2} \right) \mathbf{r}_j \\
=& (\beta_{j,t_2} - \beta_{j,t_1}) \mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2} \mathbf{r}_j + \beta_{j,t_1} \mathbf{r}_j^\star \left(\mathbf{D}_{j,t_2}^{-1} - \mathbf{D}_{j,t_1}^{-1} \right) \mathbf{D}_{j,t_1}^{-1} \mathbf{r}_j + \beta_{j,t_1} \mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-1} \left(\mathbf{D}_{j,t_2}^{-1} - \mathbf{D}_{j,t_1}^{-1} \right) \mathbf{r}_j \\
=& \left(\mathbf{r}_j^\star \mathbf{D}_{t_1}^{-1} \mathbf{r}_j - \mathbf{r}_j^\star \mathbf{D}_{t_2}^{-1} \mathbf{r}_j \right) \mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2} \mathbf{r}_j - \beta_{j,t_1} \mathbf{r}_j^\star \mathbf{D}_{j,t_1}^{-1} \Big(\sum_{k=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} \mathbf{r}_k \mathbf{r}_k^\star \Big) \mathbf{D}_{j,t_2}^{-1} \mathbf{D}_{j,t_1}^{-1} \mathbf{r}_j \\
&- \beta_{j,t_1} \mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-1} \mathbf{D}_{j,t_1}^{-1} \Big(\sum_{k=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} \mathbf{r}_k \mathbf{r}_k^\star \Big) \mathbf{D}_{j,t_2}^{-1} \mathbf{r}_j \\
=& \mathbf{r}_j^\star \mathbf{D}_{t_2}^{-1} \Big(\sum_{k=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} \mathbf{r}_k \mathbf{r}_k^\star \Big) \mathbf{D}_{t_1}^{-1} \mathbf{r}_j \mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2} \mathbf{r}_j - \beta_{j,t_1} \mathbf{r}_j^\star \mathbf{D}_{j,t_1}^{-1} \Big(\sum_{k=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} \mathbf{r}_k \mathbf{r}_k^\star \Big) \mathbf{D}_{j,t_2}^{-1} \mathbf{D}_{j,t_1}^{-1} \mathbf{r}_j
\end{aligned}$$

$$- \beta_{j,t_1}\mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-1}\mathbf{D}_{j,t_1}^{-1}\Big(\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}\mathbf{r}_k\mathbf{r}_k^\star\Big)\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j$$

$$= \sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}\Big\{\mathbf{r}_j^\star\mathbf{D}_{t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j - \beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j - \beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j\Big\}.$$

Hence, using the identity (7.40), we obtain the representation

$$\begin{aligned}
&Z_n^1(z,t_1,t_2)\\
&= \sum_{j=1}^{\lfloor nt_1\rfloor}\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j-\mathbb{E}_{j-1})\Big\{-\beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j - \beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j\\
&\quad + \mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j - \beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j\\
&\quad - \beta_{j,t_2}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j\\
&\quad + \beta_{j,t_1}\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j\Big\}.
\end{aligned}$$

We use the substitutions

$$\mathbf{D}_{j,t_2}^{-1} = \mathbf{D}_{k,j,t_2}^{-1} - \beta_{k,j,t_2}\mathbf{D}_{k,j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1} \tag{7.48}$$

and

$$\beta_{j,t} = b_{j,t} - b_{j,t}\beta_{j,t}\gamma_{j,t},\ \ \beta_{k,j,t_2} = b_{k,j,t_2} - b_{k,j,t_2}\beta_{k,j,t_2}\gamma_{k,j,t_2},$$

where $\gamma_{k,j,t}(z) = \mathbf{r}_k^\star\mathbf{D}_{k,j,t}^{-1}(z)\mathbf{r}_k - n^{-1}\mathbb{E}[\mathrm{tr}\,\mathbf{T}_n\mathbf{D}_{k,j,t}^{-1}(z)]$. This yields the representation

$$Z_n^1(z,t_1,t_2) = \sum\sum_{j=1}^{\lfloor nt_1\rfloor}\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j-\mathbb{E}_{j-1})T_{j,k}.$$

Here, the first sum corresponds to the summation with respect to a finite number of different terms $T_{j,k}$, which are of the form

$$\prod_{l_1=1}^{q}\Big(\mathbf{r}_j^\star\mathbf{A}_{l_1}\Big(\prod_{l_2=1}^{q_{l_1}}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{B}_{l_1,l_2}\Big)\mathbf{r}_j\Big),$$

$$(\beta_{j,t_1}\gamma_{j,t_1})^{X_1}\,(\beta_{j,t_2}\gamma_{j,t_2})^{X_2}\prod_{l_1=1}^{q}\Big(\mathbf{r}_j^\star\mathbf{A}_{l_1}\Big(\prod_{l_2=1}^{q_{l_1}}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{B}_{l_1,l_2}\Big)\mathbf{r}_j\Big),$$

$$(\beta_{k,j,t_2}\gamma_{k,j,t_2})^{X}\prod_{l_1=1}^{q}\Big(\mathbf{r}_j^\star\mathbf{A}_{l_1}\Big(\prod_{l_2=1}^{q_{l_1}}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{B}_{l_1,l_2}\Big)\mathbf{r}_j\Big),$$

$$(\beta_{k,j,t_2}\gamma_{k,j,t_2})^{X}\,(\beta_{j,t_1}\gamma_{j,t_1})^{X_1}\,(\beta_{j,t_2}\gamma_{j,t_2})^{X_2}\prod_{l_1=1}^{q}\Big(\mathbf{r}_j^\star\mathbf{A}_{l_1}\Big(\prod_{l_2=1}^{q_{l_1}}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{B}_{l_1,l_2}\Big)\mathbf{r}_j\Big).$$

Here, $q \in \mathbb{N}$, $q_{l_1} \in \mathbb{N}_0$, $l_1 \in \{1,\dots,q\}$, there exists an index $l_1 \in \{1,\dots,q\}$ such that $q_{l_1} \geq 1$, and the matrices $\mathbf{A}_{l_1}$ and $\mathbf{B}_{l_1,l_2}$ are products of the matrices $\mathbf{D}_{j,t_1}^{-1}, \mathbf{D}_{k,j,t_2}^{-1}$ and $\mathbf{T}_n$ for $l_2 \in \{1,\dots,q_{l_1}\}$, $l_1 \in \{1,\dots,q\}$ and of the deterministic scalars $b_{j,t_1}, b_{j,t_2}, b_{k,j,t_2}$. We assume that $X \in \mathbb{N}$ and that one of the exponents $X_1 \in \mathbb{N}_0$ and $X_2 \in \mathbb{N}_0$ is positive, that is, $X_1 + X_2 \geq 1$. Since, again by Lemma 2.1 in Li (2003),

$$\begin{aligned}\mathbb{E}|Z_n^1(z,t_1,t_2)|^4 &= \mathbb{E}\Big|\sum\sum_{j=1}^{\lfloor nt_1\rfloor}\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j-\mathbb{E}_{j-1})T_{j,k}\Big|^4\\ &\lesssim n\sum\sum_{j=1}^{\lfloor nt_1\rfloor}\mathbb{E}\Big|\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j-\mathbb{E}_{j-1})T_{j,k}\Big|^4,\end{aligned}$$

in order to prove (7.9), it suffices to show that for $j \in \{1,\dots,\lfloor nt_1\rfloor\}$ and $k \in \{\lfloor nt_1\rfloor+1,\dots,\lfloor nt_2\rfloor\}$

$$\mathbb{E}\,|(\mathbb{E}_j-\mathbb{E}_{j-1})T_{j,k}|^4 \lesssim n^{-6}.$$

In order to derive this estimate, we note that we can ignore the deterministic and bounded terms $b_{j,t_1}, b_{j,t_2}, b_{k,j,t_2}$ and denote by $\mathbf{A}_l, l \in \mathbb{N}$, a $p\times p$ (random) matrix which is a product of $\mathbf{D}_{j,t_1}^{-1}, \mathbf{D}_{k,j,t_2}^{-1}$ and $\mathbf{T}_n$. For the sake of brevity, we only consider terms of the type

$$R_1 = \mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j\gamma_{j,t_1}|^4, \tag{7.49}$$
$$R_2 = \mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j|^4, \tag{7.50}$$
$$R_3 = \mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\gamma_{j,t_2}|^4, \tag{7.51}$$
$$R_4 = \mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{k,j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j\gamma_{k,j,t_2}|^4, \tag{7.52}$$
$$R_5 = \mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{k,j,t_2}\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j\gamma_{k,j,t_2}\gamma_{j,t_2}|^4. \tag{7.53}$$

For further investigations, we observe that

$$(\mathbb{E}_j-\mathbb{E}_{j-1})\operatorname{tr}\Big(\prod_{l=1}^{q_1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_l\Big) = 0, \tag{7.54}$$

since $\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_l$, $l \in \{1,\dots,q_1\}$ does not depend on $\mathbf{r}_j$. In order to estimate the term in (7.49), we note that due to independence

$$\mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_2\left(\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3 - n^{-1}\mathbf{T}_n\mathbf{A}_3\right)\mathbf{r}_j\gamma_{j,t_1}|^4 = 0,$$

so that we obtain, using similar arguments as in the derivation of (7.41), in particular the bound in (7.46),

$$\begin{aligned}R_1 \lesssim&\,\mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_2\left(\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3 - n^{-1}\mathbf{T}_n\mathbf{A}_3\right)\mathbf{r}_j\gamma_{j,t_1}|^4\\ &+ n^{-4}\mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_2\mathbf{T}_n\mathbf{A}_3\mathbf{r}_j\gamma_{j,t_1}|^4\\ =&\,n^{-4}\mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_2\mathbf{T}_n\mathbf{A}_3\mathbf{r}_j\gamma_{j,t_1}|^4 \lesssim n^{-6}.\end{aligned}$$

For (7.50), we have using Lemma 7.7.4 and (7.54)

$$
\begin{aligned}
R_2 =& \mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\Big[\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\Big]\Big|^4 \\
\lesssim& \mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\Big[\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j \\
& - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3\Big]\Big|^4 \\
& + \mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\Big[\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \left(n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \left(n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3\Big]\Big|^4 \\
=& \mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\Big[\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j \\
& - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3\Big]\Big|^4 \\
& + n^{-4}\mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\Big[\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{T}_n \left(n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{T}_n \left(n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{A}_3\Big]\Big|^4 \\
\lesssim& n^{-4}\mathbb{E}\left(\operatorname{tr}\left(\mathbf{A}_1 \mathbf{r}_k \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3\right)\left(\mathbf{A}_1 \mathbf{r}_k \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3\right)^\star\right)^2 \\
& + n^{-6} \\
=& n^{-4}\mathbb{E}\Big(\operatorname{tr}\left(\mathbf{A}_1 \mathbf{r}_k \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right) \mathbf{r}_k^\star \mathbf{A}_3\right)\Big(\mathbf{A}_3^\star \mathbf{r}_k \overline{(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2)} \mathbf{r}_k^\star \mathbf{A}_1^\star\Big)\Big)^2 \\
& + n^{-6} \\
=& n^{-4}\mathbb{E}|\left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2\right)^2 \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{A}_3^\star \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_1^\star \mathbf{A}_1 \mathbf{r}_k|^2 + n^{-6} \\
\lesssim& n^{-6}.
\end{aligned}
$$

Next, we have for the term R_4 defined in (7.52) by similar arguments as in the derivation of (7.41)

$$
\begin{aligned}
R_4 =& \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{k,j,t_2}\{\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\}\gamma_{k,j,t_2}|^4 \\
\lesssim& n^{-4}\mathbb{E}\Big[\left(\operatorname{tr}\left(\mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\right)\left(\mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\right)^\star\right)^2 |\beta_{k,j,t_2}\gamma_{k,j,t_2}|^4\Big] \\
\lesssim& n^{-6},
\end{aligned}
$$

where we used the bound in Lemma 7.7.4 and the fact that $\mathbf{r}_j$ is independent of γ_{k,j,t_2} and β_{k,j,t_2}.
Concerning the term R_3 in (7.51), we first decompose using (7.48)

$$
\begin{aligned}
& \gamma_{j,t_2}(z) \\
=& \gamma_{j,k,t_2}(z) - \left(\beta_{k,j,t_2}(z)\mathbf{r}_j^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_j - n^{-1}\mathbb{E}\left[\beta_{k,j,t_2}(z) \operatorname{tr} \mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\right)\right] \\
=& \gamma_{j,k,t_2}(z) - \beta_{k,j,t_2}(z)\left(\mathbf{r}_j^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\right) \\
& - n^{-1}\left(\beta_{k,j,t_2}(z)\mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k - \mathbb{E}\left[\beta_{k,j,t_2}(z) \operatorname{tr} \mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\right]\right) \\
=& \gamma_{j,k,t_2}(z) - \beta_{k,j,t_2}(z)\left(\mathbf{r}_j^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\right) \\
& - n^{-1} b_{k,j,t_2}(z)\left(\mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k - n^{-1}\mathbb{E}\left[\operatorname{tr} \mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\right]\right) \\
& + n^{-1} b_{k,j,t_2}(z)\Big(\beta_{k,j,t_2}(z)\mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \gamma_{k,j,t_2}(z) \\
& - \mathbb{E}\left[\beta_{k,j,t_2}(z) \operatorname{tr} \mathbf{T}_n \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k \mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\gamma_{k,j,t_2}(z)\right]\Big)
\end{aligned}
$$

$$
\begin{aligned}
=&\gamma_{j,k,t_2}(z) - \beta_{k,j,t_2}(z)\left(\mathbf{r}_j^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\right)\\
&- n^{-1}b_{k,j,t_2}(z)\tilde{\gamma}_{k,j,t_2}(z) + n^{-1}b_{k,j,t_2}(z)\Big(\beta_{k,j,t_2}(z)\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\gamma_{k,j,t_2}(z)\\
&- \mathbb{E}\left[\beta_{k,j,t_2}(z)\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\gamma_{k,j,t_2}(z)\right]\Big)\\
=&\gamma_{j,k,t_2}(z) - b_{k,j,t_2}(z)\left(\mathbf{r}_j^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\right)\\
&+ b_{k,j,t_2}(z)\beta_{k,j,t_2}(z)\left(\mathbf{r}_j^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\mathbf{r}_k^\star \mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\right)\gamma_{k,j,t_2}(z)\\
&- n^{-1}b_{k,j,t_2}(z)\tilde{\gamma}_{k,j,t_2}(z) + n^{-1}b_{k,j,t_2}(z)\Big(\beta_{k,j,t_2}(z)\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\gamma_{k,j,t_2}(z)\\
&- \mathbb{E}\left[\beta_{k,j,t_2}(z)\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k\gamma_{k,j,t_2}(z)\right]\Big),
\end{aligned}
$$

where

$$
\tilde{\gamma}_{k,j,t_2}(z) = \mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{r}_k - n^{-1}\mathbb{E}\left[\operatorname{tr}\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}(z)\right].
$$

Thus, we conclude for (7.51), using the notations $\mathbf{A}_3 = \mathbf{D}_{k,j,t_2}^{-1}$ and $\mathbf{A}_4 = \mathbf{D}_{k,j,t_2}^{-1}\mathbf{T}_n\mathbf{D}_{k,j,t_2}^{-1}$ and the fact that b_{k,j,t_2} is deterministic and bounded,

$$
\begin{aligned}
R_3 \lesssim\ & \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\gamma_{j,k,t_2}|^4\\
&+ \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}b_{k,j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\left(\mathbf{r}_j^\star\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)|^4\\
&+ \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\beta_{k,j,t_2}b_{k,j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\left(\mathbf{r}_j^\star\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)\gamma_{k,j,t_2}|^4\\
&+ n^{-4}\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}b_{k,j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\tilde{\gamma}_{k,j,t_2}|^4\\
&+ n^{-4}\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\beta_{k,j,t_2}b_{k,j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\mathbf{r}_k^\star\mathbf{A}_4\mathbf{r}_k\gamma_{k,j,t_2}|^4\\
&+ n^{-4}\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j|^4\mathbb{E}|\beta_{k,j,t_2}b_{k,j,t_2}\mathbf{r}_k^\star\mathbf{A}_4\mathbf{r}_k\gamma_{k,j,t_2}|^4\\
\lesssim\ & R_{31} + R_{32} + R_{33} + n^{-6},
\end{aligned}
$$

where

$$
\begin{aligned}
R_{31} =&\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\gamma_{j,k,t_2}|^4, \qquad (7.55)\\
R_{32} =&\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\left(\mathbf{r}_j^\star\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)|^4,\\
R_{33} =&\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\beta_{k,j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\left(\mathbf{r}_j^\star\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)\gamma_{k,j,t_2}|^4, \qquad (7.56)
\end{aligned}
$$

and we used an analogue of the estimate (7.46) for the terms γ_{k,j,t_2} and $\tilde{\gamma}_{k,j,t_2}$ in the last step. The term R_{31} in (7.55) can be bounded using the bounds in (7.42), (7.43) and (7.45) as follows:

$$
R_{31} \lesssim R_{311} + R_{312} + o\left(n^{-l}\right),
$$

where

$$
\begin{aligned}
R_{311} =&\mathbb{E}|\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\gamma_{j,k,t_2}|^4,\\
R_{312} =&\mathbb{E}|\mathbf{r}_j^\star\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\gamma_{j,k,t_2}|^4.
\end{aligned}
$$

Since R_{312} can be handled similarly to R_{311}, we only consider R_{311} and obtain by Lemma 7.7.4

$$\begin{aligned}R_{311} \lesssim & \mathbb{E}\left| \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1\right) \gamma_{j,k,t_2}\right|^4 + n^{-4}\mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \gamma_{j,k,t_2}|^4 \\ \lesssim & n^{-6} + n^{-4}\mathbb{E}\left[|\gamma_{j,k,t_2}|^4 \left(\operatorname{tr}\left(\mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1\right)\left(\mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1\right)^\star\right)^2\right] \\ \leq & n^{-6}.\end{aligned}$$

Note that the term R_{33} defined in (7.56) can be bounded similarly. Similarly as R_{31} given in (7.55), we bound $|\beta_{j,t_2}|$ and get

$$R_{32} \lesssim R_{321} + R_{322} + o\left(n^{-l}\right),$$

where

$$\begin{aligned}R_{321} =& \mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\right)|^4, \\ R_{322} =& \mathbb{E}|\mathbf{r}_j^\star \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\right)|^4.\end{aligned}$$

For the sake of brevity, we shall limit ourselves to investigating the summand R_{321}.

$$\begin{aligned}R_{321} \lesssim & \mathbb{E}\left| \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1\right)\left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\right)\right|^4 \\ & + n^{-4}\mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\right)|^4 \\ \lesssim & n^{-6} + \left(\mathbb{E}|\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1|^8 \mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3|^8\right)^{\frac{1}{2}} \\ \lesssim & n^{-6}.\end{aligned}$$

Finally, invoking Lemma 7.7.4 and (7.46), we can show for the term R_5 defined in (7.53) that

$$R_5 \lesssim R_{51} + R_{52},$$

where

$$\begin{aligned}R_{51} =& \mathbb{E}|\beta_{k,j,t_2}\beta_{j,t_2}\left(\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3\right)\gamma_{k,j,t_2}\gamma_{j,t_2}|^4 \\ \leq & \left(\mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3|^8 \mathbb{E}|\beta_{k,j,t_2}\beta_{j,t_2}\gamma_{k,j,t_2}\gamma_{j,t_2}|^8\right)^{\frac{1}{2}} \\ \lesssim & n^{-6}, \\ R_{52} =& n^{-4}\mathbb{E}|\beta_{k,j,t_2}\beta_{j,t_2}\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \gamma_{k,j,t_2}\gamma_{j,t_2}|^4 \lesssim n^{-6}.\end{aligned}$$

Thus, the moment inequalities (7.8), (7.9) and (7.10) for M_n^1 hold true. □

Lemma 7.7.6 *For any bounded subset $S \subset \mathbb{C}^+$, we have*

$$\inf_{z \in S, t \in [t_0,1]} |\underline{\tilde{s}}_t(z)| > 0.$$

Proof of Lemma 7.7.6. Let us assume that the assertion does not hold. In this case, there exists sequences $(z_n)_{n\in\mathbb{N}}$ in S and $(t_n)_{n\in\mathbb{N}}$ in $[t_0,1]$ with the property

$$\lim_{n\to\infty} \underline{\tilde{s}}_{t_n}(z_n) = 0.$$

By choosing appropriate subsequences, we assume without loss of generality that $(z_n)_{n\in\mathbb{N}}$ converges to a limit in the closure of S and $(t_n)_{n\in\mathbb{N}}$ converges to a limit in $[t_0,1]$. From (3.2), we conclude

$$\lim_{n\to\infty} y\int \frac{\lambda \underline{\tilde{s}}_{t_n}(z_n)}{1+\lambda t_n \underline{\tilde{s}}_{t_n}(z_n)} dH(\lambda) = 1.$$

