Linear spectral statistics of sequential sample covariance matrices

Nina Dörnemann, Holger Dette
Fakultät für Mathematik
Ruhr-Universität Bochum
44801 Bochum, Germany

July 20, 2021

Abstract

Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ denote independent p-dimensional vectors with independent complex or real valued entries such that $\mathbb{E}[\mathbf{x}_i] = \mathbf{0}$, $\mathrm{Var}(\mathbf{x}_i) = \mathbf{I}_p$, $i = 1, \ldots, n$, let $\mathbf{T}_n$ be a $p \times p$ Hermitian nonnegative definite matrix and f be a given function. We prove that an approriately standardized version of the stochastic process $\big(\mathrm{tr}(f(\mathbf{B}_{n,t}))\big)_{t \in [t_0,1]}$ corresponding to a linear spectral statistic of the sequential empirical covariance estimator

$$(\mathbf{B}_{n,t})_{t \in [t_0,1]} = \Big(\frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathbf{T}_n^{1/2} \mathbf{x}_i \mathbf{x}_i^{\star} \mathbf{T}_n^{1/2} \Big)_{t \in [t_0,1]}$$

converges weakly to a non-standard Gaussian process for $n, p \to \infty$. As an application we use these results to develop a novel approach for monitoring the sphericity assumption in a high-dimensional framework, even if the dimension of the underlying data is larger than the sample size.

Keywords: linear spectral statistic, sequential sample covariance matrix, sequential process, sphericity test, Stieltjes transform, monitoring spherictiy.
AMS subject classification: Primary 15A18, 60F17; Secondary 62H15

1 Introduction

Estimation and testing of a high-dimensional covariance matrix is a fundamental problem of statistical inference with numerous applications in biostatistics, wireless communications and finance (see, e.g., Fan and Li (2006), Johnstone (2006) and the references therein). Linear spectral statistics are frequently used to construct tests for various hypotheses. For example, Mauchly (1940) proposes 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) suggest 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) consider some classical q-sample testing problems under normality in a high-dimensional setting, which are further generalized in Dette and Dörnemann (2020) for an increasing number q of groups. Other authors concentrate 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), large-dimensional matrices with bivariate dependence measures as entries (Bao et al., 2015a; Li et al., 2019) or information-plus-noise matrices (Banna et al., 2020).

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) and 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 $\mathbf{B}_n$. Several authors have followed this line of research and tried to relax the assumptions for such statements (see Pan and Zhou, 2008; Lytova and Pastur, 2009; Pan, 2014; Zheng et al., 2015; Najim and Yao, 2016, among others).

In this paper we will take a different point of view on linear spectral statistics and study these objects from a sequential perspetive. 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 non-standard Gaussian process.

Our interest in these processes is partially motivated by the fact that the sequential covariance estimator plays a central role in the construction of methodology for the detection of structural breaks in the covariance structure (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 covariance matrices. In fact, we use the probabilistic results presented in this paper to develop a procedure for monitoring deviations from sphericity, see Section 3 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 work 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 contributions the results of the present paper 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 in the case $f(x) = x$), 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.

2 A sequential look at 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 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 Section 1).
For the statement of our main result we require some notation. Let

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

be the empirical spectral distribution of a $p \times p$ Hermitian matrix $\mathbf{A}$, where $\lambda_1(\mathbf{A}), \ldots, \lambda_p(\mathbf{A})$ are the eigenvalues of $\mathbf{A}$ (often the dependence on $\mathbf{A}$ is omitted in the notation, because it is clear from the context) and δ_a denotes the Dirac measure at a point $a \in \mathbb{R}$. A useful tool in random matrix theory is the Stieltjes transform

$$s_F(z) = \int \frac{1}{\lambda - z} dF(\lambda)$$

of a distribution function F on the real line, which is here defined for $z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}$. If $F = F^{\mathbf{A}}$ is an empirical spectral distribution, then its Stieltjes transform has the form

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

Standard results on linear spectral statistics (see, for example the monograph of Bai and Silverstein, 2010) show that (for fixed $t > 0$) under certain conditions, with probability 1, the empirical spectral distribution $F^{(n/\lfloor nt \rfloor)\mathbf{B}_{n,t}}$ of the (scaled) matrix $(n/\lfloor nt \rfloor)\mathbf{B}_{n,t}$ converges weakly. The limit, say $F^{y_t,H}$, is the so-called generalized Marčenko-Pastur distribution defined by its

Stieltjes transform $s_t = s_{F^{y_t,H}}$, which is the unique solution of the equation

$$s_t(z) = \int \frac{1}{\lambda(1 - y_t - y_t z s_t(z)) - z} dH(\lambda) \tag{2.1}$$

on the set $\{s_t \in \mathbb{C}^+ : \frac{1-y_t}{z} + y_t s_t \in \mathbb{C}^+\}$. Here, H denotes the limiting spectral distribution $H_n = F^{\mathbf{T}_n}$ of the Hermitian matrix $\mathbf{T}_n$ which will be assumed to exist throughout this paper and $y_t = y/t$. Hence, we have

$$\tilde{F}^{y_t,H}(x) := \lim_{n\to\infty} F^{\mathbf{B}_{n,t}}(x) = F^{y_t,H}(x/t). \tag{2.2}$$

at all points, where $\tilde{F}^{y_t,H}$ is continuous.
For the following discussion, define for $\mathbf{B}_{n,t}$ the $(\lfloor nt \rfloor \times \lfloor nt \rfloor)$-dimensional companion matrix

$$\underline{\mathbf{B}}_{n,t} = \frac{1}{n} \mathbf{X}_{n,t}^\star \mathbf{T}_n \mathbf{X}_{n,t} \tag{2.3}$$

and denote the limit (if it exists) of its spectral distribution $F^{\underline{\mathbf{B}}_{n,t}}$ and its corresponding Stieltjes transform by

$$\underline{\tilde{F}}^{y_t,H} \quad \text{and} \quad \underline{\tilde{s}}_t(z) = s_{\underline{\tilde{F}}^{y_t,H}}(z), \tag{2.4}$$

respectively. A straightforward calculation (using (2.1)) shows that this Stieltjes transform satisfies the equation

$$z = -\frac{1}{\underline{\tilde{s}}_t(z)} + y \int \frac{\lambda}{1 + \lambda t \underline{\tilde{s}}_t(z)} dH(\lambda). \tag{2.5}$$

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,

$$X_n(f,t) = \int f(x) dG_{n,t}(x), \tag{2.6}$$

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

$$\tilde{F}^{y_{\lfloor nt \rfloor},H_n}(x) = F^{y_{\lfloor nt \rfloor},H_n}\left(\frac{n}{\lfloor nt \rfloor} x\right) \tag{2.7}$$

is a rescaled version of the generalized Marčenko-Pastur distribution defined by (2.1). The proof is challenging and therefore deferred to Section 4 and the Appendix.

Theorem 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* $\mathbb{E}x_{ij} = 0$, $\mathbb{E}|x_{ij}|^2 = 1$, $\max_{i,j,n} \mathbb{E}|x_{ij}|^{12} < \infty$*. Moreover, the condition*

$$\frac{1}{np}\sum_{i=1}^{p}\sum_{j=1}^{n} \mathbb{E}\left[|x_{ij}|^4 I(|x_{ij}| \geq \sqrt{n}\eta)\right] \to 0 \tag{2.8}$$

holds for any $\eta > 0$.

(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*

$$\left[\liminf_{n\to\infty} \lambda_{\min}(\mathbf{T}_n) I_{(0,1)}(y_{t_0}) t_0 (1-\sqrt{y_{t_0}})^2, \limsup_{n\to\infty} \lambda_{\max}(\mathbf{T}_n)(1+\sqrt{y_{t_0}})^2\right]. \tag{2.9}$$

(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*

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

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* (2.9)*,* $\mathcal{C}_1, \mathcal{C}_2$ *are non overlapping and the function* $\sigma^2_{t_1,t_2}(z_1,z_2)$ *is defined in* (4.19)*.*

(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 2.1. 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 2.1 reveals the asymptotic behaviour 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 (2.8) and (2.9) 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 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

$$\mathbf{x}_j^\star \mathbf{A} \mathbf{x}_j - \operatorname{tr}(\mathbf{A}), \tag{2.10}$$

where $\mathbf{A}$ denotes a random $p \times p$ matrix independent of $\mathbf{x}_j$, $j \in \{1, \ldots, n\}$. In particular, the existence of the 12th-moment in Theorem 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 4.3.3). Strengthening the moment conditions on the underlying random variables appears to be a convienient 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 (2.10) (see their Lemma A.2 for more details).

(2) In order to allow for non-centralized data ($\mathbb{E}[x_{ij}] \neq 0$), Zheng et al. (2015) 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 2.1 combining the tools developed in this paper with methodology used in the proof of Theorem 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 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 2.1. For instance, in the non-sequential case, Najim and Yao (2016) show that the LvyProhorov 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 paper, 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.

In general, the calculation of the limiting parameters appearing in Theorem 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 a equation involving the limiting spectral distribution H 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 C.1 in the Appendix), and for specific functions f_1 and f_2 an explicit calculation of the asymptotic expectation and variance in Theorem 2.1 is possible. In the following corollary we illustrate this for 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 (see, e.g., Bao et al. (2015b), Cai et al. (2015), Nguyen and Vu (2014), Wang et al. (2018)) and has many applications in statistics. A proof can be found in Section C.2.

Corollary 2.1. *Let $t_0 \in (0, 1]$, and assume that condition (a) of Theorem 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*

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

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

3 Monitoring sphericity in large dimension

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

$$\mathrm{H}_0 : \mathbf{T}_n = \sigma^2 \mathbf{I}_p \text{ for some } \sigma^2 > 0, \quad \text{vs.} \quad \mathrm{H}_1 : \mathbf{T}_n \neq \sigma^2 \mathbf{I}_p \text{ for all } \sigma^2 > 0. \tag{3.1}$$

In general, it is well-known that the likelihood ratio test statistic for the hypotheses in (3.1) 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

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

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 Section 2 to develop a sequential monitoring procedure for the assumption of sphericity.

To be precise, we consider random variables $\mathbf{y}_1, \ldots, \mathbf{y}_n \in \mathbb{R}^p$, 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 $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ satisfying the asssumptions stated in Section 2. We are interested in monitoring the sphericity assumption

$$\begin{aligned} &\mathrm{H}_0 : \mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_n = \sigma^2 \mathbf{I}_p \text{ for some } \sigma^2 > 0 \\ \text{vs. } &H_1 : \mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_{\lfloor n t_1^\star \rfloor} = \sigma^2 \mathbf{I}_p, \ \mathbf{\Sigma}_{\lfloor n t_1^\star \rfloor + 1} = \ldots = \mathbf{\Sigma}_n \neq \sigma^2 \mathbf{I}_p, \end{aligned} \tag{3.2}$$

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

$$U_{n,t} = \frac{\frac{1}{p} \operatorname{tr}(\hat{\mathbf{\Sigma}}_{n,t}^2)}{\big(\frac{1}{p} \operatorname{tr} \hat{\mathbf{\Sigma}}_{n,t} \big)^2} \ , \tag{3.3}$$

and investigate the asymptotic behaviour of the stochastic process $U_n = (U_{n,t})_{t \in [t_0, 1]}$ under the null hypothesis. 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,

$$\hat{\mathbf{\Sigma}}_{n,t} = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathbf{y}_i \mathbf{y}_i^\top = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \mathbf{\Sigma}_i^{\frac{1}{2}} \mathbf{x}_i \mathbf{x}_i^\top \mathbf{\Sigma}_i^{\frac{1}{2}}. \tag{3.4}$$

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. The following theorem deals with the weak convergence of $(U_n)_{n\in\mathbb{N}}$ considered as a sequence in the space of bounded functions $\ell^\infty([t_0,1])$ and its proof is postponed to Section 5. In the following discussion the symbol $\rightsquigarrow$ denotes weak convergence of processes and the symbol $\stackrel{\mathcal{D}}{\to}$ weak convergences of a real-valued random variables.

Theorem 3.1. *Let $y \in (0,\infty)$, $t_0 > 0$ and define $y_t = y/t$ for $t \in [t_0,1]$. If the random variables $\mathbf{x}_1,\dots,\mathbf{x}_n$ satisfy the assumptions (a) and (1) of Theorem 2.1, it follows under the null hypothesis* (3.2) *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 3.1.

(1) To obtain a test for the hypotheses in (3.2) we note that the continuous mapping theorem implies under the null hypothesis

$$\sup_{t\in[t_0,1]} p\left(U_{n,t} - 1 - y_{\lfloor nt\rfloor}\right) \stackrel{\mathcal{D}}{\to} \sup_{t\in[t_0,1]} U_t,\ n\to\infty\ . \tag{3.5}$$

Therefore we propose to reject the null hypothesis in (3.2) whenever

$$\sup_{t\in[t_0,1]} p\left(U_{n,t} - 1 - y_{\lfloor nt\rfloor}\right) > c_\alpha, \tag{3.6}$$

where c_α denotes the $(1-\alpha)$-quantile of the statistic $\sup_{t\in[t_0,1]} U_t$. Thus, we have by (3.5)

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

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

(2) In order to investigate the consistency of the test (3.6) assume that the matrices $\mathbf{\Sigma}_i$ in (3.2) 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 $t \in (t_1^\star, 1)$

$$\frac{1}{p}\mathbb{E}\left[\mathrm{tr}\left(\hat{\mathbf{\Sigma}}_{n,t}\right)\right] \overset{\mathbb{P}}{\longrightarrow} t_1^\star\sigma^2 + (t-t_1^\star)g,$$
$$\frac{1}{p}\mathbb{E}\left[\mathrm{tr}\left(\hat{\mathbf{\Sigma}}_{n,t}^2\right)\right] \overset{\mathbb{P}}{\longrightarrow} \left(t_1^\star\right)^2\sigma^4 + 2t_1^\star\sigma^2(t-t_1^\star)g + (t-t_1^\star)^2h + yt_1^\star\sigma^4 + y(t-t_1^\star)g.$$

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 (4.1), this implies

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

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

$$\begin{aligned}U_{n,t} \overset{\mathbb{P}}{\longrightarrow} & \frac{(t_1^\star)^2\sigma^4 + 2t_1^\star\sigma^2(t-t_1^\star)g + (t-t_1^\star)^2h + yt_1^\star\sigma^4 + y(t-t_1^\star)g^2}{(t_1^\star)^2\sigma^4 + ((t-t_1^\star)g)^2 + 2t_1^\star\sigma^2(t-t_1^\star)g} \\ = & 1 + y_t + \Delta_{1,t} + \Delta_{2,t}\end{aligned}$$

where

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

by construction, and

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

Note that under the alternative in (3.2) 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 $\mathbf{\Sigma}_1, \dots, \mathbf{\Sigma}_n$ 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 (3.6) 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 (3.2).

Fisher et al. (2010) consider several generalizations of the classical test introduced by John (1971). Motivated by this work an alternative test for the hypothesis (3.2) 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 (3.4). 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 results 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 $\big(U_{n,t}^{(1)}\big)_{t\in[t_0,1]}$ in Remark 3.1 and the details are omitted for the sake of brevity.

Theorem 3.2. *Under the assumptions of Theorem 3.1 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

$$\operatorname{cov}(U^{(2)}_{t_1}, U^{(2)}_{t_2}) = \frac{8y^2\Big\{4t_1^2(2t_2^2+3t_2y+2y^2)+6t_1y(4t_2^2+5t_2y+2y^2)+y^2(21t_2^2+24t_2y+8y^2)\Big\}}{t_1^2(t_1+y)^2(t_2+y)^2}$$

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

Example 3.1. We conclude this section with a small simulation study illustrating the finite-sample properties of the test (3.6). For this purpose, we generated centered p-dimensional normally distributed data with various covariance structures. To be precise, we consider the the alternatives

$$\mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_{\lfloor nt^\star \rfloor} = \mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star \rfloor+1} = \ldots = \mathbf{\Sigma}_n = \mathbf{I}_p + \operatorname{diag}(\underbrace{0,\ldots,0}_{p/2},\underbrace{\delta,\ldots,\delta}_{p/2}), \tag{3.7}$$

$$\mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_{\lfloor nt^\star \rfloor} = \mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star \rfloor+1} = \ldots = \mathbf{\Sigma}_n = \mathbf{I}_p + \operatorname{diag}(\underbrace{0,\ldots,0}_{p/2},\underbrace{\delta,\ldots,\delta}_{p/2}) + \tilde{\mathbf{S}}(\delta), \tag{3.8}$$

$$\mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_{\lfloor nt^\star \rfloor} = \mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star \rfloor+1} = \ldots = \mathbf{\Sigma}_n = (1+\varepsilon)\mathbf{I}_p, \tag{3.9}$$

$$\mathbf{\Sigma}_1 = \ldots = \mathbf{\Sigma}_{\lfloor nt^\star \rfloor} = \mathbf{I}_p,\ \mathbf{\Sigma}_{\lfloor nt^\star \rfloor+1} = \ldots = \mathbf{\Sigma}_n = (1+\varepsilon)\mathbf{I}_p + \mathbf{S}(\varepsilon), \tag{3.10}$$

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 (3.2)). Here, the entries of the $p \times p$ matrix $\mathbf{S}(\varepsilon)$ in (3.10) 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 $\tilde{\mathbf{S}}(\varepsilon)$ in (3.8) 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 1 and Figure 2, we display the the empirical rejection of the test (3.6) 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 grey line in each figure defines the nominal level $\alpha = 5\%$.
Note that the choices $\delta = 0$ and $\varepsilon = 0$ correspond to the null hypothesis in (3.7), (3.8), (3.9) and (3.10), 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 (3.8) compared to the alternatives in (3.7). This reflects the intuition that the alternative in (3.7) is somehow closer to sphericity than the alternative (3.8). A similar observation can be made for the alternatives (3.9) and (3.10).