But, using the fact that H is compactly supported, we see that the expression above tends to 0. Thus, we get a contradiction. □

Lemma 7.7.7 *In the real case, it holds for $i\neq l$ $(i,l\in\{1,\ldots,n\}\setminus\{j\})$*

$$\begin{aligned}\sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} \Big|\mathbb{E}\Big[&\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\beta_{l,i,j,t}(z)\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\\ &\times\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\beta_{i,l,j,t}(z)\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\Big]\Big| \qquad (7.57)\\ =&o\left(n^{-1}\right).\end{aligned}$$

Proof of Lemma 7.7.7. Denoting

$$\begin{aligned}\hat{\gamma}_{i,l,j,t}(z) =& \mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z),\\ \overline{\beta}^2_{i,j,l,t}(z) =& \frac{1}{1+n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)},\end{aligned}$$

we use the representation (7.22) in order to replace $\beta_{l,i,j,t}(z)$ and $\beta_{i,l,j,t}(z)$. Note that $\mathbb{E}[||\mathbf{D}^{-1}_{l,i,j,t}(z)||]\le K$ and $||\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}||\le K$ which follows from Lemma 7.7.3 and Lemma 7.7.2 in Section 7.7. By applying the triangle inequality, this gives us several summands for the mean in (7.57). More precisely, we can write

$$\begin{aligned}&\Big|\mathbb{E}\Big[\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\beta_{l,i,j,t}(z)\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\\ &\times\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\beta_{i,l,j,t}(z)\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\Big]\Big|\\ \le&\sum_{\zeta_1,\zeta_2}|\mathbb{E}[T(\zeta_1,\zeta_2)]|,\end{aligned}$$

where $T(\zeta_1,\zeta_2)$ has the following form

$$\begin{aligned}&\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\zeta_1\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\\ &\times\zeta_2\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right),\end{aligned}$$

and

$$\zeta_1\in\{\overline{\beta}_{l,i,j,t}(z), -\overline{\beta}^2_{l,i,j,t}(z)\hat{\gamma}_{l,i,j,t}(z), \beta_{l,i,j,t}(z)\overline{\beta}^2_{l,i,j,t}(z)\hat{\gamma}^2_{l,i,j,t}(z)\},$$

$$\zeta_2 \in \{\overline{\beta}_{i,l,j,t}(z), -\overline{\beta}_{i,l,j,t}(z)^2\hat{\gamma}_{i,l,j,t}(z), \beta_{i,l,j,t}(z)\overline{\beta}^2_{i,l,j,t}(z)\hat{\gamma}^2_{i,l,j,t}(z)\}.$$

The assertion now follows, if we show that for all ζ_1, ζ_2

$$|\mathbb{E}[T(\zeta_1,\zeta_2)]| = o\left(n^{-1}\right). \tag{7.58}$$

In the following, we restrict ourselves to three different cases noting that the remaining cases can be handled similarly.
To begin with, let $\zeta_1 = \overline{\beta}_{l,i,j,t}(z)$ and $\zeta_2 = \overline{\beta}_{i,l,j,t}(z)$. In this case, we have

$$\begin{aligned}
&|\mathbb{E}[T(\zeta_1,\zeta_2)]| \\
\le& K\Big|\mathbb{E}\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n \\
&\times \mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\Big| \\
\le& K\Big|\mathbb{E}\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1} \\
&\times \mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\Big| \\
&+Kn^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1} \\
&\times \mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\Big| \\
&+Kn^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1} \\
&\times \mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{T}_n\Big| \\
&+Kn^{-2}\Big|\mathbb{E}\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1} \\
&\times \mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{T}_n\Big| \\
=& K(T_1+T_2+T_3)+o\left(n^{-1}\right),
\end{aligned}$$

where

$$\begin{aligned}
T_1 =& \Big|\mathbb{E}\Big[\mathbf{r}_i^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z) \\
&\times \mathbf{r}_l\mathbf{r}_l^\star\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\Big]\Big|, \\
T_2 =& n^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1} \\
&\times \mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\Big|, \\
T_3 =& n^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1} \\
&\times \mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{T}_n\Big|.
\end{aligned}$$

For the first summand, we obtain using (9.8.6) in Bai and Silverstein (2010) for the real case

$$T_1 \le \Big|\mathbb{E}\Big[\big\{\mathbf{r}_i^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i$$

$$
\begin{aligned}
&- n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}^{-1}_{l,i,j,t}(z) \mathbf{r}_l \mathbf{r}_l^\star \mathbf{D}^{-1}_{l,i,j,t}(z) \left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-1} \mathbf{T}_n \mathbf{D}^{-1}_{i,l,j,t}(z) \big\} \\
&\times \left\{ \mathbf{r}_i^\star \mathbf{D}^{-1}_{i,l,j,t}(z) \mathbf{r}_l \mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-2} \mathbf{T}_n \mathbf{r}_i - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}^{-1}_{i,l,j,t}(z) \mathbf{r}_l \mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-2} \mathbf{T}_n \right\} \Big] \Big| \\
&+ o\left(n^{-1}\right) \\
=& o\left(n^{-1}\right).
\end{aligned}
$$

With similar ideas, it can be shown that $T_2 = o\left(n^{-1}\right)$ and $T_3 = o\left(n^{-1}\right)$ and that (7.58) holds true in the case $\zeta_1 = \overline{\beta}_{l,i,j,t}(z)$ and $\zeta_2 = -\overline{\beta}^2_{i,l,j,t}(z)\hat{\gamma}_{i,l,j,t}(z)$.

Finally, we consider the case $\zeta_1 = -\overline{\beta}^2_{l,i,j,t}(z)\hat{\gamma}_{l,i,j,t}(z)$ and $\zeta_2 = -\overline{\beta}^2_{i,l,j,t}(z)\hat{\gamma}_{i,l,j,t}(z)$. Note that $\overline{\beta}_{i,l,j,t}(z) = \overline{\beta}_{l,i,j,t}(z)$. We obtain (7.58), that is,

$$
\begin{aligned}
&|\mathbb{E}[T(\zeta_1,\zeta_2)]| \\
=& \Big| \mathbb{E}\Big[\overline{\beta}^4_{i,j,l,t}(z) \operatorname{tr} \left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-2} \mathbf{T}_n \left(\mathbf{r}_i \mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right) \mathbf{D}^{-1}_{l,i,j,t}(z) \mathbf{r}_l \mathbf{r}_l^\star \mathbf{D}^{-1}_{l,i,j,t}(z) \\
&\times \left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-1} \mathbf{T}_n \hat{\gamma}_{l,i,j,t}(z) \mathbf{D}^{-1}_{i,l,j,t}(z) \mathbf{r}_i \mathbf{r}_i^\star \mathbf{D}^{-1}_{i,l,j,t}(z) \left(\mathbf{r}_l \mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right) \hat{\gamma}_{i,l,j,t}(z) \Big] \Big| \\
\leq& \mathbb{E}^{\frac{1}{2}}|E_1|^2 \mathbb{E}^{\frac{1}{4}}|E_2|^4 \mathbb{E}^{\frac{1}{4}}|E_3|^4 = o\left(n^{-1}\right),
\end{aligned}
$$

if

$$\mathbb{E}^{\frac{1}{2}}|E_1|^2 \leq Kn^{-1}, \tag{7.59}$$

$$\mathbb{E}^{\frac{1}{4}}|E_2|^4 \leq K, \tag{7.60}$$

$$\mathbb{E}^{\frac{1}{4}}|E_3|^4 = o(1), \tag{7.61}$$

where

$$
\begin{aligned}
E_1 =& \overline{\beta}^4_{i,j,l,t}(z) \hat{\gamma}_{i,l,j,t}(z) \hat{\gamma}_{l,i,j,t}(z), \\
E_2 =& \mathbf{r}_l^\star \mathbf{D}^{-1}_{l,i,j,t}(z) \left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-1} \mathbf{T}_n \mathbf{D}^{-1}_{i,l,j,t}(z) \mathbf{r}_i, \\
E_3 =& \operatorname{tr} \Big\{ \left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-2} \mathbf{T}_n \left(\mathbf{r}_i \mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right) \mathbf{D}^{-1}_{l,i,j,t}(z) \mathbf{r}_l \mathbf{r}_i^\star \mathbf{D}^{-1}_{i,l,j,t}(z) \left(\mathbf{r}_l \mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right) \Big\}.
\end{aligned}
$$

We begin with a proof of (7.59). Note that $\hat{\gamma}_{l,i,j,t}(z)$ is independent of $\mathbf{r}_i$ and $\overline{\beta}_{i,j,l,t}(z)$ is independent of $\mathbf{r}_i$ and $\mathbf{r}_j$. Using (9.9.6) in Bai and Silverstein (2010) twice, we obtain

$$\mathbb{E}|E_1|^2 \leq Kn^{-2},$$

which proves (7.59). The estimate (7.60) can be proven similarly to Bai and Silverstein (2010), p. 290.

Finally, we will prove that (7.61) holds true. We obtain

$$
\begin{aligned}
E_3 =& \mathbf{r}_i^\star \mathbf{D}^{-1}_{i,l,j,t}(z) \left(\mathbf{r}_l \mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right) \left(t\underline{\tilde{s}}_t(z)\mathbf{T} + \mathbf{I}\right)^{-2} \mathbf{T} \left(\mathbf{r}_i \mathbf{r}_i^\star - n^{-1}\mathbf{T}\right) \mathbf{D}^{-1}_{l,i,j,t}(z) \mathbf{r}_l \\
=& E_{31} + E_{32} + E_{33} + E_{34},
\end{aligned}
$$

where

$$E_{31} = \mathbf{r}_i^\star \mathbf{D}^{-1}_{i,l,j,t}(z) \mathbf{r}_l \mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T} + \mathbf{I}\right)^{-2} \mathbf{T} \mathbf{r}_i \mathbf{r}_i^\star \mathbf{D}^{-1}_{l,i,j,t}(z) \mathbf{r}_l$$

$$E_{32} = - n^{-1}\mathbf{r}_i^\star \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}^2\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l,$$
$$E_{33} = - n^{-1}\mathbf{r}_i^\star \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{T}_n \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l,$$
$$E_{34} = n^{-2}\mathbf{r}_i^\star \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{T}_n \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}^2\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l.$$

For $k \in \{2,3,4\}$, it holds

$$E|E_{3k}|^4 = o(1).$$

For the first summand, we conclude

$$\begin{aligned}
&E|E_{31}|^4\\
\leq& K\mathbb{E}^{\frac{1}{2}}\left|\mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\right|^8\\
\leq& K\mathbb{E}^{\frac{1}{2}}\left|\mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l - n^{-1}\operatorname{tr}\mathbf{T}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\right|^8\\
&+K\mathbb{E}^{\frac{1}{2}}\left|n^{-1}\operatorname{tr}\mathbf{T}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\right|^8\\
\leq& K\mathbb{E}^{\frac{1}{2}}\left|\mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l - n^{-1}\operatorname{tr}\mathbf{T}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\right|^8\\
&+Kn^{-4}\\
\leq& Kn^{-\frac{1}{2}}+Kn^{-4}=o(1),
\end{aligned}$$

which proves (7.61). Hence, the proof of Lemma 7.7.7 is finished. □

Lemma 7.7.8 *It holds for sufficiently large* $N \in \mathbb{N}$

$$\inf_{n\geq N}\inf_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}}\left(\operatorname{Im}(z)+\operatorname{Im}(R_{n,t}(z))\right)\geq 0.$$

Proof of Lemma 7.7.8. We start by investigating real and imaginary part of $1/\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]$. As a preparation for the latter, one can show similarly to Lemma 7.7.6 that $\operatorname{Re}(\underline{\tilde{s}}_t(z))$ is uniformly bounded away from 0. Thus, due to Theorem 6.3.7, we also have for some sufficiently large $N \in \mathbb{N}$

$$\inf_{n\geq N}\inf_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}}|\operatorname{Re}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|>0 \text{ and } \inf_{n\geq N}\inf_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}}|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|>0.$$

Using also $|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]| \leq 1/\operatorname{Im}(z)$, this implies for the real part of the inverse for some $K_1 > 0$

$$\operatorname{Re}\left(\frac{1}{\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}\right)=\frac{\operatorname{Re}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)}{|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}\geq K_1\operatorname{Im}^2(z).$$

For the imaginary part, we conclude for some $K_2 > 0$

$$\operatorname{Im}\left(\frac{1}{\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]}\right)=\frac{-\operatorname{Im}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)}{|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}=\frac{1}{|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2}\mathbb{E}\operatorname{Im}\left(\int\frac{-1}{\lambda-z}dF^{\underline{\mathbf{B}}_{n,t}}(\lambda)\right)$$

$$=\frac{1}{|\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]|^2}\mathbb{E}\int \frac{-\operatorname{Im}(z)}{|z-\lambda|^2}dF^{\underline{\mathbf{B}}_{n,t}}(\lambda) \geq K\mathbb{E}\int \frac{-\operatorname{Im}(z)}{|\lambda - z|^2}dF^{\underline{\mathbf{B}}_{n,t}}(\lambda)$$
$$\geq - K_2 \operatorname{Im}(z).$$

By definition of $R_{n,t}(z)$, we have for all $n \geq N$

$$\operatorname{Im}(R_{n,t}(z)) = y_{\lfloor nt \rfloor} \lfloor nt \rfloor^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \operatorname{Im}\left(\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\left(\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]\right)^{-1}\right)$$
$$= y_{\lfloor nt \rfloor} \lfloor nt \rfloor^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \operatorname{Im}\left(\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) \operatorname{Re}\left(\frac{1}{\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]}\right)$$
$$+ y_{\lfloor nt \rfloor} \lfloor nt \rfloor^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \operatorname{Re}\left(\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) \operatorname{Im}\left(\frac{1}{\mathbb{E}[\tilde{\underline{s}}_{n,t}(z)]}\right)$$
$$\geq K_1 \operatorname{Im}^2(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Im}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\}$$
$$- K_2 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Re}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Re}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\},$$

which implies that

$$\operatorname{Im}(R_{n,t}(z)) + \operatorname{Im}(z)$$
$$\geq \operatorname{Im}(z) + K_1 \operatorname{Im}^2(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Im}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\}$$
$$- K_2 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Re}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Re}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\}$$
$$\geq \operatorname{Im}(z)\Bigg[1 + K_1 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Im}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\}$$
$$- K_2 \lfloor nt \rfloor^{-1} \operatorname{Re}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Re}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\}\Bigg].$$

Due to (6.54), we have for some $N \in \mathbb{N}$

$$\sup_{n \geq N} \sup_{\substack{z \in \mathcal{C}_n, \\ t \in [t_0, 1]}} \Bigg| K_1 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Im}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\}$$
$$- K_2 \lfloor nt \rfloor^{-1} \operatorname{Re}\left(y_{\lfloor nt \rfloor}\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) I\left\{\operatorname{Re}\left(\sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)]\right) > 0\right\}\Bigg| < 1.$$

Thus, we conclude that

$$\inf_{n \geq N} \inf_{\substack{z \in \mathcal{C}_n, \\ t \in [t_0, 1]}} (\mathrm{Im}(z) + \mathrm{Im}(R_{n,t}(z))) \geq 0.$$

□

Chapter 8

Proof of Theorem 5.2.1

8.1 Proof of Theorem 5.2.1

Due to the invariance of $U_{n,t}$ under H_0, we may assume w.l.o.g. that that $\mathbf{\Sigma}_1 = \ldots \mathbf{\Sigma}_n = \mathbf{I}$, which implies $\hat{\mathbf{\Sigma}}_{n,t} = \mathbf{B}_{n,t}$. Let $f_1(x) = x$ and $f_2(x) = x^2$. We recall from Corollary 4.2.1 that

$$\big((X_n(f_1,t))_{t\in[t_0,1]}, (X_n(f_2,t))_{t\in[t_0,1]}\big)_{n\in\mathbb{N}} \rightsquigarrow \big((X(f_1,t))_{t\in[t_0,1]}, (X(f_2,t))_{t\in[t_0,1]}\big)$$

in $(\ell^\infty([t_0,1]))^2$, where

$$X_n(f_1,t) = \operatorname{tr}(\mathbf{B}_{n,t}) - \lfloor nt \rfloor y_n,$$
$$X_n(f_2,t) = \operatorname{tr}\left(\mathbf{B}_{n,t}^2\right) - \lfloor nt \rfloor y_n \Big(\frac{\lfloor nt \rfloor}{n} + y_n\Big),\ t \in [t_0,1],$$

and

$$\mathbb{E}[X(f_1,t)] = 0,\ \mathbb{E}[X(f_2,t)] = ty, \tag{8.1}$$

and

$$\begin{aligned} &\operatorname{cov}(X(f_1,t_1), X(f_1,t_2)) = 2y\min(t_1,t_2),\\ &\operatorname{cov}(X(f_2,t_1), X(f_2,t_2)) = 4\min(t_1,t_2)y\left\{2t_1t_2 + [\min(t_1,t_2) + 2(t_1+t_2)]\,y + 2y^2\right\}, \\ &\operatorname{cov}(X(f_1,t_1), X(f_2,t_2)) = 4\min(t_1,t_2)y(t_2+y),\ t_1,t_2 \in [t_0,1]. \end{aligned} \tag{8.2}$$

With the definition $\phi(x,y) = \frac{y}{x^2}$, we obtain the representation

$$U_{n,t} = \phi\Big(\frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}), \frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}^2)\Big), t \in [t_0,1],\ n \in \mathbb{N},$$

for the process $U_{n,t}$ in (5.3). Consequently, the assertion can be proved by the functional delta method.

To be precise, note that it follows from $y_n = p/n$

$$p\begin{pmatrix} \frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}) - \frac{\lfloor nt \rfloor}{n} \\ \frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}^2) - \frac{\lfloor nt \rfloor}{n}\left(\frac{\lfloor nt \rfloor}{n} + y_n\right) \end{pmatrix}_{t\in[t_0,1]} \rightsquigarrow \begin{pmatrix} X(f_1,t) \\ X(f_2,t) \end{pmatrix}_{t\in[t_0,1]} \tag{8.3}$$

in $(\ell^\infty([t_0,1]))^2$. Let $a_n(t) = \frac{\lfloor nt \rfloor}{n}$ and $b_n(t) = \frac{\lfloor nt \rfloor}{n}\left(\frac{\lfloor nt \rfloor}{n} + y_n\right)$, such that

$$\lim_{n\to\infty} a_n(t) = t = a(t), \;\; \lim_{n\to\infty} b_n(t) = t(t+y) = b(t)$$

uniformly in $t \in [t_0, 1]$. For a sequence $(h_{n,1}, h_{n,2})_{n\in\mathbb{N}}$ in $(\ell^\infty([t_0,1]))^2$ converging to 0, a straightforward calculation shows that

$$p\left\{\phi\left(a_n + p^{-1}h_{n,1}, b_n + p^{-1}h_{n,2}\right) - \phi\left(a_n, b_n\right)\right\} \to \frac{h_2}{a^2} - \frac{2bh_1}{a^3} = \phi'_{(a,b)}(h_1, h_2)$$

in $l^\infty([t_0,1])$, as $n \to \infty$. Moreover, we have

$$\phi\left(a_n(t), b_n(t)\right) = \frac{\frac{\lfloor nt \rfloor}{n}\left(\frac{\lfloor nt \rfloor}{n} + y_n\right)}{\left(\frac{\lfloor nt \rfloor}{n}\right)^2} = \frac{n}{\lfloor nt \rfloor}\left(\frac{\lfloor nt \rfloor}{n} + y_n\right) = 1 + y_{\lfloor nt \rfloor}.$$

Thus, it follows from (8.3) and Theorem 3.9.5 in Van Der Vaart and Wellner (1996) that

$$\begin{aligned} p\left\{U_{n,t} - 1 - y_{\lfloor nt \rfloor}\right\}_{t\in[t_0,1]} &= p\Big\{\phi\Big(\frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}), \frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}^2)\Big) - \phi\left(a_n(t), b_n(t)\right)\Big\}_{t\in[t_0,1]} \\ &\rightsquigarrow (U_t)_{t\in[t_0,1]} \end{aligned}$$

in $\ell^\infty([t_0,1])$, where

$$U_t = \frac{X(f_2,t) - 2X(f_1,t)(t+y)}{t^2}, \;\; t \in [t_0, 1]$$

is a Gaussian process. Recalling (8.1) and (8.2) we obtain for $t, t_1, t_2 \in [t_0, 1]$ with $t_2 \le t_1$ by straightforward calculations

$$\begin{aligned} \mathbb{E}[U_t] =& \frac{1}{t^2}\left(\mathbb{E}[X(f_2,t)] - 2(t+y)\mathbb{E}[X(f_1,t)]\right) = \frac{ty}{t^2} = y_t, \\ \operatorname{cov}(U_{t_1}, U_{t_2}) =& \frac{1}{t_1^2 t_2^2}\operatorname{cov}\left(X(f_2,t_1) - 2(t_1+y)X(f_1,t_1), X(f_2,t_2) - 2(t_2+y)X(f_1,t_2)\right) \\ =& \frac{1}{t_1^2 t_2^2}\Big\{4t_2 y\left\{2t_1t_2 + \left[t_2 + 2(t_1+t_2)\right]y + 2y^2\right\} - 2(t_2+y)4\min(t_1,t_2)y(t_1+y) \\ & - 2(t_1+y)4\min(t_1,t_2)y(t_2+y) + 4(t_1+y)(t_2+y)2y\min(t_1,t_2)\Big\} \\ =& \frac{1}{t_1^2 t_2^2}\Big\{4t_2 y\left\{2t_1t_2 + \left[t_2 + 2(t_1+t_2)\right]y + 2y^2\right\} - 8(t_2+y)\min(t_1,t_2)y(t_1+y)\Big\} \\ =& 4\frac{y^2}{t_1^2} = 4y^2_{\max(t_1,t_2)}. \end{aligned}$$

which proves the assertion of Theorem 5.2.1.