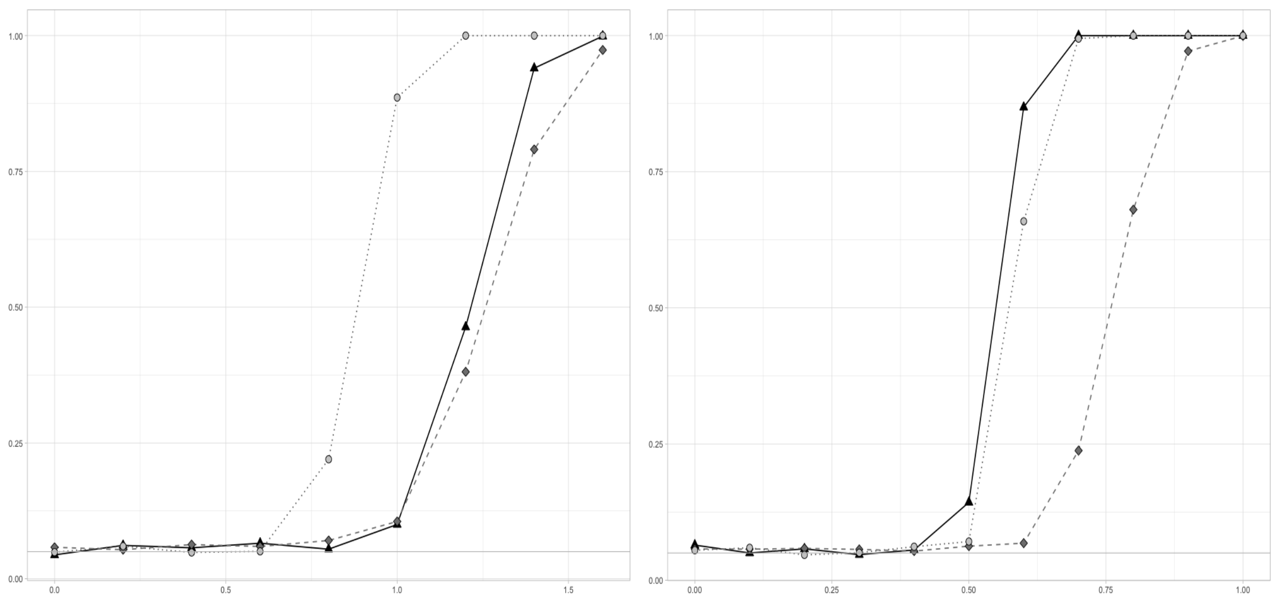

Figure 1: *Simulated rejection probabilities of the test* (3.6) *under the null hypothesis* ($\delta = 0$) *and the different alternatives in* (3.7) *(left) and* (3.8) *(right) for* $\delta > 0$ *. The circle indicates* $n = 200, p = 300$*, the triangle* $n = 200, p = 120$ *and the square* $n = 150, p = 300$*.*

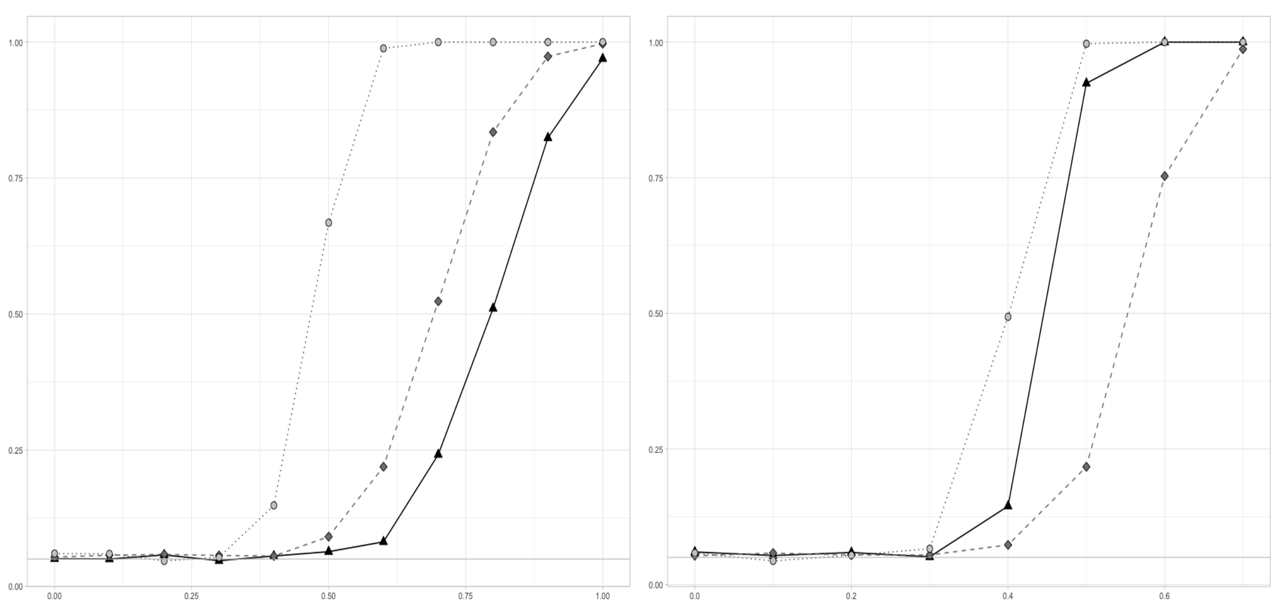

Figure 2: *Simulated rejection probabilities of the test* (3.6) *under the null hypothesis* ($\varepsilon = 0$) *and the different alternatives in* (3.9) *(left) and* (3.10) *(right) for* $\varepsilon > 0$*. The circle indicates* $n = 200, p = 300$*, the triangle* $n = 200, p = 120$ *and the square* $n = 150, p = 300$*.*

4 Proof of Theorem 2.1

4.1 Outline of the proof of Theorem 2.1

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

$$\int f(x)dG(x) = \frac{1}{2\pi i}\int\int_{\mathcal{C}} \frac{f(z)}{z-x}dzdG(x) = -\frac{1}{2\pi i}\int_{\mathcal{C}} f(z)s_G(z)dz, \tag{4.1}$$

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 (4.1) follows from Cauchys integral and Fubinis theorem. Thus invoking the continuous mapping theorem, it may suffice to prove weak convergence for the sequence $(M_n)_{n\in\mathbb{N}}$, where

$$M_n(z,t) = p\left(s_{F^{\mathbf{B}_{n,t}}}(z) - s_{\tilde{F}^{y_{\lfloor nt \rfloor},H_n}}(z)\right), \quad z \in \mathcal{C}. \tag{4.2}$$

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 (2.7) characterized through the equation

$$s_{\tilde{F}^{y_{\lfloor nt \rfloor},H_n}}(z) = \int \frac{1}{\lambda \frac{\lfloor nt \rfloor}{n}\left(1 - y_{\lfloor nt \rfloor} - y_{\lfloor nt \rfloor} z s_{\tilde{F}^{y_{\lfloor nt \rfloor},H_n}}(z)\right) - z} dH_n(\lambda), \tag{4.3}$$

and the contour $\mathcal{C}$ in (4.2) 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 all $n \in \mathbb{N}, t \in [t_0, 1]$. This idea is formalized in the proof of Theorem 2.1 in Section 4.2.

In order to prove the weak convergence of (4.2) define a contour $\mathcal{C}$ as follows. Let x_r be any number greater than the right endpoint of the interval (2.9) and $v_0 > 0$ be arbitrary. Let x_l be any negative number if the left endpoint of the interval (2.9) is zero. Otherwise, choose

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

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

$$\hat{M}_n(z,t) = \begin{cases} M_n(z,t) & \text{if } z \in \mathcal{C}_n, \\ M_n(x_r + in^{-1}\varepsilon_n, t) & \text{if } x = x_r,\ v \in [0, n^{-1}\varepsilon_n], \\ M_n(x_l + in^{-1}\varepsilon_n, t) & \text{if } x = x_l,\ v \in [0, n^{-1}\varepsilon_n]. \end{cases} \tag{4.4}$$

In Lemma A.2 in the Appendix, 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

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

in (4.1) coincide asymptotically. As a consequence the weak convergence of the process (4.2) follows from that of $\hat{M}_n$, which is established in the following theorem. The proof is given in Section 4.3.

Theorem 4.1 (Weak convergence for the process of Stieltjes transforms)**.** *Under the assumptions of Theorem 2.1, the sequence* $(\hat{M}_n)_{n\in\mathbb{N}}$ *defined in* (4.4) *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*

$$\mathbb{E}M(z,t) = \begin{cases} \dfrac{ty\int \frac{\underline{\tilde{s}}_t^3(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^3} dH(\lambda)}{\left(1 - ty\int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^2} dH(\lambda)\right)^2} & \text{for the real case,} \\ 0 & \text{for the complex case} \end{cases} \tag{4.5}$$

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

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

where $\sigma^2_{t_1,t_2}(z_1,z_2)$ *is defined in* (4.19)*. In the real case, we have*

$$\operatorname{cov}(M(z_1,t_1), M(z_2,t_2)) = 2\sigma^2_{t_1,t_2}(z_1, \overline{z_2}). \tag{4.6}$$

4.2 Proof of Theorem 2.1 using Theorem 4.1

From (4.1) we obtain

$$-\frac{1}{2\pi i}\int_{\mathcal{C}} f(z)\mathbb{E}s_G(z)dz = -\frac{1}{2\pi i}\mathbb{E}\int_{\mathcal{C}} f(z)s_G(z)dz = \mathbb{E}\int f(x)dG(x). \tag{4.7}$$

We choose v_0, x_r, x_l so that f_1 and f_2 given in Theorem 2.1 are analytic on and inside the resulting contour $\mathcal{C}$ and define

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

The almost sure convergence

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

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

$$\lambda_{\max}(\mathbf{AB}) \le \lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{B}),\ \ \lambda_{\min}(\mathbf{AB}) \ge \lambda_{\min}(\mathbf{A})\lambda_{\min}(\mathbf{B})$$

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

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

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

$$\liminf_{n\to\infty} \min\left(x_r - \lambda_{\max}(\mathbf{B}_{n,t}), \lambda_{\min}(\mathbf{B}_{n,t}) - x_l\right) > 0, \tag{4.8}$$

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 for (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 $\lfloor nt_1 \rfloor \neq \lfloor nt_2 \rfloor$. 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 independent of n and $t \in [t_0, 1]$). From Lemma A.1 in the Appendix, it follows that the support of $\tilde{F}^{y_{\lfloor nt \rfloor}, H_n}$, $t \in [t_0, 1]$, is contained in the interval

$$\Big[\frac{\lfloor nt_0 \rfloor}{n}\lambda_{\min}(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt_0 \rfloor})(1-\sqrt{y_{\lfloor nt_0 \rfloor}})^2, \lambda_{\max}(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt_0 \rfloor}})^2\Big],$$

which is enclosed by the contour $\mathcal{C}$ for sufficiently large n. Therefore, using (4.1) and (4.7), we have almost surely

$$\Big(\Big(-\frac{1}{2\pi i}\int_{\mathcal{C}} f_i(z) M_n(z,t)dz\Big)_{i=1,2}\Big)_{t\in[t_0,1]} = \big((X_n(f_i,t))_{i=1,2}\big)_{t\in[t_0,1]}$$

for sufficiently large n. Moreover, we have with probability 1 (see Lemma A.2 in the Appendix)

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

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

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

By Corollary 4.1 stated in Section 4.3.3 below and (4.5), the limiting process M in Theorem 4.1 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

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

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 integral must be Gaussian as well. The limiting expression for the mean and the covariance follow immediately from Theorem 4.1. For example, we have for the real case observing (4.6)

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

4.3 Proof of Theorem 4.1

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 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 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 4.1 describing the important steps, which are carried out in the following sections and the online appendix. We consider the stochastic processes $(M_n)_{n\in\mathbb{N}}$ and $(\hat{M}_n)_{n\in\mathbb{N}}$ (which is defined in (4.4)) as sequences in the space $\ell^\infty(\mathcal{C}^+ \times [t_0, 1])$ and use the decomposition

$$M_n = M_n^1 + M_n^2 \ , \tag{4.9}$$

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

$$\begin{aligned} M_n^1(z,t) =& p\left(s_{F^{\mathbf{B}_{n,t}}}(z) - \mathbb{E}\left[s_{F^{\mathbf{B}_{n,t}}}(z)\right]\right), & (4.10)\\ M_n^2(z,t) =& p\left(\mathbb{E}\left[s_{F^{\mathbf{B}_{n,t}}}(z)\right] - s_{\tilde{F}^{y_{\lfloor nt\rfloor}, H_n}}(z)\right), & (4.11)\end{aligned}$$

the Stieltjes transform $s_{\tilde{F}^{y_{\lfloor nt\rfloor}, H_n}}$ is defined in (4.3) 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 4.3.2.

Theorem 4.2. *Under the assumption (1) for the real case or assumption (2) for the complex case from Theorem 2.1, it holds for all* $k \in \mathbb{N}, t_1, t_2 \in [0, 1]$, $z_1, ..., z_k \in \mathbb{C}$, $\mathrm{Im}(z_i) \neq 0$

$$\begin{aligned} &(M_n^1(z_1, t_1), M_n^1(z_1, t_2), ..., M_n^1(z_k, t_1), M_n^1(z_k, t_2))^\top \\ &\xrightarrow{\mathcal{D}} (M^1(z_1, t_1), M^1(z_1, t_2), ..., M^1(z_k, t_1), M^1(z_k, t_2))^\top \ , \end{aligned} \tag{4.12}$$

where $M^1(z, t) = M(z, t) - \mathbb{E}[M(z, t)]$ *is the centered version of the Gaussian process defined in Theorem 4.1.*

Next, we define the process $\hat{M}_n^1$ in the same way as $\hat{M}_n$ in (4.4) replacing M_n by M_n^1 and show the following tightness result. The main argument in in its proof consists in establishing delicate moment inequalities for the increments of the process $(\hat{M}_n^1)_{n\in\mathbb{N}}$, see Lemma 4.1 and its proof in Appendix B.1.

Theorem 4.3. *Under the assumptions of Theorem 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 (4.5). Note that the space of bounded function is equipped with the sup-norm, which demands an 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 4.5 in Section 4.3.4.

Theorem 4.4. *Under the assumptions of Theorem 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 4.2, 4.3 and 4.4 are postponed to Section 4.3.2, 4.3.3 and 4.3.4, respectively. Using these results, we are now in the position to prove Theorem 4.1.

4.3.1 Proof of Theorem 4.1

Theorem 4.2 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 (4.4), 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 section (see Corollary 4.1 in Section 4.3.3) 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 4.3 establishes tightness, Theorem 4.1 follows from the decomposition (4.9), Theorem 1.5.6 in Van Der Vaart and Wellner (1996) and Theorem 4.4.

4.3.2 Proof of Theorem 4.2

We start by performing some preparations and by introducing notations which will remain crucial for the rest of this work. Using the CramrWold device, the convergence in (4.12) is equivalent to the weak convergence

$$\sum_{i=1}^{k} \left(\alpha_{i,1} M_n^1(z_i, t_1) + \alpha_{i,2}(M_n^1(z_i, t_2)\right) \overset{\mathcal{D}}{\to} \sum_{i=1}^{k} \left(\alpha_{i,1} M^1(z_i, t_1) + \alpha_{i,2}(M^1(z_i, t_2)\right) \tag{4.13}$$

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 2.1.