8.2 Auxiliary results for the proof of Theorem 5.2.1

8.2.1 How to calculate mean and covariance in Theorem 3.2.1

Proof of Proposition 4.1.1. If suffices to consider the real case. Since $H = \delta_1$, we obtain from Theorem 3.2.1

$$\mathbb{E}[X(f_i,t)] = -\frac{1}{2\pi i}\oint_{\mathcal{C}} f_i(z)\frac{ty\frac{\tilde{\underline{s}}_t^3(z)}{(t\tilde{\underline{s}}_t(z)+1)^3}}{\left(1-ty\frac{\tilde{\underline{s}}_t^2(z)}{(t\tilde{\underline{s}}_t(z)+1)^2}\right)^2}dz,\ i=1,2,$$

$$\operatorname{cov}(X(f_1,t_1),X(f_2,t_2)) = \frac{1}{2\pi^2}\oint_{\mathcal{C}_1}\oint_{\mathcal{C}_2} f_1(z_1)\overline{f_2(z_2)}\sigma^2_{t_1,t_2}(z_1,\overline{z_2})\overline{dz_2}dz_1,$$

where the contours $\mathcal{C},\mathcal{C}_1,\mathcal{C}_2$ enclose the interval given in (3.7) and $\mathcal{C}_1$ and $\mathcal{C}_2$ are assumed to be non-overlapping.

Step 1: *Specifying the contours*

We claim that it suffices for $\mathcal{C}=\mathcal{C}_t$ to enclose the interval $\left[t(1-\sqrt{y_t})^2, t(1+\sqrt{y_t})^2\right]$ and we will prove this assertion in a first step. Similar arguments hold true for contours $\mathcal{C}_1=\mathcal{C}_{1,t_1}$ and $\mathcal{C}_2=\mathcal{C}_{2,t_2}$.

The assertion is clear in the case $y_t<1$. In the case $y_t>1$, the transformed Marčenko-Pastur distribution $\tilde{F}^{y_{\lfloor nt\rfloor}}$ has a discrete part at the origin for sufficiently large n. A priori, the contour should enclose the whole support of $\tilde{F}^{y_t}$, including the origin. However, by the exact separation theorem in Bai and Silverstein (1999), we see that the mass at 0 of the spectral distribution $F^{\mathbf{B}_{n,t}}$ coincides with that of $\tilde{F}^{y_{\lfloor nt\rfloor}}$ for sufficiently large n. Thus, we can restrict the integration in (3.5) to the interval $\left[t(1-\sqrt{y_t})^2, t(1+\sqrt{y_t})^2\right]$ and neglect the discrete part at the origin.

Step 2: *Calculation of the mean*

Recall that $h_t=\sqrt{y_t}$. For calculation of the mean, we use a change of variables

$$z(\xi)=z=t(1+h_tr\xi+h_tr^{-1}\xi^{-1}+h_t^2),$$

where $r>1$ is close to 1 and $|\xi|=1$. It can be checked that when ξ runs anticlockwise on the unit circle, z will run a contour $\mathcal{C}$ enclosing the interval $[t(1-h_t)^2, t(1+h_t)^2]$. Using the identity (3.2), we have for $z\in\mathcal{C}$

$$\tilde{\underline{s}}_t(z) = -\frac{1}{t(1+h_tr\xi)}\ ,\ \frac{\tilde{\underline{s}}_t(z)}{t\tilde{\underline{s}}_t(z)+1} = -\frac{1}{th_tr\xi}\ ,\ dz = th_t(r-r^{-1}\xi^{-2})d\xi.$$

Thus, we can write for $i\in\{1,2\}$

$$\begin{aligned}\mathbb{E}[X(f_i,t)] &= \lim_{r\searrow 1}\frac{1}{2\pi i}\oint_{|\xi|=1} f_i(z(\xi))t^2h_t^2\frac{\left(\frac{1}{th_tr\xi}\right)^3}{\left(1-t^2h_t^2\left(\frac{1}{th_tr\xi}\right)^2\right)^2}th_t\left(r-r^{-1}\xi^{-2}\right)d\xi\\ &= \lim_{r\searrow 1}\frac{t}{2\pi i}\oint_{|\xi|=1} f_i(z(\xi))\frac{r^{-2}}{\xi t(\xi^2-r^{-2})}d\xi\end{aligned}$$

$$= -\lim_{r \searrow 1} \frac{t}{2\pi i} \oint_{|\xi|=1} f_i\left(z(\xi)\right)\left(\frac{1}{\xi t} - \frac{\xi}{t(\xi^2 - r^{-2})}\right) d\xi$$
$$= \lim_{r \searrow 1} \frac{1}{2\pi i} \oint_{|\xi|=1} f_i\left(z(\xi)\right)\left(\frac{\xi}{(\xi^2 - r^{-2})} - \frac{1}{\xi}\right) d\xi.$$

Step 3: *Calculation of the covariance function*
In order to calculate the covariance structure, we define two non-overlapping contours through

$$z_j = z_j(\xi_j) = t\left(1 + h_{t_j}\xi_j + h_{t_j} r_j^{-1}\overline{\xi}_j + h_{t_j}^2\right),\ j = 1, 2,$$

where $r_2 > r_1 > 1$. Thus, we have for $t_2 \leq t_1$

$$\operatorname{cov}(X(f_1,t_1), X(f_2,t_2)) = \lim_{\substack{r_2 > r_1,\\ r_1, r_2 \searrow 1}} \frac{1}{2\pi^2} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} f_1(z_1(\xi_1))\overline{f_2(z_2(\xi_2))}$$
$$\times \sigma^2_{t_1,t_2}(z_1(\xi_1), \overline{z_2(\xi_2)}) t_1 h_{t_1}(r_1 - r_1^{-1}\xi_1^{-2}) t_2 h_{t_2}(r_2 - r_2^{-1}\xi_2^{-2})\overline{d\xi_2} d\xi_1$$
$$= \lim_{\substack{r_2 > r_1,\\ r_1, r_2 \searrow 1}} \frac{1}{2\pi^2} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} f_1(z_1(\xi_1))\overline{f_2(z_2(\xi_2^{-1}))}$$
$$\times \sigma^2_{t_1,t_2}(z_1(\xi_1), \overline{z_2(\xi_2^{-1})}) t_1 h_{t_1}(r_1 - r_1^{-1}\xi_1^{-2}) t_2 h_{t_2}(r_2 - r_2^{-1}\xi_2^{2}) d\xi_2 d\xi_1.$$

Proceeding similarly as for the mean, we get by straightforward but tedious algebra the desired formula for the covariance (we partially used a computer algebra system). □

8.2.2 Proof of Corollary 4.2.1

We apply Theorem 3.2.1 for the special case $f_1(x) = x,\ f_2(x) = x^2, \mathbf{T}_n = \mathbf{I}$, that is

$$X_n(f_1,t) = \operatorname{tr}\left(\mathbf{B}_{n,t}\right) - \lfloor nt \rfloor y_n,$$
$$X_n(f_2,t) = \operatorname{tr}\left(\mathbf{B}_{n,t}^2\right) - \lfloor nt \rfloor y_n\left(\frac{\lfloor nt \rfloor}{n} + y_n\right),\ t \in [t_0, 1].$$

Note that all conditions from Theorem 3.2.1 are satisfied, and therefore

$$\left((X_n(f_1,t))_{t\in[t_0,1]}, (X_n(f_2,t))_{t\in[t_0,1]}\right)_{n\in\mathbb{N}} \rightsquigarrow \left((X(f_1,t))_{t\in[t_0,1]}, (X(f_2,t))_{t\in[t_0,1]}\right) \quad (8.4)$$

in the space $(\ell^\infty([t_0,1]))^2$, where $\left((X(f_1,t))_{t\in[t_0,1]}, (X(f_2,t))_{t\in[t_0,1]}\right)$ is a Gaussian process. Thus, it is left to calculate mean, covariance and the centering term appearing in Theorem 3.2.1. We begin determining the centering term. Using the moments of the Marčenko-Pastur distribution (e.g., see Example 2.12 in Yao et al. (2015)), we get

$$\int f_1(x) d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = \int x d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = \frac{\lfloor nt \rfloor}{n} \int x dF^{y_{\lfloor nt \rfloor}}(x) = \frac{\lfloor nt \rfloor}{n},$$

where F^y denotes the Marčenko-Pastur distribution with index parameter $y > 0$ and scale parameter $\sigma^2 = 1$. Similarly, we see that by using Proposition 2.13 in Yao et al. (2015)

$$\int f_2(x) d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = \int x^2 d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = \Big(\frac{\lfloor nt \rfloor}{n}\Big)^2 \big(1 + y_{\lfloor nt \rfloor}\big) = \frac{\lfloor nt \rfloor}{n}\Big(\frac{\lfloor nt \rfloor}{n} + y_n\Big).$$

We calculate the quantities given in Proposition 4.1.1 by using the residue theorem. We find for the real case

$$\begin{aligned}
\mathbb{E}[X(f_1,t)] =& \frac{t}{2\pi i} \lim_{r \searrow 1} \oint\limits_{|\xi|=1} \frac{\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi}{\xi} \Big(\frac{\xi}{\xi^2 - r^{-2}} - \frac{1}{\xi} \Big) d\xi \\
=& \frac{t}{2\pi i} \lim_{r \searrow 1} \oint\limits_{|\xi|=1} \frac{\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi}{(\xi - r^{-1})(\xi + r^{-1})} d\xi - \frac{t}{2\pi i} \oint\limits_{|\xi|=1} \frac{\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi}{\xi^2} d\xi \qquad (8.5)\\
=& \lim_{r \searrow 1} t \frac{\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi}{\xi + r^{-1}} \Big|_{\xi = r^{-1}} + \lim_{r \searrow 1} t \frac{\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi}{\xi - r^{-1}} \Big|_{\xi = -r^{-1}} \\
& - t \frac{\partial}{\partial \xi} \left(\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi \right) \Big|_{\xi = 0} \\
=& t \lim_{r \searrow 1} \frac{2 r^{-1} + 2 h_t^2 r^{-1}}{2 r^{-1}} - t(1 + h_t^2) = 0.
\end{aligned}$$

Note that $\xi = \pm r^{-1}$ are poles of order 1 for the first integrand in (8.5), since $r > 1$, while $\xi = 0$ is a pole of order 2 for the second integrand in (8.5). For the complex case, we directly have $\mathbb{E}[X(f_1,t)] = \mathbb{E}[X(f_2,t)] = 0$.
For $f_2(x) = x^2$, we have

$$\mathbb{E}[X(f_2,t)] = I_1 - I_2,$$

where

$$\begin{aligned}
I_1 =& \frac{t^2}{2\pi i} \lim_{r \searrow 1} \oint\limits_{|\xi|=1} \frac{\left(\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi\right)^2}{\xi \left(\xi - r^{-1}\right)\left(\xi + r^{-1}\right)} d\xi, \\
I_2 =& \frac{t^2}{2\pi i} \lim_{r \searrow 1} \oint\limits_{|\xi|=1} \frac{\left(\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi\right)^2}{\xi^3} d\xi.
\end{aligned}$$

The integrand in I_1 has poles which are all of the order 1 at the points $0, r^{-1}, -r^{-1}$. Thus, using the residue theorem,

$$\begin{aligned}
I_1 =& t^2 \lim_{r \searrow 1} \frac{\left(\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi\right)^2}{\left(\xi - r^{-1}\right)\left(\xi + r^{-1}\right)} \Big|_{\xi = 0} + t^2 \lim_{r \searrow 1} \frac{\left(\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi\right)^2}{\xi \left(\xi + r^{-1}\right)} \Big|_{\xi = r^{-1}} \\
& + t^2 \lim_{r \searrow 1} \frac{\left(\xi + h_t r \xi^2 + h_t r^{-1} + h_t^2 \xi\right)^2}{\xi \left(\xi - r^{-1}\right)} \Big|_{\xi = -r^{-1}}
\end{aligned}$$

$$= - t^2 h_t^2 + \frac{t^2(1+h_t)^4}{2} + \frac{t^2(1-h_t)^4}{2} = -th^2 + \frac{t^2(1+h_t)^4}{2} + \frac{t^2(1-h_t)^4}{2}.$$

Using that the integrand in I_2 has a pole at $\xi = 0$ of order 3, similar calculations yield $I_2 = (1 + 4h_t^2 + h_t^4)t^2$, which gives

$$\mathbb{E}[X(f_2, t)] = th^2 = ty.$$

For the covariance function of $(X(f_1, t))_{t\in[t_0,1]}$, we have for $t_2 \leq t_1$

$$\begin{aligned}
\operatorname{cov}(X(f_1,t_1), X(f_1,t_2)) =& \frac{1}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} t_1 \left(1 + h_{t_1}\xi_1 + h_{t_1} r_1^{-1}\xi_1^{-1} + h_{t_1}^2\right) t_2 \\
& \times \left(1 + h_{t_2}\xi_2 + h_{t_2} r_2^{-1}\xi_2^{-1} + h_{t_2}^2\right) \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1 \\
=& - \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \frac{h_{t_1} t_1 t_2 (h_1 + r_1\xi_1 + h_1^2 r_1 \xi_1 + h_1 r_1^2 \xi_1^2)}{r_1^2 r_2^2 \xi_1^3} d\xi_1 \\
=& - \frac{(2\pi i)^2}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \frac{\partial^2}{\partial^2 \xi_1} \frac{h_{t_1} t_1 t_2 (h_1 + r_1\xi_1 + h_1^2 r_1 \xi_1 + h_1 r_1^2 \xi_1^2)}{r_1^2 r_2^2} \Big|_{\xi_1=0} \\
=& 2 \lim_{\substack{r_1>r_2,\\ r_1,r_2\searrow 1}} \frac{h_1^2 t_2}{r_2^2} = 2h_1^2 t_2 = 2yt_2,
\end{aligned}$$

where we used a computer algebra system for simplifying the first integrand and then applied the residue theorem twice. Considering the function f_2, we have for $(t_2 \leq t_1)$

$$\begin{aligned}
& \operatorname{cov}(X(f_2,t_1), X(f_2,t_2)) \\
=& \frac{1}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} t_1^2 \left(1 + h_{t_1}\xi_1 + h_{t_1} r_1^{-1}\xi_1^{-1} + h_{t_1}^2\right)^2 t_2^2 \\
& \times \left(1 + h_{t_2}\xi_2 + h_{t_2} r_2^{-1}\xi_2^{-1} + h_{t_2}^2\right)^2 \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1 \\
=& \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \frac{2h_1 t_1 t_2^2}{r_1^4 r_2^4 \xi_1^5} \left(h_1 + r_1 x + h_1^2 r_1 x + h_1 r_1^2 \xi_1^2\right)^2 \\
& \times \left(-h_1 t_1 - h_1^2 r_1 t_1 \xi_1 - r_1 r_2^2 t_1 \xi_1 - h_2^2 r_1 r_2^2 t_1 \xi_1 + h_2^2 r_1 t_2 \xi_1 + h_1^2 r_1^3 t_1 \xi_1^3 - h_2^2 r_1^3 t_2 \xi_1^3\right) d\xi_1
\end{aligned} \tag{8.6}$$

$$\begin{aligned}
=& \frac{(2\pi i)^2}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \frac{1}{(5-1)!} \frac{\partial^4}{\partial^4 \xi_1} \Big\{ \frac{2h_1 t_1 t_2^2}{r_1^4 r_2^4} \left(h_1 + r_1\xi_1 + h_1^2 r_1 \xi_1 + h_1 r_1^2 \xi_1^2\right)^2 \\
& \times \left(-h_1 t_1 - h_1^2 r_1 t_1 \xi_1 - r_1 r_2^2 t_1 \xi_1 - h_2^2 r_1 r_2^2 t_1 \xi_1 + h_2^2 r_1 t_2 \xi_1 + h_1^2 r_1^3 t_1 \xi_1^3 - h_2^2 r_1^3 t_2 \xi_1^3\right) d\xi_1 \Big\} \Big|_{\xi_1=0} \\
=& 4t_2 y \left\{2t_1 t_2 + [t_2 + 2(t_1 + t_2)]\, y + 2y^2\right\}, \ t_2 \leq t_1.
\end{aligned}$$

Note that $\xi_1 = 0$ is a pole of order 5 for the integrand in (8.6) and that in the special case $t_1 = t_2 = 1$ we recover the mean and covariance given in (9.8.14) and,

respectively, (9.8.15) in Bai and Silverstein (2010).
Finally, we want to calculate the dependence structure between $X(f_1,t_1)$ and $X(f_2,t_2)$. Using similar techniques as above, we obtain for $t_2 \leq t_1$

$$\begin{aligned}
&\operatorname{cov}(X(f_1,t_1),X(f_2,t_2)) \\
=&\frac{1}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint_{|\xi_1|=1}\oint_{|\xi_2|=1} t_1\left(1+h_{t_1}\xi_1+h_{t_1}r_1^{-1}\xi_1^{-1}+h_{t_1}^2\right)t_2^2 \\
&\qquad\times\left(1+h_{t_2}\xi_2+h_{t_2}r_2^{-1}\xi_2^{-1}+h_{t_2}^2\right)^2\frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)}d\xi_2 d\xi_1 \qquad (8.7)\\
=&\frac{2\pi i}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint_{|\xi_1|=1}\frac{1}{r_1^3r_2^4\xi_1^4}2h_1t_2^2(h_1+r_1\xi_1+h_1^2r_1\xi_1+h_1r_1^2\xi_1^2) \\
&\times\left\{-h_1t_1+h_1^2r_1t_1\xi_1(-1+r_1^2\xi_1^2)-r_1\xi_1\left[(1+h_2^2)r_2^2t_1+h_2^2t_2(-1+r_1^2\xi_1^2)\right]\right\}d\xi_1 \qquad (8.8)\\
=&4t_2y(t_2+y). \qquad (8.9)
\end{aligned}$$

After simplifying the integrand in (8.7) with a computer algebra program, we see that it has a pole at $\xi_2=0$ of order 2. Note that the pole at

$$\xi_2=\frac{h_2r_2}{h_1r_1\xi_1}$$

is not relevant for an application of the residue theorem, since

$$\left|\frac{h_2r_2}{h_1r_1\xi_1}\right|=\left|\frac{\sqrt{t_1}r_2}{r_1\sqrt{t_2}\xi_1}\right|=\frac{\sqrt{t_1}r_2}{\sqrt{t_2}r_1}>1.$$

The integrand in (8.8) has a pole at $\xi_1=0$ of order 4. Similarly, we have, again for $t_2\leq t_1$,

$$\operatorname{cov}(X(f_2,t_1),X(f_1,t_2))=4t_2y(t_1+y). \qquad (8.10)$$

By combining (8.9) and (8.10), we have for $t_1,t_2\in[t_0,1]$

$$\operatorname{cov}(X(f_1,t_1),X(f_2,t_2))=4\min(t_1,t_2)y(t_2+y).$$

8.2.3 Proof of Corollary 4.3.1

We apply Theorem 3.2.1 for the choice $h(x)=\log(x)$. Note that, if $y\geq t_0$, the interval in (3.7) contains the point 0. Thus, we have to impose $y<t_0$, since h is not analytic in a neighborhood of 0.
Using Example 2.11 in Yao et al. (2015), we obtain for the centering term

$$\begin{aligned}
\int\log x d\tilde{F}^{y_{\lfloor nt\rfloor}}(x)&=\int\log x dF^{y_{\lfloor nt\rfloor}}\Big(\frac{n}{\lfloor nt\rfloor}x\Big)=\int\log x dF^{y_{\lfloor nt\rfloor}}(x)+\log\Big(\frac{\lfloor nt\rfloor}{n}\Big)\\
&=\Big(-1+\frac{y_{\lfloor nt\rfloor}-1}{y_{\lfloor nt\rfloor}}\log(1-y_{\lfloor nt\rfloor})\Big)+\log\Big(\frac{\lfloor nt\rfloor}{n}\Big)
\end{aligned}$$

$$= -1 - \frac{1}{y_{\lfloor nt \rfloor}} \log(1 - y_{\lfloor nt \rfloor}) + \log\Big(\frac{\lfloor nt \rfloor}{n} - y_n\Big),$$

which implies

$$p \int \log x d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = -p - \lfloor nt \rfloor \log(1 - y_{\lfloor nt \rfloor}) + p \log\left(\frac{\lfloor nt \rfloor}{n} - y_n\right).$$