Recalling assumption (b) in Theorem 2.1, we may assume $||\mathbf{T}_n|| \leq 1$, $n \in \mathbb{N}$. 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). By the ShermanMorrison formula we obtain the representation

$$\mathbf{D}_t^{-1}(z) - \mathbf{D}_{j,t}^{-1}(z) = -\mathbf{D}_{j,t}^{-1}(z) \mathbf{r}_j \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z) \beta_{j,t}(z). \tag{4.14}$$

In order to prove asymptotic normality of the random variable appearing in (4.13), 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 (4.10) and using the martingale decomposition, we have

$$M_n^1(z, t) = \operatorname{tr}(\mathbf{D}_t^{-1}(z) - \mathbb{E}\mathbf{D}_t^{-1}(z))$$

$$
\begin{aligned}
&= \sum_{j=1}^{\lfloor nt \rfloor} \left(\operatorname{tr} \mathbb{E}_j \mathbf{D}_t^{-1}(z) - \operatorname{tr} \mathbb{E}_{j-1} \mathbf{D}_t^{-1}(z) \right) \\
&= \sum_{j=1}^{\lfloor nt \rfloor} \left(\operatorname{tr} \mathbb{E}_j \left[\mathbf{D}_t^{-1}(z) - \mathbf{D}_{t,j}^{-1}(z) \right] - \operatorname{tr} \mathbb{E}_{j-1} \left[\mathbf{D}_t^{-1}(z) - \mathbf{D}_{j,t}^{-1}(z) \right] \right) \\
&= - \sum_{j=1}^{\lfloor nt \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) \beta_{j,t}(z) \mathbf{r}_j^{\star} \mathbf{D}_{j,t}^{-2}(z) \mathbf{r}_j. \qquad (4.15)
\end{aligned}
$$

Since, from now on, the proof is similar in spirit to Section 9.9 in Bai and Silverstein (2010), we restrict ourselves to an overview explaining the main steps and important differences. By similar arguments as given in this monograph it can be shown, that 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

$$
\begin{aligned}
Z_{nj}^{t_1,t_2} &= \sum_{i=1}^{k} \left(\alpha_{i,1} Y_{j,t_1}(z_i) + \alpha_{i,2} Y_{j,t_2}(z_i) \right) \ , \\
Y_{j,t}(z) &= -\mathbb{E}_j \left[\overline{\beta}_{j,t}(z) \alpha_{j,t}(z) - \overline{\beta}_{j,t}^2(z) \hat{\gamma}_{j,t}(z) \frac{1}{n} \operatorname{tr}(\mathbf{T}_n \mathbf{D}_{j,t}^{-2}(z)) \right] = -\mathbb{E}_j \frac{d}{dz} \overline{\beta}_{j,t}(z) \hat{\gamma}_{j,t}(z)
\end{aligned}
$$

if $j \leq \lfloor nt \rfloor$ 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). It is straightforward to show that $Z_{nj}^{t_1,t_2}$ forms a martingale difference scheme with respect to to the filtration $\mathcal{F}_{nj} = \sigma(\{\mathbf{r}_1, ..., \mathbf{r}_j\})$ and we can prove that (5.31) in this reference holds true by deriving bounds for the 4th moment of $Y_{j,t}(z)$. For a proof of condition (5.30), we note that

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

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 $t_1, t_2 \in (0, 1]$

$$V_n(z_1, z_2, t_1, t_2) = \sum_{j=1}^{\min(\lfloor nt_1 \rfloor, \lfloor nt_2 \rfloor)} \mathbb{E}_{j-1}\left[Y_{j,t_1}(z_1) Y_{j,t_2}(z_2)\right] \overset{\mathbb{P}}{\to} \sigma^2_{t_1,t_2}(z_1, z_2) \tag{4.16}$$

for an appropriate function $\sigma^2_{t_1,t_2}(z_1, z_2)$ (see equation (4.19) below for a precise definition). Note that this convergence implies condition (5.29), since

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

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 CramrWold device yields the weak convergence of the finite-dimensional distributions to a multivariate normal distribution with covariance $\sigma^2_{t_1,t_2}(z_1, \overline{z_2}) = \mathrm{cov}(M^1(z_1, t_1), M^1(z_2, t_2))$.

Hence, it is remains to show (4.16) in order to establish the convergence of the finite dimensional distributions. For this purpose, we introduce the quantity

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

and note that it can be shown by similar arguments as on p. 273 in Bai and Silverstein (2010) that

$$\frac{\partial^2}{\partial z_1 \partial z_2} V_n^{(2)}(z_1, z_2, t_1, t_2) = V_n(z_1, z_2, t_1, t_2) + o_{\mathbb{P}}(1) . \tag{4.17}$$

Moreover, we can show

$$V_n^{(2)}(z_1, z_2, t_1, t_2) \overset{\mathbb{P}}{\to} a(z_1, z_2, t_1, t_2) \int_0^{\min(t_1,t_2)} \frac{1}{1 - \lambda a(z_1, z_2, t_1, t_2)} d\lambda, \ n \to \infty, \tag{4.18}$$

where

$$a(z_1, z_2, t_1, t_2) = \frac{\tilde{\underline{s}}_{t_2}(z_2) - \tilde{\underline{s}}_{t_1}(z_1) + (z_1 - z_2)\tilde{\underline{s}}_{t_1}(z_1)\tilde{\underline{s}}_{t_2}(z_2)}{t_2 \tilde{\underline{s}}_{t_2}(z_2) - t_1 \tilde{\underline{s}}_{t_1}(z_1)}$$

and the Stieltjes transform $\underline{\tilde{s}}_t(z)$ is defined in (2.4). From (4.17) and (4.18) it follows that

$$\sigma^2_{t_1,t_2}(z_1,z_2) = \frac{\partial^2}{\partial z_1 \partial z_2} \int\limits_0^{\min(t_1,t_2)a(z_1,z_2,t_1,t_2)} \frac{1}{1-\lambda} d\lambda = \frac{\partial}{\partial z_2}\Big(\frac{\min(t_1,t_2)\frac{\partial}{\partial z_1}a(z_1,z_2,t_1,t_2)}{1-\min(t_1,t_2)a(z_1,z_2,t_1,t_2)}\Big)$$
$$= \frac{\text{numerator}}{\text{denominator}}, \tag{4.19}$$

where

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

The proofs of (4.17) and (4.18) are very similar to Bai and Silverstein (2010) and omitted for the sake of brevity. Note also, that for the special case $t_1 = t_2 = 1$, this covariance structure coincides with formula (9.8.4) in this monograph.

4.3.3 Proof of Theorem 4.3 and continuity of the limiting process

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

$$m^1(z,t,z',z'') = \min\{|\hat{M}_n^1(z,t) - \hat{M}_n^1(z',t)|, |\hat{M}_n^1(z,t) - \hat{M}_n^1(z'',t)|\}, \tag{4.20}$$
$$m^2(z,t,t',t'') = \min\{|\hat{M}_n^1(z,t) - \hat{M}_n^1(z,t')|, |\hat{M}_n^1(z,t) - \hat{M}_n^1(z,t'')|\}, \tag{4.21}$$

where $t,t',t'' \in [t_0,1]$ and $z,z',z'' \in \mathcal{C}^+$. In order to find estimates for the tails of (4.20) and (4.21), we establish in the following lemma estimates on the moments of the increments of $\hat{M}_n^1(z,t)$, which are proved in Appendix B.1. For this purpose note that it follows from (4.15) that

$$\hat{M}_n^1(z,t_1) - \hat{M}_n^1(z,t_2) = \hat{Z}_n^1(z,t_1,t_2) + \hat{Z}_n^2(z,t_1,t_2).$$

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

$$Z_n^1(z,t_1,t_2) = \sum_{j=1}^{\lfloor nt_1 \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1}) \left(\beta_{j,t_2}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j - \beta_{j,t_1}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t_1}^{-2}(z)\mathbf{r}_j \right), \tag{4.22}$$

$$Z_n^2(z,t_1,t_2) = \sum_{j=\lfloor nt_1 \rfloor+1}^{\lfloor nt_2 \rfloor} (\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)\mathbf{r}_j^\star \mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j \tag{4.23}$$

using the definition (4.4).

Lemma 4.1. *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 4.3*

$$\mathbb{E}|\hat{M}_n^1(z_1,t) - \hat{M}_n^1(z_2,t)|^{2+\delta} \leq K|z_1 - z_2|^{2+\delta}, \tag{4.24}$$

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

$$\mathbb{E}|\hat{Z}_n^1(z,t_1,t_2)|^4 \leq K\Big(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\Big)^4, \tag{4.25}$$

$$\mathbb{E}|\hat{Z}_n^2(z,t_1,t_2)|^{4+\delta} \leq K\Big(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\Big)^{2+\delta/2}. \tag{4.26}$$

In order to simplify notation, we write $a \lesssim b$ for $a \leq Kb$, where $a, b \geq 0$ 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 4.3 by using results from Lemma 4.1.

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

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

In the case $t'' - t' \geq 1/n$, we use Lemma 4.1 and obtain

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

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 $t' \leq t \leq t'' \leq 1$

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

In order to derive a similar estimate for the term m^1, we note that it follows for $z, z', z'' \in \mathcal{C}_n$

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

where we used Lemma 4.1 in the last line. Moreover, we have

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

$$
\begin{aligned}
&\lesssim C_{1,\lambda}\Big[\Big|\frac{\lfloor nt_2\rfloor - \lfloor nt_1\rfloor}{n}\Big|^{2+\delta/2} + |z_1 - z_2|^{2+\delta}\Big] \\
&\leq C_{1,\lambda}\Big[\Big(|t_2 - t_1| + \frac{1}{n}\Big)^{2+\delta/2} + |z_1 - z_2|^{2+\delta}\Big] \\
&\leq C_{1,\lambda}\Big(\Big\|(z_1,t_1)^\top - (z_2,t_2)^\top\Big\|_\infty + \frac{1}{n}\Big)^{2+\delta/2},
\end{aligned}
$$

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 4.3.

Corollary 4.1. *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 $(z,t,\omega) \in (\mathcal{C}^+ \backslash \{x_l, x_r\}) \times [t_0,1] \times \Omega \mapsto \hat{M}^1(z,t)(\omega)$ are continuous. Since $(\mathcal{C}^+ \backslash \{x_l, x_r\}) \times [t_0,1] \subset \mathcal{C}^+ \times [t_0,1]$ 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. □

4.3.4 Proof of Theorem 4.4

Let

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

be the Stieltjes transform of the empirical spectral distribution $F^{\underline{\mathbf{B}}_{n,t}}$ of the matrix $\underline{\mathbf{B}}_{n,t}$ defined in (2.3), and let

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

be the Stieltjes transform of the distribution

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

with $\underline{F}^{y_{\lfloor nt \rfloor},H_n} - y_{\lfloor nt \rfloor}F^{y_{\lfloor nt \rfloor},H_n} = (1-y_{\lfloor nt \rfloor})I_{[0,\infty)}$. Recalling the definition (4.11) we have

$$M_n^2(z,t) = p\left(\mathbb{E}[\tilde{s}_{n,t}(z)] - \tilde{s}^0_{n,t}(z)\right) = \lfloor nt \rfloor\left(\mathbb{E}\left[\underline{\tilde{s}}_{n,t}(z)\right] - \underline{\tilde{s}}^0_{n,t}(z)\right). \tag{4.27}$$

We begin with a lemma which can be 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 B.2 of the online supplement.

Lemma 4.2.

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

where

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

$$\mathbf{q}_j = \frac{1}{\sqrt{p}}\mathbf{x}_j.$$

The next main step is the following result, which is proved in Section B.3.

Theorem 4.5. *Under the assumptions of Theorem 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* (2.4)*.*

The third step in the proof of Theorem 4.4 is the following result, which is proved in Section B.4 of the online supplement.

Theorem 4.6. *Under the assumptions of Theorem 2.1, we have*

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

With these preparations we show in Section B.5 that

$$\lfloor nt \rfloor R_{n,t}(z)\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] \to \begin{cases} \frac{y\int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^3}dH(\lambda)}{1-ty\int \frac{\underline{\tilde{s}}_t^2(z)\lambda^2}{(t\underline{\tilde{s}}_t(z)\lambda+1)^2}dH(\lambda)} & \text{for the real case,} \\ 0 & \text{for the complex case,} \end{cases} \tag{4.28}$$

uniformly with respect to $z \in \mathcal{C}_n, t \in [t_0, 1]$. Combining this result with Theorem 4.5 and Lemma B.8 yields

$$\lim_{n\to\infty} \sup_{\substack{z\in\mathcal{C}_n\\ t\in[t_0,1]}} |R_{n,t}(z)| = 0 \tag{4.29}$$

This result and Lemma B.8, Theorem 4.5, Theorem 4.6, Proposition B.1 and the equation (2.5) show that

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

Observing the representation in (4.27), Lemma 4.2 and Theorem 4.6, 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 4.4.

5 Proof of Theorem 3.1

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 $\hat{\mathbf{\Sigma}}_{n,t} = \mathbf{B}_{n,t}$. We apply Theorem 2.1 for the special case $f_1(x) = x,\ f_2(x) = x^2, \mathbf{T}_n = \mathbf{I}$, that is

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

Note that all conditions from Theorem 2.1 are satisfied, and therefore

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

in the space $(\ell^\infty([t_0, 1]))^2$, where $\big((X(f_1, t))_{t \in [t_0,1]}, (X(f_2, t))_{t \in [t_0,1]} \big)$ is a Gaussian process. Thus, it is left to calculate mean, covariance and the centering term appearing in Theorem 2.1. A tedious calculation in Section C.2 shows

$$\mathbb{E}[X(f_1, t)] = 0,\ \mathbb{E}[X(f_2, t)] = ty, \tag{5.1}$$

and

$$\begin{aligned} \operatorname{cov}(X(f_1, t_1), X(f_1, t_2)) &= 2y \min(t_1, t_2), \\ \operatorname{cov}(X(f_2, t_1), X(f_2, t_2)) &= 4 \min(t_1, t_2) y \left\{ 2t_1 t_2 + \left[\min(t_1, t_2) + 2(t_1 + t_2) \right] y + 2y^2 \right\}, \\ \operatorname{cov}(X(f_1, t_1), X(f_2, t_2)) &= 4 \min(t_1, t_2) y (t_2 + y). \end{aligned} \tag{5.2}$$

In Section C.2, we also calculate the centering terms for $X_n(f_1, t)$ and $X_n(f_2, t)$. With the definition $\phi(x, y) = \frac{y}{x^2}$, we obtain the representation

$$U_{n,t} = \phi \Big(\frac{1}{p} \operatorname{tr}(\mathbf{B}_{n,t}), \frac{1}{p} \operatorname{tr}(\mathbf{B}_{n,t}^2) \Big).$$

for the process $U_{n,t}$ in (3.3). Consequently, the assertion can be proved by the functional delta method.

To be precise, note that it follows from $y_n = p/n$

$$p \begin{pmatrix} \frac{1}{p} \operatorname{tr}(\mathbf{B}_{n,t}) - \frac{\lfloor nt \rfloor}{n} \\ \frac{1}{p} \operatorname{tr}(\mathbf{B}_{n,t}^2) - \frac{\lfloor nt \rfloor}{n} \left(\frac{\lfloor nt \rfloor}{n} + y_n \right) \end{pmatrix}_{t \in [t_0,1]} \rightsquigarrow \begin{pmatrix} X(f_1, t) \\ X(f_2, t) \end{pmatrix}_{t \in [t_0,1]} \tag{5.3}$$

in $(\ell^\infty([t_0,1]))^2$. Let $a_n(t) = \frac{\lfloor nt \rfloor}{n}$ and $b_n(t) = \frac{\lfloor nt \rfloor}{n}\big(\frac{\lfloor nt \rfloor}{n} + y_n\big)$, such that

$$\lim_{n\to\infty} a_n(t) = t = a(t), \quad \lim_{n\to\infty} b_n(t) = t(t+y) = b(t)$$

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

$$\phi\left(a_n(t), b_n(t)\right) = \frac{\frac{\lfloor nt \rfloor}{n}\big(\frac{\lfloor nt \rfloor}{n} + y_n\big)}{\big(\frac{\lfloor nt \rfloor}{n}\big)^2} = \frac{n}{\lfloor nt \rfloor}\Big(\frac{\lfloor nt \rfloor}{n} + y_n\Big) = 1 + y_{\lfloor nt \rfloor}.$$

Thus, it follows from (5.3) and Theorem 3.9.5 in Van Der Vaart and Wellner (1996) that

$$p\left\{U_{n,t} - 1 - y_{\lfloor nt \rfloor}\right\}_{t\in[t_0,1]} = p\Big\{\phi\Big(\frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}), \frac{1}{p}\operatorname{tr}(\mathbf{B}_{n,t}^2)\Big) - \phi\left(a_n(t), b_n(t)\right)\Big\}_{t\in[t_0,1]} \rightsquigarrow (U_t)_{t\in[t_0,1]}$$

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 (5.1) and (5.2) 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 3.1.

Acknowledgements. The work of N. Dörnemann was supported by the Deutsche Forschungsgemeinschaft (RTG 2131). We are grateful to Zhidong Bai and Jack Silverstein for the discussion

of the proof of formula (9.11.1) in Bai and Silverstein (2010).

References

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.

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., and Zhou, W. (2015b). The logarithmic law of random determinant. *Bernoulli*, 21(3):1600–1628.

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.

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. (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.

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.

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.

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.

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.

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.

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.

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.

Mauchly, J. W. (1940). Significance test for sphericity of a normal N-variate distribution. *Annals of Mathematical Statistics*, 11:204–209.

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.

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.

Van Der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence*. Springer.

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.

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.

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. (2015). 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.