By Proposition 4.1.1, we have for the mean of the limiting process $\mathbb{D}$ in the real case

$$\mathbb{E}[\mathbb{D}(t)] = I_1 + I_2, \tag{8.11}$$

where

$$\begin{aligned} I_1 =& \frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|\xi|=1} \log(t(1 + h_t r\xi + h_t r^{-1}\xi^{-1} + h_t^2)) \frac{\xi}{\xi^2 - r^{-2}} d\xi \\ =& \frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|\xi|=1} \log(t|1 + h_t \xi|^2) \frac{\xi}{\xi^2 - r^{-2}} d\xi, \end{aligned} \tag{8.12}$$

$$\begin{aligned} I_2 =& -\frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|\xi|=1} \log(t(1 + h_t r\xi + h_t r^{-1}\xi^{-1} + h_t^2)) \frac{1}{\xi} d\xi \\ =& -\frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|\xi|=1} \log(t|1 + h_t \xi|^2) \frac{1}{\xi} d\xi \end{aligned}$$

(see also Wang and Yao (2013) for a similar representation). Beginning with I_1, we further decompose (note that for $|\xi| = 1$, it holds $\xi^{-1} = \overline{\xi}$)

$$I_1 = I_{11} + I_{12},$$

where

$$\begin{aligned} I_{11} =& \frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|\xi|=1} \log(t(1 + h_t \xi)) \frac{\xi}{(\xi - r^{-1})(\xi + r^{-1})} d\xi, \\ =& \frac{1}{2} \lim_{r \searrow 1} \left\{ \log\left(t\left(1 + h_t r^{-1}\right)\right) + \log\left(t\left(1 - h_t r^{-1}\right)\right) \right\} = \frac{1}{2} \log\left(t^2\left(1 - h_t^2\right)\right), \\ I_{12} =& \frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|\xi|=1} \log(t(1 + h_t \xi^{-1})) \frac{\xi}{(\xi - r^{-1})(\xi + r^{-1})} d\xi \\ =& \frac{1}{2\pi i} \lim_{r \searrow 1} \oint_{|z|=1} \log(t(1 + h_t z)) \frac{r^2}{z(z - r)(r + z)} dz \\ =& \lim_{r \searrow 1} \log(t(1 + h_t z)) \frac{r^2}{(z - r)(z + r)} \Big|_{z=0} = -\log(t). \end{aligned}$$

These calculations imply

$$I_1 = \frac{1}{2}\log\left(t^2\left(1-h_t^2\right)\right) - \log(t). \tag{8.13}$$

The quantity I_2 in (8.11) can be determined similarly using the decomposition

$$I_2 = I_{21} + I_{22},$$

where

$$I_{21} = -\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1} \log(t(1+h_t\xi))\frac{1}{\xi}d\xi = -\log t$$

$$I_{22} = -\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1} \log(t(1+h_t\xi^{-1}))\frac{1}{\xi}d\xi = \log t.$$

This gives $I_2 = 0$, and by (8.13) and (8.11), we obtain

$$\mathbb{E}[\mathbb{D}(t)] = \frac{1}{2}\log\left(t^2\left(1-h_t^2\right)\right) - \log(t) = \frac{1}{2}\log\left(1-h_t^2\right) = \frac{1}{2}\log\left(1-y_t\right).$$

Next, we calculate the covariance structure. Similarly to (8.12), we obtain for $t_2 \le t_1$

$$\begin{aligned}
&\operatorname{cov}(\mathbb{D}(t_1),\mathbb{D}(t_2))\\
=&\frac{1}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint\limits_{|\xi_1|=1}\oint\limits_{|\xi_2|=1} \log(t_1|1+h_{t_1}\xi_1|^2)\overline{\log(t_2|1+h_{t_2}\xi_2|^2)}\frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)}d\xi_2 d\xi_1\\
=&\frac{1}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint\limits_{|\xi_1|=1}\oint\limits_{|\xi_2|=1} \log(t_1|1+h_{t_1}\xi_1|^2)\log(t_2|1+h_{t_2}\xi_2|^2)\frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)}d\xi_2 d\xi_1\\
=&I_3+I_4,
\end{aligned}$$

where (note that $|1+h_{t_2}\xi_2|^2 = (1+h_{t_2}\xi_2)(1+h_{t_2}\xi_2^{-1})$)

$$\begin{aligned}
I_3 =&\frac{1}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint\limits_{|\xi_1|=1} \log(t_1|1+h_{t_1}\xi_1|^2)\oint\limits_{|\xi_2|=1} \log(t_2(1+h_{t_2}\xi_2))\frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)}d\xi_2 d\xi_1,\\
I_4 =&\frac{1}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint\limits_{|\xi_1|=1} \log(t_1|1+h_{t_1}\xi_1|^2)\oint\limits_{|\xi_2|=1} \log(t_2(1+h_{t_2}\xi_2^{-1}))\frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)}d\xi_2 d\xi_1.
\end{aligned}$$

Using a computer algebra program for simplifying I_3 and I_4, we see that $I_3 = 0$ and for I_4, and we perform the substitution $\xi_2 = z_2^{-1}$, which yields

$$I_4 = \frac{2\pi i}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint\limits_{|\xi_1|=1} \log(t_1|1+h_{t_1}\xi_1|^2)\frac{h_1 r_1}{r_2+h_1 r_1\xi_1}d\xi_1 = I_{41}+I_{42},$$

where

$$I_{41} = \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2 > r_1, \\ r_1, r_2 \searrow 1}} \oint_{|\xi_1|=1} \log(t_1(1 + h_{t_1}\xi_1)) \frac{h_1 r_1}{r_2 + h_1 r_1 \xi_1} d\xi_1,$$

$$I_{42} = \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2 > r_1, \\ r_1, r_2 \searrow 1}} \oint_{|\xi_1|=1} \log(t_1(1 + h_{t_1}\xi_1^{-1})) \frac{h_1 r_1}{r_2 + h_1 r_1 \xi_1} d\xi_1.$$

It holds $I_{41} = 0$, since we have for the pole at $\xi_1 = -r_2/(h_1 r_1)$ that $|\xi_1|^2 > \frac{1}{h_{t_1}^2} = \frac{t_1}{y} \geq \frac{t_0}{y} \geq 1$.

As above, we perform for I_{42} the substitution $\xi_1^{-1} = z_1$ and obtain

$$\begin{aligned} I_{42} = & -\frac{2\pi i}{2\pi^2} \lim_{\substack{r_2 > r_1, \\ r_1, r_2 \searrow 1}} \oint_{|z_1|=1} \log(t_1(1 + h_{t_1} z_1)) \frac{h_{t_1} r_1}{h_{t_1} r_1 z_1 + r_2 z_1^2} dz_1 \\ = & -\frac{(2\pi i)^2}{2\pi} \lim_{\substack{r_2 > r_1, \\ r_1, r_2 \searrow 1}} \left\{ -\log(t_1) + \log\left(t_1\left(1 - \frac{h_{t_1}^2 r_1}{r_2}\right)\right)\right\} \\ = & -2\log(1 - h_{t_1}^2). \end{aligned}$$

Finally, we obtain for $t_2 \leq t_1$

$$\operatorname{cov}(\mathbb{D}(t_1), \mathbb{D}(t_2)) = I_3 + I_4 = -2\log(1 - h_{t_1}^2) = -2\log(1 - y_{t_1}).$$

Chapter 9

Gaussian fluctuations for diagonal entries of a large sample precision matrix

In this chapter, we prove joint asymptotic normality for several diagonal entries of the inverse of the sample covariance matrix, the so-called sample precision matrix. An introduction to this problem and its connection to linear spectral statistics of sample covariance matrices can be found in Section 9.1. Some notation is introduced in 9.2. The CLT for a single diagonal entry given in Section 9.3 is generalized in Section 9.4 to the joint convergence of several diagonal entries. All proofs are provided in Section 9.6. Section 9.7 sheds light on the QR-decomposition of the data matrix, which is an important tool used in the proofs.

9.1 Introduction

In this chapter, we establish a central limit theorem for the diagonal entries of a large sample precision matrix. In order to ensure that this object is well-defined, we assume that $p < n$ and the data generating process follows a continuous distribution, so that the resulting sample covariance matrix is an invertible matrix. When investigating the diagonal entries of the sample precision matrix, an immediate connection to linear spectral statistics of the sample covariance matrix is noteworthy. Recall Cramer's rule, which gives the representation

$$(\mathbf{A}^{-1})_{qq} = \frac{|\mathbf{A}^{(-q)}|}{|\mathbf{A}|}, \ 1 \leq q \leq p,$$

for the diagonal elements of $\mathbf{A}^{-1}$ in terms of minors of an invertible matrix $\mathbf{A} \in \mathbb{R}^{p \times p}$. Here, $\mathbf{A}^{(-q)}$ denotes the $(p-1) \times (p-1)$ submatrix of $\mathbf{A}$ where the qth row and qth column are deleted. If $\hat{\mathbf{\Sigma}} = \mathbf{B}_{n,1}$ denotes some sample covariance matrix, then we obtain by an application of Cramer's rule

$$\log(\hat{\mathbf{\Sigma}}^{-1})_{qq} = \log |\hat{\mathbf{\Sigma}}^{(-q)}| - \log |\hat{\mathbf{\Sigma}}|, \tag{9.1}$$

which turns out to be a difference of linear spectral statistics of $\hat{\mathbf{\Sigma}}$ and its submatrix $\hat{\mathbf{\Sigma}}^{(-q)} \in \mathbb{R}^{(p-1)\times(p-1)}$. A further interesting connection to linear spectral statistics of the sample covariance matrix is the fact that the limiting variance of $(\hat{\mathbf{\Sigma}}^{-1})_{qq}$ is determined by the fourth moment of the underlying data generating distribution. Due to the strong dependence between the eigenvalues of $\hat{\mathbf{\Sigma}}$ and $\hat{\mathbf{\Sigma}}^{(-q)}$, the asymptotic behavior of (9.1) cannot be investigated by using techniques established for the proof of Theorem 3.2.1, and in particular, such statistics are not covered by Bai and Silverstein's CLT (Bai and Silverstein, 2010). In fact, we will observe that the difference in (9.1) fluctuates on a scale $1/\sqrt{n}$ which is of significantly smaller order than the fluctuations of each single linear spectral statistic $\log|\hat{\mathbf{\Sigma}}|$ and $\log|\hat{\mathbf{\Sigma}}^{(-q)}|$. In order to tackle the difficulties arising for this difference of two dependent linear spectral statistics, we perform a QR-decomposition for the data matrix which is useful in a broader context in random matrix theory: Wang et al. (2018) used this tool in order to derive the logarithmic law of the sample covariance matrix for the case $p/n \to 1$ near singularity, while Heiny and Parolya (2021) investigated the log-determinant of the sample correlation matrix under infinite fourth moment. These papers were partially inspired by works of Nguyen and Vu (2014) and Bao et al. (2015b), in which the authors proved Girko's logarithmic law for a general random matrix with independent entries and brought his "method of perpendiculars" (see Girko, 1998) to a mathematically rigorous level. Using such a QR-decomposition for the data matrix in our setting, we derive a representation of the diagonal entry as the inverse of a quadratic form. With this knowledge in hand, we show that a central limit theorem for martingale difference schemes is applicable to the quadratic form. By the delta method, we finally get asymptotic normality for $(\hat{\mathbf{\Sigma}}^{-1})_{qq}$ being the inverse of this quadratic form. In this chapter, we will also consider the joint asymptotic distribution of several diagonal entries, which calls for particular attention due to the dependencies of the diagonal elements of the sample precision matrix.
Considering the special case of a sample precision matrix based on a sample following a multivariate normal distribution, the exact distribution of $(\hat{\mathbf{\Sigma}}^{-1})_{qr}$ is well-understood in the literature for fixed dimension and sample size ($1 \leq q, r \leq p$). In fact, $n^{-1}\hat{\mathbf{\Sigma}}$ is said to follow an inverse Wishart distribution. For more details on this matrix-valued distribution, we refer the reader to Von Rosen (1988), Nydick (2012) and Gupta and Nagar (2018).
In the Gaussian case, the entries of the population precision matrix capture the dependence of two components of a random vector conditionally on all others. More precisely, if $\mathbf{x} = (x_1, \dots, x_p)^\top \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$, then the coordinates x_i and x_j are independent conditionally on $\{x_1, \dots, x_p\} \setminus \{x_i, x_j\}$ if and only if $(\mathbf{\Sigma}^{-1})_{ij} = 0$ (see Lauritzen, 1996). A popular field using precision matrices is Gaussian graphical models, where the vertices of a graph represent the different coordinates. Two vertices i and j are connected with an edge if and only if the (i, j)th entry of the estimated precision matrix does not vanish. Beyond normally distributed data, this one-to-one correspondence between conditional independence and precision matrix does not hold true in general. However, for various statistical problems such as linear regression, linear prediction, kriging and partial correlation, the behavior of the precision matrix is crucial (see Huang et al., 2010; Van de Geer et al., 2014; Chang

et al., 2018; Huang et al., 2021, among many others).
Motivated by its importance in statistics, several authors investigated the sample precision matrix using techniques from random matrix theory. While the asymptotic behavior of specific entries of the sample precision matrix has received less attention in the literature so far, some works are devoted to the investigation of its spectral properties. Zheng et al. (2015a) established a central limit theorem for linear spectral statistics of a rescaled version of the sample precision matrix. If the dimension exceeds the sample size, the sample covariance matrix is singular and in this case, the sample precision matrix can be defined as a generalized inverse of the sample precision matrix: Bodnar et al. (2016) concentrate on linear spectral statistics of the Moore-Penrose inverse of the sample covariance matrix.

9.2 Notation

In this section, we introduce the sample precision matrix formally in order to formulate the central limit theorem for its diagonal entries. In contrast to previous chapters, note that we use for the sample covariance matrix $\mathbf{B}_{n,1}$ either the notation $\hat{\mathbf{\Sigma}}$ or $\hat{\mathbf{I}}$ in order to differentiate clearly between the cases of a general population covariance matrix $\mathbf{T} = \mathbf{\Sigma}$ and the special case $\mathbf{T} = \mathbf{\Sigma} = \mathbf{I}$.
Let

$$\mathbf{X}_n = (x_{ij})_{\substack{i=1,\ldots,p \\ j=1,\ldots,n}} = (\mathbf{b}_1, \ldots, \mathbf{b}_p)^\top = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^{p\times n}$$

be a random matrix with i.i.d. entries following a continuous distribution, $\mathbf{\Sigma} = \mathbf{\Sigma}_n \in \mathbb{R}^{p\times p}$ nonrandom and (symmetric) positive definite matrix with symmetric square root $\mathbf{\Sigma}^{1/2}$. The matrix $\mathbf{\Sigma}$ denotes the population covariance matrix and for most of the following results, it is assumed to be a diagonal matrix (except for the normal case). Define

$$\hat{\mathbf{I}} = \frac{1}{n}\mathbf{X}_n\mathbf{X}_n^\top = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^\top \in \mathbb{R}^{p\times p}.$$

If we set for some $q \in \{1, \ldots, p\}$

$$\tilde{\mathbf{X}}_n^{(-q)} = (\mathbf{b}_1, \ldots, \mathbf{b}_{q-1}, \mathbf{b}_{q+1}, \ldots, \mathbf{b}_p)^\top \in \mathbb{R}^{(p-1)\times n},$$

then

$$\hat{\mathbf{I}}^{(-q)} = \frac{1}{n}\tilde{\mathbf{X}}_n^{(-q)} \left(\tilde{\mathbf{X}}_n^{(-q)}\right)^\top \in \mathbb{R}^{(p-1)\times(p-1)}$$

can be obtained from $\hat{\mathbf{I}}$ by deleting the qth row and the qth column. Similarly, if we set $\mathbf{Y}_n = \mathbf{\Sigma}^{\frac{1}{2}}\mathbf{X}_n = (\mathbf{d}_1, \ldots, \mathbf{d}_p)^\top \in \mathbb{R}^{p\times n}$ and $\tilde{\mathbf{Y}}_n^{(-q)} = (\mathbf{d}_1, \ldots, \mathbf{d}_{q-1}, \mathbf{d}_{q+1}, \ldots, \mathbf{d}_p)^\top$, we define

$$\hat{\mathbf{\Sigma}} = \frac{1}{n}\mathbf{Y}_n\mathbf{Y}_n^\top \text{ and } \hat{\mathbf{\Sigma}}^{(-q)} = \frac{1}{n}\tilde{\mathbf{Y}}_n^{(-q)}\left(\tilde{\mathbf{Y}}_n^{(-q)}\right)^\top.$$

Additionally, the matrix $\mathbf{\Sigma}^{(-q)} \in \mathbb{R}^{(p-1)\times(p-1)}$ can be obtained from $\mathbf{\Sigma}$ by deleting the qth row and the qth column. Moreover, we denote by

$$\mathbf{P}(q) = \mathbf{I} - \tilde{\mathbf{X}}_{n,q}^\top \left(\tilde{\mathbf{X}}_{n,q}\tilde{\mathbf{X}}_{n,q}^\top\right)^{-1} \tilde{\mathbf{X}}_{n,q} \in \mathbb{R}^{n\times n},$$

the projection matrix on the orthogonal complement of the subspace generated by the first q rows of $\mathbf{X}_n$, that is,

$$\tilde{\mathbf{X}}_{n,q} = (\mathbf{b}_1, \ldots, \mathbf{b}_q)^\top \in \mathbb{R}^{q\times n}. \tag{9.2}$$

9.3 CLT for a single diagonal entry

All proofs for results of this section are deferred to Section 9.5.

Theorem 9.3.1 (CLT for diagonal entries of full-sample precision matrix) *Let* $\mathbf{\Sigma} \in \mathbb{R}^{p\times p}$ *be a diagonal matrix with positive diagonal entries. Assume that the random variables* x_{ij} *are i.i.d. with* $\mathbb{E}[x_{11}] = 0$, $\operatorname{Var}(x_{11}) = 1$ *and* $1 < \mathbb{E}[x_{11}^4] = \nu_4 < \infty$ *for* $1 \le i \le p$, $1 \le j \le n$*. Then, it holds for* $n \to \infty, p/n \to y \in [0, 1)$ *and* $q \in \{1, \ldots, p\}$

$$\frac{\sqrt{n-p+1}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}} \left(\frac{n-p+1}{n} \left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq} - \left(\mathbf{\Sigma}^{-1}\right)_{qq} \right) \xrightarrow{\mathcal{D}} W \sim \mathcal{N}(0, \rho),\ n \to \infty,$$

where $\rho = 2 + (\nu_4 - 3)(1 - y)$.

Remark 9.3.2 It is of interest to compare this statement with related results in the literature. The statistic

$$\log \left(\hat{\mathbf{I}}^{-1}\right)_{qq} = \log \left|\hat{\mathbf{I}}\right| - \log \left|\hat{\mathbf{I}}^{(-q)}\right| \tag{9.3}$$

can be interpreted as a difference of two linear spectral statistics of sample covariances matrices and a CLT for this random variable would yield a CLT for $\left(\hat{\mathbf{I}}^{-1}\right)_{qq}$ via the delta method. Recently, Cipolloni and Erdős (2018) developed a CLT for the difference of general linear spectral statistics of a sample covariance matrix and its minor, which is applicable to a standardized and centered version of $\log \left(\hat{\mathbf{I}}^{-1}\right)_{qq}$. To be precise, when applying Theorem 2.2 of their work, the random variables x_{ij} are assumed to admit finite moments of all order. Moreover, the asymptotic regime of p, n is more restrictive in comparison to Theorem 9.3.1 and does not include the case $p/n \to 0$ for $n \to \infty$. In contrast to Theorem 9.3.1, the authors do not need the limit y of p/n to exist. While this assumption allows us to determine the limiting variance ρ, it is not necessary for proving a CLT as in Theorem 9.3.1 (one could instead divide by $1/\sqrt{\rho_n}$ defined in (9.9)). We would also like to emphasize that the technique used for proving Theorem 9.3.1 sets us in the position to investigate the joint convergence of several diagonal elements of the sample precision matrix given in Theorem 9.4.1.

We may allow for a general form of the population covariance matrix $\mathbf{\Sigma}$ when imposing a normal assumption on the data. The following corollary follows directly from Theorem 9.3.1 and Lemma 9.5.2 given in Section 9.5.1.

Corollary 9.3.3 *Let $\mathbf{\Sigma} \in \mathbb{R}^{p\times p}$ be a symmetric positive definite matrix and assume that $x_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ for $1 \le i \le p,\ 1 \le j \le n$. Then, it holds for $n \to \infty, p/n \to y \in [0,1)$ and $q \in \{1,\dots,p\}$*

$$\frac{\sqrt{n-p+1}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}}\left(\frac{n-p+1}{n}\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq} - \left(\mathbf{\Sigma}^{-1}\right)_{qq}\right) \xrightarrow{\mathcal{D}} W \sim \mathcal{N}(0,2),\ n \to \infty.$$

Concluding this section, we investigate a simple subsampling strategy (see, e.g., Ma et al., 2014; Raskutti and Mahoney, 2016; Wang et al., 2019; Wang and Ma, 2021, for subsampling procedures in various settings). For this purpose, we need to introduce some further notation. Choose $m \in \mathbb{N}$ with $m \le n$ for the size of the subsample and define the set

$$\mathcal{P}(n,m) = \{A \in 2^{[n]} : |A| = m\}.$$

Let U_n be the uniform distribution on $\mathcal{P}(n,m)$, that is, each subset with cardinality m of $\{1,\dots,n\}$ occurs with probability $1/\binom{n}{m}$. We define

$$\delta_i = I_{\{i\in U_n\}}, 1 \le i \le n.$$

The subsample covariance matrix is then defined as

$$\breve{\mathbf{I}} = \frac{1}{m}\sum_{i=1}^{n} \delta_i \mathbf{x}_i\mathbf{x}_i^\top = \frac{1}{m}\sum_{i\in U_n} \mathbf{x}_i\mathbf{x}_i^\top.$$

Let

$$\breve{\mathbf{X}}_n = (\mathbf{x}_j)_{j\in U_n} = (\breve{\mathbf{b}}_1,\dots,\breve{\mathbf{b}}_p)^\top \in \mathbb{R}^{m\times p},$$

where $\breve{\mathbf{b}}_1,\dots,\breve{\mathbf{b}}_p$ denote the subsampled data, that is, $\breve{\mathbf{b}}_j = (x_{ij})_{i\in U_n}^\top \in \mathbb{R}^m,\ 1 \le i \le p$. We denote the projection matrix on the orthogonal complement of the subspace generated by the first q rows of $\breve{\mathbf{X}}$, that is,

$$\breve{\mathbf{P}}(q) = \mathbf{I} - \breve{\mathbf{X}}_{n,q}^\top\left(\breve{\mathbf{X}}_{n,q}\breve{\mathbf{X}}_{n,q}^\top\right)^{-1}\breve{\mathbf{X}}_{n,q} \in \mathbb{R}^{m\times m},$$

where we set similarly to (9.2)

$$\breve{\mathbf{X}}_{n,q} = (\breve{\mathbf{b}}_1,\dots,\breve{\mathbf{b}}_q)^\top \in \mathbb{R}^{q\times m}.$$

Similarly as before, if we set $\breve{\mathbf{Y}}_n = \mathbf{\Sigma}^{\frac{1}{2}}\breve{\mathbf{X}}_n = (\breve{\mathbf{d}}_1,\dots,\breve{\mathbf{d}}_p)^\top \in \mathbb{R}^{p\times m}$ and $\breve{\mathbf{Y}}_n^{(-q)} = (\breve{\mathbf{d}}_1,\dots,\breve{\mathbf{d}}_{q-1},\breve{\mathbf{d}}_{q+1},\dots,\breve{\mathbf{d}}_p)^\top \in \mathbb{R}^{(p-1)\times m}$, we define

$$\breve{\mathbf{\Sigma}} = \frac{1}{m}\breve{\mathbf{Y}}_n\breve{\mathbf{Y}}_n^\top \text{ and } \breve{\mathbf{\Sigma}}^{(-q)} = \frac{1}{m}\breve{\mathbf{Y}}_n^{(-q)}\left(\breve{\mathbf{Y}}_n^{(-q)}\right)^\top.$$

Theorem 9.3.4 (CLT for diagonal entries of subsample precision matrix) *Let* $\mathbf{\Sigma} \in \mathbb{R}^{p\times p}$ *be a diagonal matrix with positive diagonal entries. Assume that the random variables* x_{ij} *are i.i.d. with* $\mathbb{E}[x_{11}] = 0$, $\operatorname{Var}(x_{11}) = 1$ *and* $1 < \mathbb{E}[x_{11}^4] = \nu_4 < \infty$ *for* $1 \le i \le p$, $1 \le j \le n$. *Then, if* $p/m \overset{n\to\infty}{\to} \gamma \in [0,1), m = m_n \overset{n\to\infty}{\to} \infty$ *and* $q \in \{1,\dots,p\}$

$$\frac{\sqrt{m-p+1}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}}\left(\frac{m-p+1}{m}\left(\breve{\mathbf{\Sigma}}^{-1}\right)_{qq} - \left(\mathbf{\Sigma}^{-1}\right)_{qq}\right) \overset{\mathcal{D}}{\to} \mathcal{N}(0,\breve{\rho}),\ n\to\infty.$$

where $\breve{\rho} = 2 + (\nu_4 - 3)(1-\gamma)$.

9.4 Joint convergence of diagonal entries

The next theorem presents the joint asymptotic distribution of two diagonal entries and is proven in Section 9.6.

Theorem 9.4.1 *Let* $\mathbf{\Sigma} \in \mathbb{R}^{p\times p}$ *be a diagonal matrix with positive diagonal entries. Assume that the random variables* x_{ij} *are i.i.d. with* $\mathbb{E}[x_{11}] = 0$, $\operatorname{Var}(x_{11}) = 1$ *and* $1 < \mathbb{E}[x_{11}^4] = \nu_4 < \infty$ *for* $1 \le i \le p$, $1 \le j \le n$. *Then, it holds for* $n \to \infty, p/n \to y \in [0,1)$ *and* $1 \le q_1 \ne q_2 \le p$

$$\left\{\frac{\sqrt{n-p+1}}{\left(\mathbf{\Sigma}^{-1}\right)_{ii}}\left(\frac{n-p+1}{n}\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{ii} - \left(\mathbf{\Sigma}^{-1}\right)_{ii}\right)\right\}_{i=q_1,q_2}^{\top} \overset{\mathcal{D}}{\to} \mathcal{N}_2\left(\mathbf{0},\rho\mathbf{I}_2\right),\ n\to\infty,$$

where $\rho = 2 + (\nu_4 - 3)(1-y)$.

Remark 9.4.2 Note that Theorem 9.4.1 provides a nontrivial generalization of Theorem 9.3.1 since the diagonal entries are not independent. For more details on the concrete dependence structure, we refer the reader to Lemma 9.6.1 and Lemma 9.6.2. However, it is noteworthy that these random variables are asymptotically independent in the sense that

$$\lim_{n\to\infty}\mathbb{P}(Z_{n,q_1}\in A, Z_{n,q_2}\in B) = \lim_{n\to\infty}\mathbb{P}(Z_{n,q_1}\in A)\lim_{n\to\infty}\mathbb{P}(Z_{n,q_2}\in B),\ 1\le q_1\ne q_2\le p,$$

for any Borel sets $A, B \subset \mathbb{R}$, which is a consequence of Theorem 9.3.1 and Theorem 9.4.1. Here, Z_{n,q_1} denotes a transformation of the q_1th diagonal element of the sample precision matrix, that is,

$$Z_{n,q_1} = \frac{\sqrt{n-p+1}}{\left(\mathbf{\Sigma}^{-1}\right)_{q_1,q_1}}\left(\frac{n-p+1}{n}\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{q_1,q_1} - \left(\mathbf{\Sigma}^{-1}\right)_{q_1,q_1}\right).$$

9.5 Proofs of results in Section 9.3

We continue by proving Theorem 9.3.1 using a CLT for martingale difference schemes. Subsequently, these ideas are generalized for the subsampling case in Theorem 9.3.4. The auxiliary results for these proofs can be found in Section 9.5.1.

Proof of Theorem 9.3.1. Using Lemma 9.5.1 and noting that the distribution of $\mathbf{X}_n$ is invariant under a permutation of the qth and the pth row, we see that

$$\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}} = \left(\hat{\mathbf{I}}^{-1}\right)_{qq} \stackrel{\mathcal{D}}{=} \left(\hat{\mathbf{I}}^{-1}\right)_{pp} = \frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{pp}}{\left(\mathbf{\Sigma}^{-1}\right)_{pp}}.$$

Thus, we may assume $q = p$ without loss of generality. From now on, the proof is divided in several steps.

Step 1: QR decomposition

In this step, we rewrite $|\hat{\mathbf{I}}|$ and $|\hat{\mathbf{I}}^{(-p)}|$ in a more handy form via the QR decomposition. More details on this decomposition can be found in Section 9.7.

As explained in detail in Section 9.7, we get by proceeding the QR-decomposition for $\mathbf{X}_n^\top$

$$\mathbf{X}_n^\top = \mathbf{Q}\mathbf{R},\ \mathbf{X}_n = \mathbf{R}^\top\mathbf{Q}^\top, \tag{9.4}$$

where $\mathbf{Q} = (\mathbf{e}_1, \ldots, \mathbf{e}_p) \in \mathbb{R}^{n\times p}$ denotes a matrix with orthonormal columns satisfying $\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} \in \mathbb{R}^{p\times p}$ is an upper triangular matrix with entries $r_{ij} = (\mathbf{e}_i, \mathbf{b}_j)$ for $i \le j$ and $r_{ij} = 0$ for $i > j$, $i,j \in \{1,\ldots,p\}$. Note that, since $\left(\tilde{\mathbf{X}}_n^{(-p)}\right)^\top$ is the same as $\mathbf{X}_n^\top$ but with the pth column $\mathbf{b}_p$ removed, we have

$$\left(\tilde{\mathbf{X}}_n^{(-p)}\right)^\top = \mathbf{Q}\tilde{\mathbf{R}},\ \tilde{\mathbf{X}}_n^{(-p)} = \tilde{\mathbf{R}}^\top\mathbf{Q}^\top, \tag{9.5}$$

where $\tilde{\mathbf{R}} = (r_{ij})_{\substack{1\le i\le p,\\ 1\le j\le p-1}} \in \mathbb{R}^{p\times(p-1)}$ and we set $\tilde{\mathbf{R}}^{(-p)} = (r_{ij})_{1\le i,j\le p-1} \in \mathbb{R}^{(p-1)\times(p-1)}$. Using (9.4), we write

$$|\mathbf{X}_n\mathbf{X}_n^\top| = |\mathbf{R}^\top\mathbf{Q}^\top\mathbf{Q}\mathbf{R}| = |\mathbf{R}^\top\mathbf{R}| = |\mathbf{R}|^2 = \prod_{i=1}^p r_{ii}^2$$

and similarly, by using (9.5) and the Cauchy-Binet formula,

$$\left|\tilde{\mathbf{X}}_n^{(-p)}\left(\tilde{\mathbf{X}}_n^{(-p)}\right)^\top\right| = |\tilde{\mathbf{R}}^\top\tilde{\mathbf{R}}| = |\tilde{\mathbf{R}}^{(-p)}|^2 = \prod_{\substack{i=1,\\ i\ne p}}^p r_{ii}^2.$$

Thus, we obtain from Lemma 9.5.1 and Cramer's rule

$$\left(\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}}\right)^{-1} = \left(\left(\hat{\mathbf{I}}^{-1}\right)_{pp}\right)^{-1} = \frac{|\hat{\mathbf{I}}|}{|\hat{\mathbf{I}}^{(-q)}|} = \frac{1}{n}r_{pp}^2. \tag{9.6}$$

Before continuing with Step 2 of the proof of Theorem 9.3.1, we visit as an illustrating example the normal case where the distribution of r_{pp}^2 is explicitly known.

Illustration: The normal case

If we assume additionally that $x_{ij} \sim \mathcal{N}(0,1)$ i.i.d. for $i \in \{1,\dots,p\}, j \in \{1,\dots,n\}$, then it is well-known that $r_{pp}^2 \sim \mathcal{X}_{n-p+1}$ (see, e.g., Goodman (1963) or directly use (9.25)), that is,

$$r_{pp}^2 \stackrel{\mathcal{D}}{=} \sum_{j=1}^{n-p+1} Z_j^2,$$

where Z_j are i.i.d. standard normal distributed random variables, $j \in \{1,\dots,n-p+1\}$. Thus, we are able to apply a CLT for r_{pp}^2, namely,

$$\sqrt{n-p+1}\left(\frac{1}{n-p+1}r_{pp}^2 - 1\right) = \frac{1}{\sqrt{n-p+1}}\sum_{j=1}^{n-p+1}(Z_j^2 - 1) \xrightarrow{\mathcal{D}} \mathcal{N}(0,2).$$

Applying the delta method, we get

$$\sqrt{n-p+1}\left(\frac{n-p+1}{r_{pp}^2} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,2).$$

Thus, using (9.6), we conclude

$$\begin{aligned}
&\frac{\sqrt{n-p+1}}{\left(\mathbf{\Sigma}^{-1}\right)_{pp}}\left(\frac{n-p+1}{n}\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{pp} - \left(\mathbf{\Sigma}^{-1}\right)_{pp}\right) \\
=&\sqrt{n-p+1}\left(\frac{n-p+1}{n}\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{pp}}{\left(\mathbf{\Sigma}^{-1}\right)_{pp}} - 1\right) \\
=&\sqrt{n-p+1}\left(\frac{n-p+1}{r_{pp}^2} - 1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,2). \qquad (9.7)
\end{aligned}$$

Note that in the normal case, we have $\nu_4 = 3$. Thus, we have recovered the assertion of Theorem 9.3.1 in this special case.

Step 2: CLT for quadratic forms

In this step, we will show that the random variable r_{pp}^2 meets the conditions of a CLT for martingale difference schemes. In Section 9.7, it is shown that (see (9.25))

$$r_{pp}^2 = \mathbf{b}_p^\top \mathbf{P}(p-1)\mathbf{b}_p,$$

where $\mathbf{P}(0) = \mathbf{I}_n$ and for $p > 1$

$$\mathbf{P}(q) = \mathbf{I} - \tilde{\mathbf{X}}_{n,q}^\top\left(\tilde{\mathbf{X}}_{n,q}\tilde{\mathbf{X}}_{n,q}^\top\right)^{-1}\tilde{\mathbf{X}}_{n,q} \in \mathbb{R}^{n\times n} \qquad (9.8)$$

denotes the projection matrix on the orthogonal complement of the subspace generated by the first q rows of $\mathbf{X}_n$. Note that the matrix $\tilde{\mathbf{X}}_{n,q}$ is defined in (9.2). For the following analysis, we denote $\mathbf{P}(p-1) = \mathbf{P} = (p_{ik})_{1\le i,k\le n}$, which only depends on the random variables $\mathbf{b}_1, \dots, \mathbf{b}_{p-1}$ and is independent of $\mathbf{b}_p$.

We write

$$\sqrt{\frac{n-p+1}{\rho_n}} \frac{1}{n-p+1} \left(r_{pp}^2 - (n-p+1)\right) = \frac{1}{\sqrt{\rho_n(n-p+1)}} \left(\mathbf{b}_p^\top \mathbf{P} \mathbf{b_p} - \mathbb{E}\left[\mathbf{b}_p^\top \mathbf{P} \mathbf{b_p}\right]\right)$$
$$= \frac{1}{\sqrt{\rho_n(n-p+1)}} \sum_{i=1}^{n} Z_{pi},$$

where for $i \in \{1, \dots, n\}$, $n \in \mathbb{N}$

$$Z_{pi} = 2b_{pi} \sum_{k=1}^{i-1} p_{ki} b_{pk} + p_{ii} \left(b_{pi}^2 - \mathbb{E}[b_{pi}^2]\right),$$
$$\rho_n = 2 + \frac{\nu_4 - 3}{n-p+1} \sum_{i=1}^{n} p_{ii}^2. \tag{9.9}$$

For $i \in \{1, \dots, n\}$, let $\mathbb{E}_i$ denote the conditional expectation with respect to the σ-field $\mathcal{F}_{pi}$ generated by $\{\mathbf{b}_1, \dots, \mathbf{b}_{p-1}\} \cup \{b_{pk} : 1 \le k \le i\}$. Furthermore, $\mathbb{E}_0[X] = \mathbb{E}[X]$ denotes the usual expectation.
Since b_{pk} is measurable with respect to $\mathcal{F}_{p,i-1}$ for $k \in \{1, \dots, i-1\}$ and b_{pj} is independent of $\mathcal{F}_{p,i-1}$ for $j \in \{i, \dots, n\}$, and $\mathbf{P}$ is measurable with respect to $\mathcal{F}_{pi}$ for all $i \in \{1, \dots, n\}$, we obtain

$$\mathbb{E}_{i-1}[Z_{pi}] = 2\sum_{k=1}^{i-1} \mathbb{E}_{i-1}[b_{pi} p_{ki}] b_{pk} + \mathbb{E}_{i-1}\left[p_{ii}\left(b_{pi}^2 - \mathbb{E}[b_{pi}^2]\right)\right]$$
$$= 2\mathbb{E}[b_{pi}] \sum_{k=1}^{i-1} p_{ki} b_{pk} + p_{ii}\left(\mathbb{E}[b_{pi}^2] - \mathbb{E}[b_{pi}^2]\right) = 0, \ 2 \le i \le n.$$

Note that Z_{pi} is measurable with respect to $\mathcal{F}_{pi}$ $(1 \le i \le n)$. These observations imply that for each $n \in \mathbb{N}$, $(Z_{pi})_{1\le i\le n}$ forms a martingale difference sequence with respect to the filtration $(\mathcal{F}_{pi})_{1\le i\le n}$. This representation of a random quadratic form as a martingale difference scheme generalizes the one of Bhansali et al. (2007). Note that we are not able to apply their Theorem 2.1 directly in order to prove asymptotic normality, since in our case $\mathbf{P}$ is a random matrix and the random vectors $\mathbf{b}_p$ vary with $n \in \mathbb{N}$. Thus, we have to give a direct proof showing that it satisfies the conditions of the central limit theorem for martingale difference sequences provided in Lemma 9.5.3 in Section 9.5.1. More precisely, we will show that for all $\delta > 0$

$$\sigma_n^2 = \frac{1}{\rho_n(n-p+1)} \sum_{i=1}^{n} \mathbb{E}_{i-1}[Z_{pi}^2] \xrightarrow{\mathbb{P}} 1, \tag{9.10}$$

$$r_n(\delta) = \frac{1}{\rho_n(n-p+1)} \sum_{i=1}^{n} \mathbb{E}\left[Z_{pi}^2 I_{\{|Z_{pi}| \ge \delta\sqrt{(n-p+1)\rho_n}\}}\right] \to 0, \tag{9.11}$$

as $n \to \infty$.
As a preparation for the following steps, we note that

$$\max_{l=1,\dots,n} \sum_{m=1}^{n} p_{lm}^2 \le ||\mathbf{P}||^2 \le 1, \tag{9.12}$$

$$\operatorname{tr}\left(\mathbf{P}^2\right) = \sum_{i,k=1}^{n} p_{ki}p_{ik} = ||\mathbf{P}||_2^2 = \operatorname{tr}\mathbf{P} = n-p+1, \tag{9.13}$$

where $||\mathbf{P}||$ denotes the spectral norm of $\mathbf{P}$ and $||\mathbf{P}||_2$ denotes the Euclidean norm of $\mathbf{P}$. The first inequality in (9.12) is a well-known estimate for general symmetric matrices and can be shown by choosing the unit vectors for the maximum appearing in the definition of the spectral norm, while the equality in (9.13) follows from the fact that $\mathbf{P}^2 = \mathbf{P}$.