Online supplement: technical details

A Details for the arguments in Section 4.2

Lemma A.1. *Let Γ_F denote the support of a cdf F. Then it holds*

$$\Gamma_{\tilde{F}^{y_{\lfloor nt \rfloor}, H_n}} \subset \Big[\frac{\lfloor nt_0 \rfloor}{n} \lambda_{\min}(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt_0 \rfloor})(1 - \sqrt{y_{\lfloor nt_0 \rfloor}})^2, \lambda_{\max}(\mathbf{T}_n)(1 + \sqrt{y_{\lfloor nt_0 \rfloor}})^2 \Big].$$

The proof of Lemma A.1 follows from Lemma 6.1, Bai and Silverstein (2010) or Proposition 2.17, Yao et al. (2015) and is therefore omitted.
The following lemma ensures that the process $(\hat{M}_n(z,t))_{z \in \mathcal{C}^+, t \in [t_0,1]}$ defined in (4.4) provides an appropriate approximation for the process $(M_n(z,t))_{z \in \mathcal{C}^+, t \in [t_0,1]}$.

Lemma A.2. *Let $i \in \{1,2\}$. It holds for all large n and for all $z \in \mathcal{C}^+, t \in [t_0,1]$ with probability 1 (uniformly in $t \in [t_0,1]$)*

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

Proof of Lemma A.2. 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 $\hat{M}_n$)

$$\begin{aligned} \Big| \int_{\mathcal{C}} f(z) \Big(M_n(z,t) - \hat{M}_n(z,t) \Big) dz \Big| - +* \leq K \int_{[0,n^{-1}\varepsilon_n]} \Big\{ & |M_n(x_r + iv, t) - M_n(x_r + in^{-1}\varepsilon_n, t)| \\ & + |M_n(x_l + iv, t) - M_n(x_l + in^{-1}\varepsilon_n, t)| \Big\} dv. \end{aligned}$$

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

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

where $z \in \mathbb{C} \setminus \Gamma_F$ and s_F is the Stieltjes transform of F. Using (4.8) and Lemma A.1, we have for $v \in [0, n^{-1}\varepsilon_n]$ and sufficiently large n

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

Similarly, one can show that for sufficiently large n

$$\operatorname{dist}\left(x_r + iv, \Gamma_{\tilde F^{y_{\lfloor nt \rfloor}},H_n}\right) \geq \left|x_r - \max\left(\lambda_{\max}(\mathbf{B}_{n,t}), \lambda_{\max}(\mathbf{T}_n)(1+\sqrt{y_{\lfloor nt \rfloor}})^2\right)\right|,$$
$$\operatorname{dist}\left(x_l + iv, \Gamma_{F^{\mathbf{B}_{n,t}}}\right) \geq \left|x_l - \min\left(\lambda_{\min}(\mathbf{B}_{n,t}), \frac{\lfloor nt_0 \rfloor}{n}\lambda_{\min}(\mathbf{T}_n) I_{(0,1)}(y_{\lfloor nt \rfloor})(1-\sqrt{y_{\lfloor nt \rfloor}})^2\right)\right|.$$

Recall the definition of M_n, then (A.1) implies

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

Due to (4.8), 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)

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

□

B More details for the proof of Theorem 4.1

In this section we provide the remaining arguments in the proof of Theorem 4.1 in Section 4.3. Several further very technical results are given in Section B.6.

B.1 Proof of Lemma 4.1

To be precise, recall the definition of Z_n^1 and Z_n^2 in (4.22) and (4.23) 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 (4.4). The bounds (4.25) and (4.26) for the moments of $\hat Z_n^1$ and $\hat Z_n^2$ follow directly from corresponding bounds (B.27) and (B.28) in Lemma B.5.

We continue by proving the first assertion (4.24). If z_1 and z_2 are both contained in $\mathcal{C}_n$, the assertion directly follows from (B.26). 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, $0 \leq \mathrm{Im}(z_2) \leq \varepsilon_n n^{-1} \leq \mathrm{Im}(z_1)$. With the notation $\mathrm{Re}(z_2) = x \in \{x_l, x_r\}$ we have from (B.26)

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

Finally, if both $z_1, z_2 \in \mathcal{C}^+ \setminus \mathcal{C}_n$, it follows from (B.26) that

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

which completes the proof of Lemma 4.1.

B.2 Proof of Lemma 4.2

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

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

and Lemma B.7, we have

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

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

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

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

B.3 Proof of Theorem 4.5

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 an uniform limit.

Lemma B.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 $z \in \mathcal{C}^+, t \in [t_0,1]$,*

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

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* (2.2)

Proof of Lemma B.1. We will start showing 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

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

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

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

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

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

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

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 Zhou (2008) for their estimate (3.4) 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 $l\in\{0,1\}$

$$\frac{1}{p}\Big(\mathbb{E}\operatorname{tr}\mathbf{T}_n^l\Big(\frac{\lfloor nt\rfloor}{n}\mathbf{K}-z\mathbf{I}\Big)^{-1}-\mathbb{E}\operatorname{tr}\mathbf{T}_n^l\mathbf{D}_t^{-1}(z)\Big)\to 0,\ n\to\infty. \tag{B.2}$$

Using (B.2) with $l=0$ for the first line and $l=1$ for the second one, we have

$$\frac{1}{p}\mathbb{E}\operatorname{tr}\Big(\frac{\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}-z\mathbf{I}\Big)^{-1}-\mathbb{E}\tilde{s}_{n,t}(z)\to 0, \tag{B.3}$$

$$\frac{1}{p}\mathbb{E}\operatorname{tr}\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n\Big(\frac{\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n}{1+y_{\lfloor nt\rfloor}a_{n,t}(z)}-z\mathbf{I}\Big)^{-1}-a_{n,t}(z)\to 0, \tag{B.4}$$

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 (B.4)

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

Combining this with (B.3) 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 (B.3), we get

$$\frac{1}{p}\mathbb{E}\,\mathrm{tr}\left(\frac{\lfloor nt\rfloor}{n}\mathbf{T}_n\left(1-y_{\lfloor nt\rfloor}(1+z\mathbb{E}\tilde{s}_{n,t}(z))\right)-z\mathbf{I}\right)^{-1}-\mathbb{E}\tilde{s}_{n,t}(z)\to 0. \tag{B.5}$$

Due to (B.5), 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 $t\in[0,1], z\in\mathcal{C}^+$

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

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 B.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 B.2. The sequence $(\mathbb{E}\hat{m}_{n,\cdot}(\cdot))_{n\in\mathbb{N}}$ is bounded, since by Lemma B.4, we get uniformly with respect to $t \in [t_0, 1], z \in \mathcal{C}_n, n \in \mathbb{N}$

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

We observe for $t_2 \geq t_1$ and $z \in \mathcal{C}_n$

$$\begin{aligned}
|\mathbb{E}[\tilde{s}_{n,t_1}(z)] - \mathbb{E}[\tilde{s}_{n,t_2}(z)]| =& \left|\frac{1}{p}\mathbb{E}\left[\operatorname{tr}\left(\mathbf{D}_{t_1}^{-1}(z) - \mathbf{D}_{t_2}^{-1}(z)\right)\right]\right| = \left|\frac{1}{p}\sum_{j=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}\mathbb{E}\left[\mathbf{r}_j^\star\mathbf{D}_{t_1}^{-1}(z)\mathbf{D}_{t_2}^{-1}(z)\mathbf{r}_j\right]\right| \\
=& \left|\frac{1}{p}\sum_{j=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}\left\{\mathbb{E}\left[\mathbf{r}_j^\star\mathbf{D}_{t_1}^{-1}(z)\mathbf{D}_{j,t_2}^{-1}(z)\mathbf{r}_j\right]\right.\right. \\
&\left.\left. - \mathbb{E}\left[\beta_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{t_1}^{-1}(z)\mathbf{D}_{j,t_2}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}(z)\mathbf{r}_j\right]\right\}\right| \\
\leq& K y_n\frac{\lfloor nt_2\rfloor - \lfloor nt_1\rfloor}{n} \leq K\frac{\lfloor nt_2\rfloor - \lfloor nt_1\rfloor}{n},
\end{aligned}$$

where the constant K is independent of t_1, t_2, z, n. Thus, we have

$$\left|\mathbb{E}[\hat{s}_{n,t_1}(z)] - \mathbb{E}[\hat{s}_{n,t_2}(z)]\right| \leq K\frac{\lfloor nt_2\rfloor - \lfloor nt_1\rfloor}{n}.$$