Step 2.1: Calculation of the variance

We begin with a proof of (9.10). For this purpose, we calculate

$$\begin{aligned}
\sigma_n^2 =& \frac{1}{\rho_n(n-p+1)} \sum_{i=1}^{n} \mathbb{E}_{i-1}\left[Z_{pi}^2\right] \\
=& \frac{4}{\rho_n(n-p+1)} \sum_{i=1}^{n} \mathbb{E}_{i-1}\left[\left(\sum_{k=1}^{i-1} p_{ki}b_{pk}\right)^2\right] \\
&+ \frac{4}{\rho_n(n-p+1)} \sum_{i=1}^{n} \left\{\left(\mathbb{E}\left[b_{pi}^3\right] - \mathbb{E}[b_{pi}]\mathbb{E}\left[b_{pi}^2\right]\right) \sum_{k=1}^{i-1} b_{pk}\mathbb{E}_{i-1}[p_{ki}p_{ii}]\right\} \\
&+ \frac{1}{\rho_n(n-p+1)} \sum_{i=1}^{n} \mathbb{E}_{i-1}[p_{ii}^2]\mathbb{E}\left[b_{pi}^2 - \mathbb{E}[b_{pi}^2]\right]^2 \\
=& \frac{4}{\rho_n(n-p+1)} \sum_{i=1}^{n} \mathbb{E}_{i-1}\left[\left(\sum_{k=1}^{i-1} p_{ki}b_{pk}\right)^2\right] + \frac{4\mathbb{E}\left[b_{p1}^3\right]}{\rho_n(n-p+1)} \sum_{i=1}^{n} \left\{\sum_{k=1}^{i-1} b_{pk}p_{ki}p_{ii}\right\} \\
&+ \frac{(\nu_4 - 1)}{\rho_n(n-p+1)} \sum_{i=1}^{n} p_{ii}^2.
\end{aligned} \tag{9.14}$$

Here, we used that b_{pk} is measurable with respect to $\mathcal{F}_{pi}$ for $k \in \{1,\dots,i\}$ and b_{pj} is independent of $\mathcal{F}_{pi}$ for $j \in \{i+1,\dots,n\}$, and $\mathbf{P}$ is measurable with respect to $\mathcal{F}_{pi}$ for all $i \in \{1,\dots,n\}$. Moreover, we obtain using (9.13)

$$\begin{aligned}
1 =& \rho_n^{-1}\left(2 + \frac{\nu_4 - 3}{n-p+1} \sum_{i=1}^{n} p_{ii}^2\right) \\
=& \frac{2}{\rho_n(n-p+1)} \sum_{\substack{i,k=1,\\ i\neq k}}^{n} p_{ki}^2 + \frac{\nu_4 - 1}{\rho_n(n-p+1)} \sum_{i=1}^{n} p_{ii}^2
\end{aligned}$$

$$= \frac{4}{\rho_n(n-p+1)} \sum_{i=1}^{n} \sum_{k=1}^{i-1} p_{ki}^2 + \frac{\nu_4 - 1}{\rho_n(n-p+1)} \sum_{i=1}^{n} p_{ii}^2. \tag{9.15}$$

Denoting $\nu_4 = 1 + \varepsilon$ for some small $\varepsilon > 0$, we note that ρ_n is uniformly bounded away from 0, since for all $n \in \mathbb{N}$

$$\rho_n = 2 - \frac{2-\varepsilon}{n-p+1} \sum_{i=1}^{n} p_{ii}^2 \geq 2 - \frac{2-\varepsilon}{n-p+1} \sum_{i=1}^{n} p_{ii} = \varepsilon > 0. \tag{9.16}$$

In the following, we will show that (9.10) holds true with $\sigma^2 = 1$. For this purpose, we write using (9.14), (9.15) and (9.16)

$$\begin{aligned} |\sigma_n^2 - 1| \leq & \frac{4}{\rho_n(n-p+1)} \left| \sum_{i=1}^{n} \left(\mathbb{E}_{i-1} \left[\sum_{k=1}^{i-1} p_{ki} b_{pk} \right]^2 - \sum_{k=1}^{i-1} p_{ki}^2 \right) \right| \\ & + \frac{4\mathbb{E}|b_{p1}|^3}{\rho_n(n-p+1)} \left| \sum_{i=1}^{n} \left\{ \sum_{k=1}^{i-1} b_{pk} p_{ki} p_{ii} \right\} \right| \\ \lesssim & \frac{1}{n-p+1} \left(\delta_{n,1} + \delta_{n,2} + \delta_{n,3} \right), \end{aligned} \tag{9.17}$$

where

$$\begin{aligned} \delta_{n,1} &= \left| \sum_{i=1}^{n} \sum_{1 \leq k < j \leq i-1} p_{ki} p_{ji} b_{pk} b_{pj} \right|, \\ \delta_{n,2} &= \left| \sum_{i=1}^{n} \sum_{k=1}^{i-1} \left(b_{pk}^2 - 1 \right) p_{ki}^2 \right|, \\ \delta_{n,3} &= \left| \sum_{i=1}^{n} \sum_{k=1}^{i-1} b_{pk} p_{ki} p_{ii} \right|. \end{aligned}$$

Similarly as in Bhansali et al. (2007), one can show that $\delta_{n,i}/(n-p+1) = o_{\mathbb{P}}(1)$, as $n \to \infty$ for $i \in \{1, 2, 3\}$, by bounding the second moments of $\delta_{n,1}, \delta_{n,2}, \delta_{n,3}$. Exemplarily, we demonstrate this for the term $\delta_{n,3}$. Notice that an application of Lemma 2.1 in Bhansali et al. (2007) and (9.13) yields

$$\left(\sum_{i,i'=1}^{n} \left(\sum_{k=1}^{\min(i,i')-1} p_{ik} p_{i'k} \right)^2 \right)^{\frac{1}{2}} \lesssim \sqrt{n-p+1} ||\mathbf{P}|| \leq \sqrt{n-p+1}. \tag{9.18}$$

Using the Cauchy-Schwarz inequality, (9.18) and (9.13),

$$\mathbb{E}[\delta_{n,3}^2] = \mathbb{E} \left[\sum_{i,i'=1}^{n} p_{ii} p_{i'i'} \sum_{k=1}^{\min(i,i')-1} p_{ki} p_{ki'} \right]$$

$$
\begin{aligned}
&\leq \mathbb{E}\left[\left(\sum_{i}^{n} p_{ii}^2\right)\left(\sum_{i,i'=1}^{n}\left(\sum_{k=1}^{\min(i,i')-1} p_{ki}p_{ki'}\right)^2\right)^{\frac{1}{2}}\right] \\
&\lesssim (n-p+1)^{\frac{3}{2}} = o\left((n-p+1)^2\right), \ n \to \infty.
\end{aligned}
$$

Proceeding similarly for the remaining terms $\delta_{n,1}$ and $\delta_{n,2}$, we get $\sigma_n^2 = 1 + o_{\mathbb{P}}(1)$ as $n \to \infty$. By an application of Lemma 9.5.4 given at the end of this section, the normalizing term ρ_n converges in probability towards ρ as $n \to \infty$.

Step 2.2: Verifying the Lindeberg-type condition (9.11)

Using a truncation argument as in Bhansali et al. (2007), it is sufficient to prove (9.11) under the assumption $\mathbb{E}[b_{11}^8] < \infty$. Then, we obtain by using (9.16)

$$
r_n(\delta) \leq \frac{1}{(n-p+1)^2\rho_n^2\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[Z_{pi}^4\right] \lesssim J_1 + J_2,
$$

where

$$
\begin{aligned}
J_1 =& \frac{1}{(n-p+1)^2\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[b_{pi}^4\left(\sum_{k=1}^{i-1} p_{ki}b_{pk}\right)^4\right] \\
\lesssim& \frac{1}{(n-p+1)^2\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{j,k=1}^{i-1} p_{ki}p_{ji}b_{pk}b_{pj}\right)^2\right] \\
\lesssim& \frac{1}{(n-p+1)\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{k=1}^{i-1} p_{ki}^2 b_{pk}^2\right)^2\right] \\
&+ \frac{1}{(n-p+1)\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[\left(\sum_{\substack{j,k=1\\ j<k}}^{i-1} p_{ki}p_{ji}b_{pk}b_{pj}\right)^2\right], \\
J_2 =& \frac{1}{(n-p+1)^2\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[p_{ii}^4\left(b_{pi}^2 - \mathbb{E}[b_{pi}^2]\right)^4\right] \lesssim \frac{1}{(n-p+1)\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[p_{ii}^4\right].
\end{aligned}
$$

This implies using (9.12) and (9.13)

$$
\begin{aligned}
J_1 + J_2 \lesssim& \frac{1}{(n-p+1)^2\delta^2}\sum_{i=1}^{n}\left(\sum_{j,k=1}^{i-1}\mathbb{E}[p_{ki}^2p_{ji}^2] + \mathbb{E}[p_{ii}^4]\right) \lesssim \frac{1}{(n-p+1)^2\delta^2}\sum_{i,j,k=1}^{n}\mathbb{E}[p_{ki}^2p_{ji}^2] \\
\lesssim& \frac{1}{(n-p+1)^2\delta^2}\sum_{j,k=1}^{n}\mathbb{E}\left[p_{jk}^2\max_{l=1,\dots,n}\sum_{m=1}^{n} p_{lm}^2\right] \\
\lesssim& \frac{1}{(n-p+1)^2\delta^2}\sum_{j,k=1}^{n}\mathbb{E}[p_{jk}^2] = o(1).
\end{aligned}
$$

Step 3: Conclusion via delta method

In Step 2, we have shown that an appropriately centered and standardized version of r_{pp}^2 satisfies a CLT. By applying the delta method and using (9.6), we conclude that

$$\frac{\sqrt{n-p+1}}{\left(\mathbf{\Sigma}^{-1}\right)_{pp}}\left(\frac{n-p+1}{n}\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{pp}-\left(\mathbf{\Sigma}^{-1}\right)_{pp}\right)=\sqrt{n-p+1}\left(\frac{n-p+1}{r_{pp}^2}-1\right)$$
$$\xrightarrow{\mathcal{D}}\mathcal{N}(0,\rho),\ n\to\infty,$$

which finishes the proof of Theorem 9.3.1. □

Next, we prove the corresponding result for the subsampling procedure.

Proof of Theorem 9.3.4. The proof of Theorem 9.3.1 can be generalized to this subsampling strategy. In the following, we discuss only the main steps. Again, we may assume w.l.o.g. that $q=p$, since

$$\frac{\left(\breve{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}}=\left(\breve{\mathbf{I}}^{-1}\right)_{qq}\overset{\mathcal{D}}{=}\left(\breve{\mathbf{I}}^{-1}\right)_{pp}=\frac{\left(\breve{\mathbf{\Sigma}}^{-1}\right)_{pp}}{\left(\mathbf{\Sigma}^{-1}\right)_{pp}}.$$

Proceeding with a similar QR decomposition as in the proof of Theorem 9.3.1 (see also Section 9.7), we obtain

$$\left(\breve{\mathbf{I}}^{-1}\right)_{pp}^{-1}=\frac{1}{m}\breve{r}_{pp}^2,$$

where $\breve{r}_{qq}^2$ has the representation

$$\breve{r}_{pp}^2=\breve{\mathbf{b}}_p^\top\breve{\mathbf{P}}\breve{\mathbf{b}}_p.$$

We write

$$\sqrt{\frac{m-p+1}{\breve{\rho}_n}}\frac{1}{m-p+1}\left(\breve{r}_{pp}^2-(m-p+1)\right)$$
$$=\frac{1}{\sqrt{\breve{\rho}_n(m-p+1)}}\left(\breve{\mathbf{b}}_p^\top\breve{\mathbf{P}}\breve{\mathbf{b}}_p-\mathbb{E}\left[\breve{\mathbf{b}}_p^\top\breve{\mathbf{P}}\breve{\mathbf{b}}_p\right]\right)=\frac{1}{\sqrt{\breve{\rho}_n(m-p+1)}}\sum_{i=1}^m\breve{Z}_{pi},$$

where for $i\in\{1,\dots,m\}$, $n\in\mathbb{N}$

$$\breve{Z}_{pi}=2\breve{b}_{pi}\sum_{k=1}^{i-1}\breve{p}_{ki}\breve{b}_{pk}+\breve{p}_{ii}\left(\breve{b}_{pi}^2-\mathbb{E}[\breve{b}_{pi}^2]\right),$$
$$\breve{\rho}_n=2+\frac{\nu_4-3}{m-p+1}\sum_{i=1}^m\breve{p}_{ii}^2.$$

When defining $\breve{\mathcal{F}}_{pi}$ as the $\sigma-$field generated by $\{\mathbf{b}_1,\dots,\mathbf{b}_{p-1}\}\cup\{b_{pk}:1\le k\le i\}\cup\{U_n:n\in\mathbb{N}\}$, $1\le i\le n$, one can show that for each $n\in\mathbb{N}$, $(\breve{Z}_{pi})_{1\le i\le m}$ forms a martingale difference sequence with respect to the σ-fields $(\breve{\mathcal{F}}_{pi})_{1\le i\le m}$ defined above. Similarly as in the proof of Theorem 9.3.1, it is seen that

$$\begin{aligned}\breve{\sigma}_n^2 =& \frac{1}{\breve{\rho}_n(m-p+1)}\sum_{i=1}^m \mathbb{E}[\breve{Z}_{pi}^2|\breve{\mathcal{F}}_{pi}] \\ =& \frac{4}{\breve{\rho}_n(m-p+1)}\sum_{i=1}^m \mathbb{E}\left[\left(\sum_{k=1}^{i-1}\breve{p}_{ki}\breve{b}_{pk}\right)^2 \Big|\breve{\mathcal{F}}_{p,i-1}\right] \\ &+ \frac{4}{\breve{\rho}_n(m-p+1)}\sum_{i=1}^m \left\{\mathbb{E}\left[\breve{b}_{pi}^3\right]\sum_{k=1}^{i-1}\breve{b}_{pk}\breve{p}_{ki}\breve{p}_{ii}\right\} \\ &+ \frac{1}{\breve{\rho}_n(m-p+1)}\sum_{i=1}^m \breve{p}_{ii}^2\left(\mathbb{E}[\breve{b}_{pi}^4]-1\right)\end{aligned}$$

and

$$\begin{aligned}1 =& \breve{\rho}_n^{-1}\left(2+\frac{\nu_4-3}{m-p+1}\sum_{i=1}^m \breve{p}_{ii}^2\right) \\ =& \frac{2}{\breve{\rho}_n(m-p+1)}\sum_{\substack{i,k=1,\\ i\neq k}}^m \breve{p}_{ki}^2 + \frac{\nu_4-1}{\breve{\rho}_n(m-p+1)}\sum_{i=1}^m \breve{p}_{ii}^2 \\ =& \frac{4}{\breve{\rho}_n(m-p+1)}\sum_{i=1}^m\sum_{k=1}^{i-1} \breve{p}_{ki}^2 + \frac{\nu_4-1}{\breve{\rho}_n(m-p+1)}\sum_{i=1}^m \breve{p}_{ii}^2.\end{aligned}$$

Then, one can show in a similar fashion that $\breve{\sigma}_n^2 = 1 + o_{\mathbb{P}}(1)$, as $n\to\infty$, meaning that condition (9.22) in Lemma 9.5.3 holds true. Condition (9.23) is proven similarly as (9.11) in the proof of Theorem 9.3.1. Similarly to Lemma 9.5.4, we also get the convergence $\breve{\rho}_n \xrightarrow{\mathbb{P}} \breve{\rho}$ for $n\to\infty$, and thus, by applying Lemma 9.5.3, we conclude

$$\sqrt{m-p+1}\left(\frac{\breve{r}_{pp}^2}{m-p+1}-1\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\breve{\rho}), n\to\infty.$$

Similarly to Step 3 in the proof of Theorem 9.3.1, we conclude that

$$\frac{\sqrt{m-p+1}}{(\mathbf{\Sigma}^{-1})_{pp}}\left(\frac{m-p+1}{m}\left(\breve{\mathbf{\Sigma}}^{-1}\right)_{pp} - \left(\mathbf{\Sigma}^{-1}\right)_{pp}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\breve{\rho}), n\to\infty.$$

This finishes the proof of Theorem 9.3.4. □

9.5.1 Auxiliary results

As the following result reveals, the diagonal entries of the sample precision matrix for standardized data are closely connected to those for data with inhomogeneous variances.

Lemma 9.5.1 *For $1 \leq q \leq p$ and a diagonal matrix $\mathbf{\Sigma} \in \mathbb{R}^{p\times p}$, it holds*

$$\left(\hat{\mathbf{I}}^{-1}\right)_{qq} = \frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}}.$$

Proof of Lemma 9.5.1. Applying Cramer's rule and noting that $|\hat{\mathbf{\Sigma}}| = |\mathbf{\Sigma}||\hat{\mathbf{I}}|$, we get

$$\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}} = \frac{|\mathbf{\Sigma}|}{|\hat{\mathbf{\Sigma}}|}\frac{|\hat{\mathbf{\Sigma}}^{(-q)}|}{|\mathbf{\Sigma}^{(-q)}|} = \frac{1}{|\hat{\mathbf{I}}|}\frac{|\hat{\mathbf{\Sigma}}^{(-q)}|}{|\mathbf{\Sigma}^{(-q)}|}. \tag{9.19}$$

Let $(\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}$ denote the $(p-1)\times p$ submatrix of $\mathbf{\Sigma}^{1/2}$ where the qth row is deleted. Similarly, $(\mathbf{\Sigma}^{1/2})^{(\cdot,-q)}$ denotes the $p\times(p-1)$ submatrix of $\mathbf{\Sigma}^{1/2}$ where the qth column is deleted. Using these definitions, we see that

$$\hat{\mathbf{\Sigma}}^{(-q)} = (\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\mathbf{X}_n\mathbf{X}_n^\top(\mathbf{\Sigma}^{1/2})^{(\cdot,-q)} = (\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\mathbf{X}_n\left((\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\mathbf{X}_n\right)^\top. \tag{9.20}$$

(In order to enforce (9.20), $\mathbf{\Sigma}$ does not need to be a diagonal matrix.) Since $\mathbf{\Sigma}$ is a diagonal matrix, it holds

$$(\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\mathbf{X}_n = \left(\mathbf{\Sigma}^{(-q)}\right)^{1/2}\tilde{\mathbf{X}}_n^{(-q)},$$

which implies

$$|\hat{\mathbf{\Sigma}}^{(-q)}| = |\mathbf{\Sigma}^{(-q)}||\hat{\mathbf{I}}^{(-q)}|. \tag{9.21}$$

Using (9.19), (9.21) and Cramer's rule again, we obtain

$$\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}} = \frac{|\hat{\mathbf{I}}^{(-q)}|}{|\hat{\mathbf{I}}|} = \left(\hat{\mathbf{I}}^{-1}\right)_{qq}.$$

□

The connection given in Lemma 9.5.1 can be generalized to the case of dependent coordinates if we assume that the data follows a standard normal distribution.

Lemma 9.5.2 *If $\mathbf{\Sigma}$ is a general (not necessarily diagonal) $p\times p$ population covariance matrix and $x_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ $(1\leq i\leq p,\ 1\leq j\leq n)$, then for any $1\leq q\leq p$*

$$\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}} \overset{\mathcal{D}}{=} \left(\hat{\mathbf{I}}^{-1}\right)_{qq}.$$

Proof of Lemma 9.5.2. Recall formula (9.20) from the proof of Lemma 9.5.1. It follows from our normal assumption that

$$(\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\mathbf{X}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}^{(-q)}),$$

where we used that

$$(\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\left((\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\right)^{\top}=(\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}(\mathbf{\Sigma}^{1/2})^{(\cdot,-q)}=\mathbf{\Sigma}^{(-q)}.$$

This implies that

$$(\mathbf{\Sigma}^{1/2})^{(-q,\cdot)}\mathbf{X}_n\overset{\mathcal{D}}{=}\tilde{\mathbf{Y}}_n^{(-q)}.$$

Using Cramers rule, we get

$$\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{qq}}{\left(\mathbf{\Sigma}^{-1}\right)_{qq}}=\frac{|\mathbf{\Sigma}|}{|\hat{\mathbf{\Sigma}}|}\frac{|\hat{\mathbf{\Sigma}}^{(-q)}|}{|\mathbf{\Sigma}^{(-q)}|}\overset{\mathcal{D}}{=}\frac{|\hat{\mathbf{I}}^{(-q)}|}{|\hat{\mathbf{I}}|}=\left(\hat{\mathbf{I}}^{-1}\right)_{qq}.$$

The proof of Lemma 9.5.2 concludes. □

In order to prove asymptotic normality of the quadratic forms appearing in the previous proofs, we make use of the following CLT for martingale difference schemes.

Lemma 9.5.3 (Theorem 35.12 in Billingsley (1995)) *Suppose that for each $n\in\mathbb{N}$, $Z_{n1},...,Z_{nr_n}$ form a real martingale difference sequence with respect to the increasing σ-field (F_{nj}) having second moments. If, as $n\to\infty$*

$$\sum_{j=1}^{r_n}\mathbb{E}[Z_{nj}^2|F_{n,j-1}]\overset{\mathbb{P}}{\to}\sigma^2, \tag{9.22}$$

where $\sigma^2>0$, and for each $\varepsilon>0$,

$$\sum_{j=1}^{r_n}\mathbb{E}[Z_{nj}^2 I_{\{|Z_{nj}|>\varepsilon\}}]\to 0, \tag{9.23}$$

then

$$\sum_{j=1}^{r_n}Z_{nj}\overset{\mathcal{D}}{\to}\mathcal{N}(0,\sigma^2).$$

We conclude this section by proving the following lemma which was used in the proof of Theorem 9.3.1 and provides the limiting variance.

Lemma 9.5.4 *It holds*

$$\rho_n\overset{\mathbb{P}}{\to}\rho,\ n\to\infty,$$

where ρ is defined in Theorem 9.3.1 and ρ_n in (9.9).

Proof of Lemma 9.5.4. Assume that $y=0$. For this case, we note that

$$\frac{1}{n}\sum_{i=1}^{n}p_{ii}^2=\frac{1}{n}\sum_{i=1}^{n}(1-p_{ii})^2-1+\frac{2}{n}\sum_{i=1}^{n}p_{ii}=\frac{2(n-p+1)}{n}-1+o_{\mathbb{P}}(1)$$

$$= 1 + o_{\mathbb{P}}(1), \ n \to \infty, \tag{9.24}$$

where we used

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(1-p_{ii})^2 \le \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[1-p_{ii}] = \frac{1}{n}\operatorname{tr}(\mathbf{I}-\mathbf{P}) = \frac{p-1}{n} = o(1), \ n \to \infty.$$

Then, (9.24) implies

$$\rho_n = 2 + \frac{(\nu_4 - 3)n}{n-p+1} + o_{\mathbb{P}}(1) = \nu_4 - 1 = \rho.$$

Let $y \in (0,1)$. Then we have from Theorem 3.2 in Anatolyev and Yaskov (2017)

$$\frac{1}{n}\sum_{i=1}^{n}(1-p_{ii}-y)^2 \xrightarrow{\mathbb{P}} 0, \ n \to \infty,$$

which implies

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} p_{ii}^2 &= \frac{1}{n}\sum_{i=1}^{n}(1-p_{ii}-y)^2 - (1-y)^2 + \frac{2(1-y)}{n}\sum_{i=1}^{n} p_{ii} \\ &= \frac{2(1-y)(n-p+1)}{n} - (1-y)^2 + o_{\mathbb{P}}(1) = (1-y)^2 + o_{\mathbb{P}}(1), \ n \to \infty.\end{aligned}$$

We conclude for $n \to \infty$

$$\rho_n = 2 + \frac{(\nu_4-3)(1-y)^2 n}{n-p+1} + o_{\mathbb{P}}(1) = 2 + (\nu_4 - 3)(1-y) + o_{\mathbb{P}}(1) = \rho + o_{\mathbb{P}}(1).$$

□

9.6 Proofs of results in Section 9.4

9.6.1 Auxiliary results

The following lemma gives a concrete representation for any diagonal element of the sample precision matrix in terms of the entries of the triangular matrix $\mathbf{R}$.

Lemma 9.6.1 *For $1 \le q \le p$, it holds*

$$n\left(\hat{\mathbf{I}}^{-1}\right)^{-1}_{qq} = r_{qq}^2 \prod_{i=q+1}^{p} \frac{r_{ii}^2}{r_{ii,q}^2},$$

where the matrix $\mathbf{R}$ is defined in the proof of Theorem 9.3.1 and

$$r_{ii,q}^2 = \mathbf{b}_i^\top \mathbf{P}(i-1,q)\mathbf{b}_i, \ 1 \le i \ne q \le p.$$

Here, $\mathbf{P}(i-1,q)$ denotes the projection matrix on the orthogonal complement of span($\{\mathbf{b}_1, \dots, \mathbf{b}_{i-1}\} \setminus \{\mathbf{b}_q\}$). In particular, if $q = p-1$, we obtain

$$n\left(\hat{\mathbf{I}}^{-1}\right)^{-1}_{p-1,p-1} = \frac{r_{pp}^2 r_{p-1,p-1}^2}{\mathbf{b}_p^\top \mathbf{P}(p-2)\mathbf{b}_p}.$$

Proof of Lemma 9.6.1. Recall the QR-decomposition of $\mathbf{X}_n^\top$ given in Section 9.7 and the resulting formula

$$|\mathbf{X}_n\mathbf{X}_n^\top| = \prod_{i=1}^{p} r_{ii}^2.$$

Note that the first $(q-1)$ step in the QR-decomposition of the matrices $\tilde{\mathbf{X}}_n^\top = (\tilde{\mathbf{X}}_n^{(-q)})^\top$ and $\mathbf{X}_n^\top$ coincide, which implies

$$|\tilde{\mathbf{X}}_n\tilde{\mathbf{X}}_n^\top| = \prod_{i=1}^{q-1} r_{ii}^2 \prod_{i=q+1}^{p} r_{ii,q}^2.$$

Combining these formulas with Cramer's rule, we conclude

$$n\left(\hat{\mathbf{I}}^{-1}\right)_{qq}^{-1} = \frac{|\mathbf{X}_n\mathbf{X}_n^\top|}{|\tilde{\mathbf{X}}_n\tilde{\mathbf{X}}_n^\top|} = r_{qq}^2 \prod_{i=q+1}^{p} \frac{r_{ii}^2}{r_{ii,q}^2}.$$

□

Recall from the proof of Theorem 9.3.1 (or see Section 9.7 for more details) that

$$\left(\hat{\mathbf{I}}^{-1}\right)_{pp}^{-1} = \frac{1}{n} r_{pp}^2 = \frac{1}{n}\mathbf{b}_p^\top \mathbf{P}(p-1)\mathbf{b}_p,$$

while it follows from the fact the entries x_{ij} of the matrix $\mathbf{X}_n$ are i.i.d. random variables that

$$\left(\hat{\mathbf{I}}^{-1}\right)_{qq}^{-1} \stackrel{\mathcal{D}}{=} \left(\hat{\mathbf{I}}^{-1}\right)_{pp}^{-1}, \; 1 \le q \le p,$$

These quantities can also be written as a quadratic form, but its concrete structure is unknown so far. The next lemma provides such a representation and specifies the dependency structure between two diagonal elements. For convenience, we restrict ourselves to the case $q = p-1$.

Lemma 9.6.2 *It holds*

$$n\left(\hat{\mathbf{I}}^{-1}\right)_{p-1,p-1}^{-1} = \mathbf{b}_{p-1}^\top \left(\mathbf{P}(p-2) - \mathbf{Q}(p)\right)\mathbf{b}_{p-1},$$

where $\mathbf{P}(p-2) - \mathbf{Q}(p)$ is a projection matrix of rank $n-p+1$ and independent of $\mathbf{b}_{p-1}$. More precisely, $\mathbf{Q}(p)$ denotes the matrix corresponding to the projection to $\mathbf{P}(p-2)\mathbf{b}_p$, that is,

$$\mathbf{Q}(p) = \frac{\mathbf{P}(p-2)\mathbf{b}_p\mathbf{b}_p^\top\mathbf{P}(p-2)}{\mathbf{b}_p^\top\mathbf{P}(p-2)\mathbf{b}_p}.$$

Proof of Lemma 9.6.2. Recall from Lemma 9.6.1 that

$$n\left(\hat{\mathbf{I}}^{-1}\right)^{-1}_{p-1,p-1} = \frac{r_{pp}^2 r_{p-1,p-1}^2}{\mathbf{b}_p^\top \mathbf{P}(p-2)\mathbf{b}_p}.$$

Note that $\mathbf{P}(p-1)\mathbf{b}_p = \mathbf{P}(p-2)\mathbf{b}_p - \mathrm{proj}_{e_{p-1}}(\mathbf{b}_p)$, where the projection of a vector $\mathbf{a} \in \mathbb{R}^n$ to a vector $\mathbf{e} \in \mathbb{R}^n$ is given by

$$\mathrm{proj}_{\mathbf{e}}(\mathbf{a}) = \frac{(\mathbf{e}, \mathbf{a})}{(\mathbf{e}, \mathbf{e})}\mathbf{e}$$

and (for details, see Section 9.7)

$$\mathbf{u}_{p-1} = \mathbf{P}(p-2)\mathbf{b}_{p-1}, \; \mathbf{e}_{p-1} = \frac{\mathbf{u}_{p-1}}{||\mathbf{u}_{p-1}||_2}.$$

Thus, we obtain

$$\begin{aligned} n\left(\hat{\mathbf{I}}^{-1}\right)^{-1}_{p-1,p-1} =& \mathbf{b}_{p-1}^\top \mathbf{P}(p-2)\mathbf{b}_{p-1}\left(1 - \frac{\mathbf{b}_p^\top \,\mathrm{proj}_{\mathbf{e}_{p-1}}(\mathbf{b}_p)}{\mathbf{b}_p^\top \mathbf{P}(p-2)\mathbf{b}_p}\right) \\ =& \mathbf{b}_{p-1}^\top \mathbf{P}(p-2)\mathbf{b}_{p-1}\left(1 - \mathbf{b}_p^\top \frac{(\mathbf{u}_{p-1}, \mathbf{b}_p)}{(\mathbf{u}_{p-1}, \mathbf{u}_{p-1})\mathbf{b}_p^\top \mathbf{P}(p-2)\mathbf{b}_p}\mathbf{u}_{p-1}\right) \\ =& \mathbf{b}_{p-1}^\top \mathbf{P}(p-2)\mathbf{b}_{p-1} - \frac{\mathbf{b}_p^\top (\mathbf{u}_{p-1}, \mathbf{b}_p)\mathbf{u}_{p-1}}{\mathbf{b}_p^\top \mathbf{P}(p-2)\mathbf{b}_p} \\ =& \mathbf{b}_{p-1}^\top \mathbf{P}(p-2)\mathbf{b}_{p-1} - \mathbf{b}_{p-1}^\top \mathbf{Q}(p)\mathbf{b}_{p-1}. \end{aligned}$$

Note that $\mathbf{Q}(p)^2 = \mathbf{Q}(p)$ and $\mathbf{P}(p-2)\mathbf{Q}(p) = \mathbf{Q}(p)\mathbf{P}(p-2) = \mathbf{Q}(p)$. Consequently, we obtain

$$\begin{aligned} \left(\mathbf{P}(p-2) - \mathbf{Q}(p)\right)^2 =& \mathbf{P}(p-2)^2 + \mathbf{Q}(p)^2 - \mathbf{P}(p-2)\mathbf{Q}(p) - \mathbf{Q}(p)\mathbf{P}(p-2) \\ =& \mathbf{P}(p-2) + \mathbf{Q}(p) - 2\mathbf{Q}(p) = \mathbf{P}(p-2) - \mathbf{Q}(p). \end{aligned}$$

This implies that $\mathbf{P}(p-2) - \mathbf{Q}(p)$ is a projection matrix independent of $\mathbf{b}_{p-1}$ of rank

$$\mathrm{tr}\left(\mathbf{P}(p-2) - \mathbf{Q}(p)\right) = n - p + 2 - 1 = n - p + 1.$$

□

Lemma 9.6.2 helps us to understand the dependence structure between two diagonal entries and thus, sets us in the position to prove Theorem 9.4.1, which is done in the following section.

9.6.2 Proof of Theorem 9.4.1

Since the distribution of $\hat{\mathbf{I}}^{-1}$ is invariant under interchanging rows of $\mathbf{X}_n$, we have using Lemma 9.5.1

$$\left(\frac{\left(\hat{\boldsymbol{\Sigma}}^{-1}\right)_{q_1,q_1}}{\left(\boldsymbol{\Sigma}^{-1}\right)_{q_1,q_1}}, \frac{\left(\hat{\boldsymbol{\Sigma}}^{-1}\right)_{q_2,q_2}}{\left(\boldsymbol{\Sigma}^{-1}\right)_{q_2,q_2}}\right) = \left(\left(\hat{\mathbf{I}}^{-1}\right)_{q_1,q_1}, \left(\hat{\mathbf{I}}^{-1}\right)_{q_2,q_2}\right) \stackrel{\mathcal{D}}{=} \left(\left(\hat{\mathbf{I}}^{-1}\right)_{p-1,p-1}, \left(\hat{\mathbf{I}}^{-1}\right)_{pp}\right)$$

$$= \left(\left(\hat{\mathbf{I}}^{-1}\right)_{p-1,p-1}, \left(\hat{\mathbf{I}}^{-1}\right)_{pp} \right).$$

Thus, we may assume $q_1 = p-1$ and $q = p$ without loss of generality. Similarly as in the proof of Theorem 9.3.1, we start by investigating the asymptotic properties of

$$\begin{aligned} W_n =& \left\{ \frac{1}{\sqrt{n-p+1}} \left(n \left(\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{pp}}{\left(\mathbf{\Sigma}^{-1}\right)_{pp}} \right)^{-1} - (n-p+1) \right), \right. \\ & \left. \frac{1}{\sqrt{n-p+1}} \left(n \left(\frac{\left(\hat{\mathbf{\Sigma}}^{-1}\right)_{p-1,p-1}}{\left(\mathbf{\Sigma}^{-1}\right)_{p-1,p-1}} \right)^{-1} - (n-p+1) \right) \right\}^\top \\ =& \left\{ \frac{1}{\sqrt{n-p+1}} \left(n \left(\hat{\mathbf{I}}^{-1}\right)_{pp}^{-1} - (n-p+1) \right), \right. \\ & \left. \frac{1}{\sqrt{n-p+1}} \left(n \left(\hat{\mathbf{I}}^{-1}\right)_{p-1,p-1}^{-1} - (n-p+1) \right) \right\}^\top \\ =& \frac{1}{\sqrt{n-p+1}} \left\{ \mathbf{b}_p^\top \mathbf{P}(p-1) \mathbf{b}_p - (n-p+1), \right. \\ & \left. \mathbf{b}_{p-1}^\top (\mathbf{P}(p-2) - \mathbf{Q}(p)) \mathbf{b}_{p-1} - (n-p+1) \right\}^\top, \end{aligned}$$

where we used Lemma 9.5.1 and Lemma 9.6.2. From now on, the proof is divided in several steps.

Approximation and MDS

Note that for any rank-one projection matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ independent of $\mathbf{b}_p$, we have

$$\operatorname{Var}(\mathbf{b}_p^\top \mathbf{Q} \mathbf{b}_p) \lesssim 1 \ \forall n \in \mathbb{N},$$

and consequently, by Slutsky's lemma, it is sufficient to investigate

$$\begin{aligned} W_n^{(1)} &= \frac{1}{\sqrt{n-p+1}} \left\{ \mathbf{b}_p^\top \mathbf{P}(p-2) \mathbf{b}_p - (n-p+2), \mathbf{b}_{p-1}^\top \mathbf{P}(p-2) \mathbf{b}_{p-1} - (n-p+2) \right\} \\ &= W_n + o_{\mathbb{P}}(1). \end{aligned}$$

Throughout the rest of this proof, we denote $\mathbf{P}(p-2) = \mathbf{P} = (p_{ij})_{1 \leq i,j \leq n}$. By an application of the Cramer-Wold device, we note that it is sufficient to prove a one-dimensional central limit theorem for

$$W_n^{(2)} = \frac{1}{\sqrt{n-p+1}} \left\{ a \left(\mathbf{b}_p^\top \mathbf{P} \mathbf{b}_p - (n-p+2) \right) + b \left(\mathbf{b}_{p-1}^\top \mathbf{P} \mathbf{b}_{p-1} - (n-p+2) \right) \right\}, \ a, b \in \mathbb{R},$$

in order to ensure that the vector $W_n^{(1)}$ converges to a two-dimensional normal distribution. We write

$$\frac{1}{\sqrt{\rho_n}}W_n^{(2)} = \frac{1}{\sqrt{(n-p+1)\rho_n}}\sum_{i=1}^{n} W_{pi},$$

where

$$\begin{aligned} W_{pi} =& a\left(2b_{pi}\sum_{k=1}^{i-1} p_{ki}b_{pk} + p_{ii}\left(b_{pi}^2-1\right)\right) + b\left(2b_{p-1,i}\sum_{k=1}^{i-1} p_{ki}b_{p-1,k} + p_{ii}\left(b_{p-1,i}^2-1\right)\right), \\ \rho_n =& 2 + \frac{\nu_4-3}{n-p+1}\sum_{i=1}^{n} p_{ii}^2. \end{aligned}$$

For $p \in \mathbb{N}$, $1 \le i \le n$, let $\mathcal{A}_{pi}$ denote the σ field generated by $\{\mathbf{b}_1, \ldots, \mathbf{b}_{p-2}\} \cup \{b_{pk}, b_{p-1,k} : 1 \le k \le i\}$. Similarly as in the proof of Theorem 9.3.1, one can show that $(W_{pi})_{1\le i\le n}$ forms a martingale difference sequence with respect to the σ-fields $(\mathcal{A}_{pi})_{1\le i\le n}$ for each $p \in \mathbb{N}$. In order to apply the central limit theorem given in Lemma 9.5.3, we need to verify the conditions (9.22) and (9.23).

Calculation of the variance

We begin with a proof of condition (9.22). Note that

$$\begin{aligned} &\frac{1}{\rho_n(n-p+1)}\sum_{i=1}^{n}\mathbb{E}[W_{pi}^2|\mathcal{A}_{i-1}] \\ =& \frac{a^2}{\rho_n(n-p+1)}\mathbb{E}\left[\left(2b_{pi}\sum_{k=1}^{i-1} p_{ki}b_{pk} + p_{ii}\left(b_{pi}^2-1\right)\right)^2\middle|\mathcal{A}_{i-1}\right] \\ &+ \frac{b^2}{\rho_n(n-p+1)}\mathbb{E}\left[\left(2b_{p-1,i}\sum_{k=1}^{i-1} p_{ki}b_{p-1,k} + p_{ii}\left(b_{p-1,i}^2-1\right)\right)^2\middle|\mathcal{A}_{i-1}\right] \\ &+ \frac{2ab}{\rho_n(n-p+1)}\mathbb{E}\Bigg[\left(2b_{p-1,i}\sum_{k=1}^{i-1} p_{ki}b_{p-1,k} + p_{ii}\left(b_{p-1,i}^2-1\right)\right) \\ &\times\left(2b_{pi}\sum_{k=1}^{i-1} p_{ki}b_{pk} + p_{ii}\left(b_{pi}^2-1\right)\right)\middle|\mathcal{A}_{i-1}\Bigg] \\ =& \frac{a^2}{\rho_n(n-p+1)}\mathbb{E}\left[\left(2b_{pi}\sum_{k=1}^{i-1} p_{ki}b_{pk} + p_{ii}\left(b_{pi}^2-1\right)\right)^2\middle|\mathcal{A}_{i-1}\right] \\ &+ \frac{b^2}{\rho_n(n-p+1)}\mathbb{E}\left[\left(2b_{p-1,i}\sum_{k=1}^{i-1} p_{ki}b_{p-1,k} + p_{ii}\left(b_{p-1,i}^2-1\right)\right)^2\middle|\mathcal{A}_{i-1}\right] \\ =& a^2+b^2+o_{\mathbb{P}}(1),\ n\to\infty, \end{aligned}$$

where we used (9.10) from the proof of Theorem 9.3.1.

Verification of the Lindeberg-type condition

For a proof of condition (9.23), we use the results from Step 2.2 in the proof of Theorem 9.3.1 and obtain

$$\frac{1}{\rho_n(n-p+1)}\sum_{i=1}^{n}\mathbb{E}\left[W_{pi}^2 I_{\{|W_{pi}|\geq\delta\sqrt{(n-p+1)\rho_n}\}}\right] \leq \frac{1}{(n-p+1)^2\rho_n^2\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[W_{pi}^4\right]$$
$$\lesssim \frac{a^4}{(n-p+1)^2\rho_n^2\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[\left(2b_{pi}\sum_{k=1}^{i-1}p_{ki}b_{pk}+p_{ii}\left(b_{pi}^2-1\right)\right)^4\right]$$
$$+\frac{b^4}{(n-p+1)^2\rho_n^2\delta^2}\sum_{i=1}^{n}\mathbb{E}\left[\left(2b_{p-1,i}\sum_{k=1}^{i-1}p_{ki}b_{p-1,k}+p_{ii}\left(b_{p-1,i}^2-1\right)\right)^4\right] = o(1),\ n\to\infty.$$

Conclusion via delta method

Summarizing the steps above, we obtain from Lemma 9.5.3

$$W_n \xrightarrow{\mathcal{D}} \mathcal{N}_2(\mathbf{0},\rho\mathbf{I}_2), n\to\infty.$$

By an application of the multivariate delta method, we have

$$\begin{aligned}(Z_{n,p},Z_{n,p-1})^\top =&\Bigg\{\sqrt{n-p+1}\left(\frac{n-p+1}{\mathbf{b}_p^\top\mathbf{P}(p-1)\mathbf{b}_p}-1\right),\\ &\sqrt{n-p+1}\left(\frac{n-p+1}{\mathbf{b}_{p-1}^\top(\mathbf{P}(p-2)-\mathbf{Q}(p))\mathbf{b}_{p-1}}-1\right)\Bigg\}^\top\\ &\xrightarrow{\mathcal{D}}\mathcal{N}_2(\mathbf{0},\rho\mathbf{I}_2), n\to\infty.\end{aligned}$$

9.7 Details on the QR-decomposition of $\mathbf{X}_n^\top$

In this section, we give more details on the QR-decomposition of the matrix $\mathbf{X}_n^\top$ (compare Section 2 in Wang et al., 2018) and provide an explicit representation of the diagonal elements of $\mathbf{R}$ as a quadratic form in the rows of $\mathbf{X}_n$.

To begin with, we describe the QR-decomposition of a general full-column rank matrix $\mathbf{A}=(\mathbf{a}_1,\ldots,\mathbf{a}_p)\in\mathbb{R}^{n\times p}$ by applying the Gram-Schmidt procedure to the vectors $\mathbf{a}_1,\ldots,\mathbf{a}_p$. Recall the definition of the projection of a vector $\mathbf{a}\in\mathbb{R}^n$ on a vector $\mathbf{e}\in\mathbb{R}^n,\mathbf{e}\neq\mathbf{0}$, is given by

$$\mathrm{proj}_{\mathbf{e}}(\mathbf{a})=\frac{(\mathbf{e},\mathbf{a})}{(\mathbf{e},\mathbf{e})}\mathbf{e}.$$

It holds

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{a}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \\
\mathbf{u}_2 &= \mathbf{a}_2 - \operatorname{proj}_{\mathbf{u}_1} \mathbf{a}_2, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \\
\mathbf{u}_3 &= \mathbf{a}_3 - \operatorname{proj}_{\mathbf{u}_1} \mathbf{a}_3 - \operatorname{proj}_{\mathbf{u}_2} \mathbf{a}_3, & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}, \\
&\vdots & &\vdots \\
\mathbf{u}_n &= \mathbf{a}_n - \sum_{j=1}^{n-1} \operatorname{proj}_{\mathbf{u}_j} \mathbf{a}_n, & \mathbf{e}_n &= \frac{\mathbf{u}_n}{\|\mathbf{u}_n\|}.
\end{aligned}$$

Rearranging these equations, we may write $\mathbf{A} = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q} = (\mathbf{e}_1, \dots, \mathbf{e}_p) \in \mathbb{R}^{n \times p}$ denotes a matrix with orthonormal columns satisfying $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$ and $\mathbf{R} \in \mathbb{R}^{p \times p}$ is an upper triangular matrix with entries $r_{ij} = (\mathbf{e}_i, \mathbf{a}_j)$ for $i \le j$ and $r_{ij} = 0$ for $i > j$, $i, j \in \{1, \dots, p\}$.

In order to ensure formal correctness of the QR decomposition for the matrix $\mathbf{X}_n^\top = (\mathbf{b}_1, \dots, \mathbf{b}_p)$, we note that the matrix $\mathbf{X}_n^\top$ has full column rank since we assumed that each x_{ij} follows a continuous distribution for $1 \le i \le p$, $1 \le j \le n$. Performing the QR decomposition for the special choice $\mathbf{A} = \mathbf{X}_n^\top = (\mathbf{b}_1, \dots, \mathbf{b}_p)$, we get

$$\mathbf{X}_n^\top = \mathbf{Q}\mathbf{R},$$

where $\mathbf{Q} = (\mathbf{e}_1, \dots, \mathbf{e}_p) \in \mathbb{R}^{n \times p}$ denotes a matrix with orthonormal columns satisfying $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$ and $\mathbf{R} \in \mathbb{R}^{p \times p}$ is an upper triangular matrix with entries $r_{ij} = (\mathbf{e}_i, \mathbf{b}_j)$ for $i \le j$ and $r_{ij} = 0$ for $i > j$, $i, j \in \{1, \dots, p\}$. Using the definitions $r_{qq}^2 = (\mathbf{e}_i, \mathbf{b}_i)^2$ for $1 \le q \le p$ and $\mathbf{P}(0) = \mathbf{I}$, we have

$$r_{11}^2 = (\mathbf{e}_1, \mathbf{b}_1)^2 = ||\mathbf{b}_1||_2^2 = \mathbf{b}_1^\top \mathbf{P}(0) \mathbf{b}_1,$$

and for $2 \le q \le p$

$$r_{qq}^2 = (\mathbf{e}_q, \mathbf{b}_q)^2 = \left(\frac{\mathbf{u}_q^\top \mathbf{b}_q}{||\mathbf{u}_q||_2} \right)^2 = \left(\frac{\mathbf{b}_q^\top \mathbf{P}(q-1) \mathbf{b}_q}{||\mathbf{P}(q-1)\mathbf{b}_q||_2} \right)^2 = \mathbf{b}_q^\top \mathbf{P}(q-1) \mathbf{b}_q, \tag{9.25}$$

where the projection matrix $\mathbf{P}(q-1)$ is defined in (9.8) and satisfies $\mathbf{P}(q-1)^2 = \mathbf{P}(q-1)$.

List of symbols

$\lvert \cdot \rvert$	determinant of a matrix or absolute value of a complex number
$\lvert\lvert \cdot \rvert\rvert$	spectral norm of a matrix
$\mathbf{0}$	vector filled with zeros of appropriate dimension
$a \lesssim b$	a is smaller than b up to a positive constant or $a = b$
$a \vee b$	maximum of a and b
$a \wedge b$	minimum of a and b
$\mathbf{B}_{n,t}$	sequential sample covariance matrix
$\mathcal{X}_d^2$	chi-squared distribution with d degrees of freedom
$\mathbb{C}$	set of complex numbers
$\mathbb{C}^+$	set of complex numbers with positive imaginary part
$\xrightarrow{\mathcal{D}}$	weak convergence (convergence in distribution) in $\mathbb{R}^d$
$\rightsquigarrow$	weak convergence in a metric space
$F^{\mathbf{A}}$	empirical spectral distribution of a quadratic matrix $\mathbf{A}$ with real eigenvalues
F^{y,σ^2}	Marčenko–Pastur law with index y and scale parameter σ^2
F^y	Marčenko–Pastur law with index y and scale parameter $\sigma^2 = 1$
$F^{y,H}$	generalized Marčenko–Pastur law for $y > 0$ and a distribution H
Γ_F	support of a c.d.f. F
H	limiting spectral distribution of $\mathbf{T}_n$
$\mathbf{I}$	identity matrix
I	indicator function
$\operatorname{Im}(\cdot)$	imaginary part of a complex number
$\lambda_1(\cdot)$	largest eigenvalue of a $(p \times p)$ Hermitian matrix
$\lambda_p(\cdot)$	smallest eigenvalue of a $(p \times p)$ Hermitian matrix

$\ell^\infty(I)$	space of bounded functions mapping from $I \in \{[0,1],[t_0,1]\}$ into $\mathbb{R}$ or $\mathbb{C}$
$\mathbb{N}$	set of all positive integers
n	sample size
$\mathcal{N}(\mu,\Sigma)$	multivariate normal distribution with mean vector μ and covariance matrix Σ
p	dimension
$\mathbf{P}$, $\mathbf{P}(\cdot)$	projection matrix
$\mathbb{R}$	set of real numbers
$\mathrm{Re}(\cdot)$	real part of a complex number
s_F	Stieltjes transform of a measure F
$\mathbf{T}_n, \mathbf{T}$	population covariance matrix
$(X(f,t))$	Gaussian process defined in Theorem 3.2.1
$\mathbf{X}_n$	$(p \times n)$ data matrix
y	limit of the dimension-to-sample-size ratio

Bibliography

Ahlfors, L. V. (1953). Complex analysis: an introduction to the theory of analytic functions of one complex variable. *New York, London*, 177.

Anatolyev, S. and Yaskov, P. (2017). Asymptotics of diagonal elements of projection matrices under many instruments/regressors. *Econometric Theory*, 33(3):717–738.

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, second edition.

Aue, A., Hörmann, S., Horváth, L., and Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics*, 37(6B):4046–4087.

Bai, Z., Jiang, D., Yao, J.-F., and Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Annals of Statistics*, 37:3822–3840.

Bai, Z. and Silverstein, J. W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *The Annals of Probability*, 26(1):316–345.

Bai, Z. and Silverstein, J. W. (1999). Exact separation of eigenvalues of large dimensional sample covariance matrices. *Annals of Probability*, pages 1536–1555.

Bai, Z. and Silverstein, J. W. (2004). Clt for linear spectral statistics of large-dimensional sample covariance matrices. *Annals of Probability*, 32(1):553–605.

Bai, Z. and Silverstein, J. W. (2010). *Spectral analysis of large dimensional random matrices*, volume 20. Springer.

Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. *Statistica Sinica*, pages 425–442.

Bai, Z. D. and Yin, Y. Q. (1988). Convergence to the semicircle law. *The Annals of Probability*, 16(2):863–875.

Bai, Z. D., Yin, Y. Q., and Krishnaiah, P. R. (1986). On limiting spectral distribution of product of two random matrices when the underlying distribution is isotropic. *Journal of multivariate analysis*, 19(1):189–200.

Banna, M., Najim, J., and Yao, J. (2020). A clt for linear spectral statistics of large random information-plus-noise matrices. *Stochastic Processes and their Applications*, 130(4):2250–2281.

Bao, Z., Lin, L.-C., Pan, G., and Zhou, W. (2015a). Spectral statistics of large dimensional Spearman's rank correlation matrix and its application. *The Annals of Statistics*, 43(6):2588 – 2623.

Bao, Z., Pan, G., Zhou, W., et al. (2015b). The logarithmic law of random determinant. *Bernoulli*, 21(3):1600–1628.

Bhansali, R., Giraitis, L., and Kokoszka, P. (2007). Convergence of quadratic forms with nonvanishing diagonal. *Statistics & probability letters*, 77(7):726–734.

Bickel, P. J. and Wichura, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *The Annals of Mathematical Statistics*, 42(5):1656–1670.

Billingsley, P. (1968). *Probability and Measure*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., first edition.

Billingsley, P. (1995). *Probability and measure*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., third edition.

Birke, M. and Dette, H. (2005). A note on testing the covariance matrix for large dimension. *Statistics & Probability Letters*, 74(3):281–289.

Bodnar, T., Dette, H., and Parolya, N. (2016). Spectral analysis of the moore–penrose inverse of a large dimensional sample covariance matrix. *Journal of Multivariate Analysis*, 148:160–172.

Bodnar, T., Dette, H., and Parolya, N. (2019). Testing for independence of large dimensional vectors. *The Annals of Statistics*, 47(5):2977 – 3008.

Cai, T. T., Liang, T., and Zhou, H. H. (2015). Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional gaussian distributions. *Journal of Multivariate Analysis*, 137:161–172.

Chang, J., Qiu, Y., Yao, Q., and Zou, T. (2018). Confidence regions for entries of a large precision matrix. *Journal of Econometrics*, 206(1):57–82.

Chen, B. and Pan, G. (2015). Clt for linear spectral statistics of normalized sample covariance matrices with the dimension much larger than the sample size. *Bernoulli*, 21(2):1089–1133.

Chen, S. X., Zhang, L.-X., and Zhong, P.-S. (2010). Testing high dimensional covariance matrices. *Journal of the American Statistical Association*, 105:810–819.

Cipolloni, G. and Erdős, L. (2018). Fluctuations for linear eigenvalue statistics of sample covariance random matrices. *arXiv preprint arXiv:1806.08751*.

D'Aristotile, A. (2000). An invariance principle for triangular arrays. *Journal of Theoretical Probability*, 13(2):327–341.

D'Aristotile, A., Diaconis, P., and Newman, C. M. (2003). Brownian motion and the classical groups. *IMS Lecture Notes-Monograph Series*, 41:97–116.

Dette, H. and Dörnemann, N. (2020). Likelihood ratio tests for many groups in high dimensions. *Journal of Multivariate Analysis*, 178:104605.

Dette, H. and Gösmann, J. (2020). A likelihood ratio approach to sequential change point detection for a general class of parameters. *Journal of the American Statistical Association*, 115(531):1361–1377.

Dette, H. and Tomecki, D. (2019). Determinants of block hankel matrices for random matrix-valued measures. *Stochastic Processes and their Applications*, 129(12):5200–5235.

Dörnemann, N. and Dette, H. (2021). Linear spectral statistics of sequential sample covariance matrices. *arXiv preprint arXiv:2107.10036*.

Dörnemann, N. (2022). Likelihood ratio tests under model misspecifications in high dimensions. *arXiv preprint arXiv:2203.05423*.

Fan, J. and Li, R. (2006). Statistical challenges with high dimensionality: feature selection in knowledge discovery. In Sanz-Solé, M., Soria, J., Varona, J. L., and Verdera, J., editors, *Proceedings of the International Congress of Mathematicians (p. 595-622)*. European Mathematical Society, Madrid, Spain.

Fisher, T. J., Sun, X., and Gallagher, C. M. (2010). A new test for sphericity of the covariance matrix for high dimensional data. *Journal of Multivariate Analysis*, 101(10):2554–2570.

Frisch, R. (1929). Correlation and scatter in statistical variables. *Nordic statistical journal*, 8:36–102.

Gao, J., Han, X., Pan, G., and Yang, Y. (2017). High dimensional correlation matrices: The central limit theorem and its applications. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3):677–693.

Girko, V. L. (1998). A refinement of the central limit theorem for random determinants. *Theory of Probability & Its Applications*, 42(1):121–129.

Goodman, N. (1963). The distribution of the determinant of a complex wishart distributed matrix. *The Annals of mathematical statistics*, 34(1):178–180.

Grenander, U. and Silverstein, J. W. (1977). Spectral analysis of networks with random topologies. *SIAM Journal on Applied Mathematics*, 32(2):499–519.

Guo, W. and Qi, Y. (2021). Asymptotic distributions for likelihood ratio tests for the equality of covariance matrices. *arXiv preprint arXiv:2110.02384*.

Gupta, A. K. and Nagar, D. K. (2018). *Matrix variate distributions*. Chapman and Hall/CRC.

Gupta, A. K. and Xu, J. (2006). On some tests of the covariance matrix under general conditions. *Annals of the Institute of Statistical Mathematics*, 58:101–114.

Heiny, J. and Parolya, N. (2021). Log determinant of large correlation matrices under infinite fourth moment. *arXiv preprint arXiv:2112.15388*.

Huang, D., Yao, Q., and Zhang, R. (2021). Krigings over space and time based on latent low-dimensional structures. *Science China Mathematics*, 64(4):823–848.

Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., Chen, K., Reiman, E., Initiative, A. D. N., et al. (2010). Learning brain connectivity of alzheimer's disease by sparse inverse covariance estimation. *NeuroImage*, 50(3):935–949.

Ji, H. C. and Lee, J. O. (2020). Gaussian fluctuations for linear spectral statistics of deformed wigner matrices. *Random Matrices: Theory and Applications*, 9(03):2050011.

Jiang, T. and Qi, Y. (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scandinavian Journal of Statistics*, 42(4):988–1009.

Jiang, T. and Yang, F. (2013). Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Annals of Statistics*, 41:2029–2074.

Jin, B., Wang, C., Bai, Z. D., Nair, K. K., and Harding, M. (2014). Limiting spectral distribution of a symmetrized auto-cross covariance matrix. *The Annals of Applied Probability*, 24(3):1199 – 1225.

John, S. (1971). Some optimal multivariate tests. *Biometrika*, 58(1):123–127.

Johnstone, I. (2006). High dimensional statistical inference and random matrices.

Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *Journal of Multivariate Analysis*, 12(1):1–38.

Lauritzen, S. L. (1996). *Graphical models*, volume 17. Clarendon Press.

Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Annals of statistics*, 30(4):1081–1102.

Li, Y. (2003). A martingale inequality and large deviations. *Statistics & Probability Letters*, 62(3):317–321.

Li, Z., Wang, Q., and Li, R. (2019). Central limit theorem for linear spectral statistics of large dimensional kendall's rank correlation matrices and its applications.

Li, Z., Wang, Q., and Li, R. (2021). Central limit theorem for linear spectral statistics of large dimensional kendall's rank correlation matrices and its applications. *The Annals of Statistics*, 49(3):1569–1593.

Lytova, A. and Pastur, L. (2009). Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices. *Metrika*, 69:153–172.

Ma, P., Mahoney, M., and Yu, B. (2014). A statistical perspective on algorithmic leveraging. In *International Conference on Machine Learning*, pages 91–99. PMLR.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457.

Mauchly, J. W. (1940). Significance test for sphericity of a normal N-variate distribution. *Annals of Mathematical Statistics*, 11:204–209.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* John Wiley & Sons, New York.

Muirhead, R. J. (2009). *Aspects of Multivariate Statistical Theory*, volume 197. John Wiley & Sons.

Nagel, J. (2020). A functional CLT for partial traces of random matrices. *Journal of Theoretical Probability*, 98.

Najim, J. and Yao, J. (2016). Gaussian fluctuations for linear spectral statistics of large random covariance matrices. *Annals of Applied Probability*, 26(3):1837–1887.

Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *The Annals of Mathematical Statistics*, 42(4):1285–1295.

Nguyen, H. H. and Vu, V. (2014). Random matrices: Law of the determinant. *The Annals of Probability*, 42(1):146–167.

Nydick, S. W. (2012). The wishart and inverse wishart distributions. *Electronic Journal of Statistics*, 6(1-19).

Pan, G. (2014). Comparison between two types of large sample covariance matrices. *Annales de l'IHP Probabilités et Statistiques*, 50(2):655–677.

Pan, G. and Zhou, W. (2008). Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *The Annals of Applied Probability*, 18(3):1232–1270.

Parolya, N., Heiny, J., and Kurowicka, D. (2021). Logarithmic law of large random correlation matrix. *arXiv preprint arXiv:2103.13900.*

Paul, D. and Aue, A. (2014). Random matrix theory in statistics: A review. *Journal of Statistical Planning and Inference*, 150:1–29.

Qi, Y., Wang, F., and Zhang, L. (2019). Limiting distributions of likelihood ratio test for independence of components for high-dimensional normal vectors. *Annals of the Institute of Statistical Mathematics*, 71(4):911–946.

Qiu, J., Li, Z., and Yao, J. (2021). Asymptotic normality for eigenvalue statistics of a general sample covariance matrix when $p/n \to \infty$ and applications. *arXiv preprint arXiv:2109.06701.*

Raskutti, G. and Mahoney, M. W. (2016). A statistical perspective on randomized sketching for ordinary least-squares. *The Journal of Machine Learning Research*, 17(1):7508–7538.

Shcherbina, M. (2011). Central limit theorem for linear eigenvalue statistics of the wigner and sample covariance random matrices. *Journal of Mathematical Physics, Analysis, Geometry*, 7(2):176–192.

Van de Geer, S., Bühlmann, P., Ritov, Y., and Dezeure, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *The Annals of Statistics*, 42(3):1166–1202.

Van Der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence*. Springer.

Von Rosen, D. (1988). Moments for the inverted wishart distribution. *Scandinavian Journal of Statistics*, pages 97–109.

Wachter, K. W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *The Annals of Probability*, pages 1–18.

Wachter, K. W. (1980). The limiting empirical measure of multiple discriminant ratios. *The Annals of Statistics*, pages 937–957.

Wang, H. and Ma, Y. (2021). Optimal subsampling for quantile regression in big data. *Biometrika*, 108(1):99–112.

Wang, H., Yang, M., and Stufken, J. (2019). Information-based optimal subdata selection for big data linear regression. *Journal of the American Statistical Association*, 114(525):393–405.

Wang, Q. and Yao, J. (2013). On the sphericity test with large-dimensional observations. *Electronic Journal of Statistics*, 7:2164–2192.

Wang, X., Han, X., and Pan, G. (2018). The logarithmic law of sample covariance matrices near singularity. *Bernoulli*, 24(1):80–114.

Wang, Z. and Yao, J. (2021). Central limit theorem for linear spectral statistics of block-wigner-type matrices. *arXiv preprint arXiv:2110.12171*.

Wigner, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, pages 325–327.

Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika*, pages 471–494.

Yao, J., Zheng, S., and Bai, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Yin, Y., Bai, Z., and Krishnaiah, P. (1983). Limiting behavior of the eigenvalues of a multivariate f matrix. *Journal of multivariate analysis*, 13(4):508–516.

Yin, Y. Q. (1986). Limiting spectral distribution for a class of random matrices. *Journal of multivariate analysis*, 20(1):50–68.

Yin, Y. Q. and Krishnaiah, P. R. (1983). A limit theorem for the eigenvalues of product of two random matrices. *Journal of Multivariate Analysis*, 13(4):489–507.

Zheng, S. (2012). Central limit theorems for linear spectral statistics of large dimensional f-matrices. *Annales de l'IHP Probabilités et Statistiques*, 48(2):444–476.

Zheng, S., Bai, Z., and Yao, J. (2015a). Clt for linear spectral statistics of a rescaled sample precision matrix. *Random Matrices: Theory and Applications*, 4(04):1550014.

Zheng, S., Bai, Z., and Yao, J. (2015b). Substitution principle for clt of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Annals of Statistics*, 43(2):546–591.

Zheng, S., Bai, Z., and Yao, J. (2017). CLT for eigenvalue statistics of large-dimensional general Fisher matrices with applications. *Bernoulli*, 23(2):1130 – 1178.

Zou, T., Zheng, S., Bai, Z., Yao, J., and Zhu, H. (2021). Clt for linear spectral statistics of large dimensional sample covariance matrices with dependent data. *Statistical Papers*, pages 1–60.