We aim to show

$$\lim_{\delta\to 0}\sup_{n\in\mathbb{N}}\sup_{\substack{(t_1,t_2,t)\in\mathcal{A}_\delta,\\ z\in\mathcal{C}^+}}\min\left(|\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_1}(z)]|, |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t_2}(z)]|\right) = 0, \tag{B.6}$$

where

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

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 $n \in \{1, \ldots, N\}$

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

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

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

For $n \geq N$ we conclude

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

and obtain

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

Thus, (B.6) 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

$$\begin{aligned}0 = & \lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \sup_{z \in \mathcal{C}^+} \inf_{\substack{(t_0,\ldots,t_r) \in \mathcal{B}_{\delta,r}, \\ r \in \mathbb{N}}} \max_{0 \leq i \leq r} \sup_{t,t' \in [t_{i-1},t_i)} |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t'}(z)]| \\ = & \lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \sup_{z \in \mathcal{C}^+} \inf_{\substack{(t_0,\ldots,t_r) \in \mathcal{B}_{\delta,r}, \\ r \in \mathbb{N}}} \max_{0 \leq i \leq r} \sup_{t,t' \in [t_{i-1},t_i\rangle} |\mathbb{E}[\hat{m}_{n,t}(z)] - \mathbb{E}[\hat{m}_{n,t'}(z)]|, \end{aligned} \tag{B.7}$$

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

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

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

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

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

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

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

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

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

$$\lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \sup_{\substack{t \in [0,1], \\ z_1, z_2 \in \mathcal{C}^+, \\ |z_1 - z_2| < \delta}} |\mathbb{E}[\hat{s}_{n,t}(z_1)] - \mathbb{E}[\hat{s}_{n,t}(z_2)]| = 0. \tag{B.8}$$

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 (B.7) and (B.8)

$$\begin{aligned}&\lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \inf_{\substack{((t_0,z_0),\ldots,(t_r,z_r)) \in \mathcal{B}^{(2)}_{\delta,r}, \\ r \in \mathbb{N}}} \max_{0 \leq i \leq r} \sup_{\substack{t,t' \in [t_{i-1}, t_i) \\ z,z' \in [z_{i-1}, z_i\rangle}} |\mathbb{E}[\hat{s}_{n,t}(z)] - \mathbb{E}[\hat{s}_{n,t'}(z')]| \\ &= \lim_{\delta \to 0} \sup_{n \in \mathbb{N}} \inf_{\substack{((t_0,z_0),\ldots,(t_r,z_r)) \in \mathcal{B}^{(2)}_{\delta,r}, \\ r \in \mathbb{N}}} \max_{0 \leq i \leq r} \sup_{\substack{t,t' \in [t_{i-1}, t_i) \\ z,z' \in [z_{i-1}, z_i\rangle}} |\mathbb{E}[\hat{m}_{n,t}(z)] - \mathbb{E}[\hat{m}_{n,t'}(z')]| \\ &= 0,\end{aligned} \tag{B.9}$$

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 (B.9) 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 4.5. From Lemma B.1 and Lemma B.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

$$\lim_{n\to\infty} d_2|_{\mathcal{C}_n\times[t_0,1]}(\mathbb{E}[\tilde{s}_{n,\cdot}(\cdot)], \tilde{s}_{\cdot}(\cdot)) = 0. \tag{B.10}$$

The considerations in the proof of Lemma B.2 reveal that $\mathbb{E}[\tilde{s}_{\cdot}(\cdot)] \in C(\mathcal{C}^+ \times [t_0,1])$. In this case, the convergence in the Skorokhod space in (B.10) 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

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

we also have

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

□

B.4 Proof of Theorem 4.6

The second assertion directly follows from the first one combined with Theorem 4.5. 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 B.3. *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 B.3. We have for sufficiently large n

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

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 A.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 4.6, we note that it follows from (B.1) in the proof of Lemma 4.2 in Section B.2 that

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

which is equivalent to

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

Note that $\tilde{\underline{s}}^0_{n,t}(z)$ is uniformly bounded which follows by a similar argument as given in the proof of Lemma B.3. In order to show that the the sequence $(M_n^2)_{n\in\mathbb{N}}$ is uniformly bounded, by using (4.28), 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

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

holds uniformly. Similarly to the proof of Lemma B.8, we conclude that for any bounded subset

$S \subset \mathbb{C}^+$

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

Using this, Hlder's inequality, Lemma B.3 and the identities

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

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

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

where we used the fact that $\operatorname{Im}(R_{n,t}(z)) + \operatorname{Im}(z) \ge 0$ for sufficiently large n, which follows from Lemma B.10. This finishes the proof of Theorem 4.6.

B.5 Proof of the statement (4.28)

As a preparation, we need the following proposition. The proof is omitted for the sake of brevity.

Proposition B.1.

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

Using (4.14) and the representation (B.51) we obtain

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

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

$$\begin{aligned}
T_{n,1}(z,t)=&\ y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\overline{\beta}_{j,t}^2(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j\\
&-\frac{1}{p}\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big]\Big\}\hat{\gamma}_{j,t}(z)\Big], \qquad \text{(B.11)}\\
T_{n,2}(z,t)=&-\frac{1}{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\beta_{j,t}(z)\right]\mathbb{E}\Big[\beta_{j,t}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z)\Big(\frac{\lfloor nt\rfloor}{n}\mathbb{E}\underline{\tilde{s}}_{n,t}(z)\mathbf{T}_n+\mathbf{I}\Big)^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j\Big], \qquad \text{(B.12)}
\end{aligned}$$

For this argument we used the fact

$$\mathbb{E}\Big[\overline{\beta}_{j,t}(z)\Big\{\mathbf{q}_j^\star\mathbf{T}_n^{\frac{1}{2}}\mathbf{D}_{j,t}^{-1}(z)(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n^{\frac{1}{2}}\mathbf{q}_j \\ -\frac{1}{p}\mathbb{E}\Big[\operatorname{tr}(\frac{\lfloor nt\rfloor}{n}\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n+\mathbf{I})^{-1}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z)\Big]\Big\}\Big]=0$$

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

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

For the term in (B.11) 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 (B.12)

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

We now consider the 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 (4.28) in this case.

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

$$\begin{aligned}
\lfloor nt \rfloor R_{n,t}(z)\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)] &= T_{n,1}(z,t) + T_{n,2}(z,t) + o(1) \\
&= \frac{z^2\underline{\tilde{s}}_t^2(z)}{\lfloor nt \rfloor n} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}\Big[\operatorname{tr} \mathbf{D}_{j,t}^{-1}(z) \Big(\frac{\lfloor nt \rfloor}{n} \mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\mathbf{T}_n + \mathbf{I} \Big)^{-1} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n \Big] + o(1) \\
&= \frac{z^2\underline{\tilde{s}}_t^2(z)}{\lfloor nt \rfloor n} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}\big[\operatorname{tr} \mathbf{D}_{j,t}^{-1}(z) \big(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I} \big)^{-1} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z)\mathbf{T}_n \big] + o(1). \quad \text{(B.13)}
\end{aligned}$$

For a detailed analysis of the random variable in (B.13) we use the decomposition

$$\mathbf{D}_{j,t}^{-1}(z) = -\Big(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n \Big)^{-1} + b_{j,t}(z)\mathbf{A}_t(z) + \mathbf{B}_t(z) + \mathbf{C}_t(z), \quad \text{(B.14)}$$

where

$$\begin{aligned}
\mathbf{A}_t(z) &= \sum_{i \neq j, 1 \leq i \leq \lfloor nt \rfloor} \Big(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n \Big)^{-1} \left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n \right) \mathbf{D}_{i,j,t}^{-1}(z), \quad \text{(B.15)} \\
\mathbf{B}_t(z) &= \sum_{i \neq j, 1 \leq i \leq \lfloor nt \rfloor} (\beta_{i,j,t}(z) - b_{j,t}(z)) \Big(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n \Big)^{-1} \mathbf{r}_i\mathbf{r}_i^\star \mathbf{D}_{i,j,t}^{-1}(z), \\
\mathbf{C}_t(z) &= b_{j,t}(z) \Big(z\mathbf{I} - \frac{\lfloor nt \rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n \Big)^{-1} \mathbf{T}_n n^{-1} \sum_{i \neq j} \left(\mathbf{D}_{i,j,t}^{-1}(z) - \mathbf{D}_{j,t}^{-1}(z) \right).
\end{aligned}$$

(here, we do not reflect the dependence on index j in our notation). We now investigate these terms in more detail.
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 $\mathbf{M}$. Then, one can show the following bounds (similarly to the inequalities (9.9.14) and (9.9.15) in Bai and Silverstein (2010))

$$\mathbb{E}|\operatorname{tr}(\mathbf{B}_t(z)\mathbf{M})| \leq K||\mathbf{M}||n^{\frac{1}{2}}, \tag{B.16}$$
$$|\operatorname{tr}(\mathbf{C}_t(z)\mathbf{M})| \leq K||\mathbf{M}||. \tag{B.17}$$

Moreover, we have for any nonrandom $\mathbf{M}$

$$\mathbb{E}\,|\operatorname{tr}\mathbf{A}_t(z)\mathbf{M}| \leq K||\mathbf{M}||,$$

which follows using formula (9.9.6) in Bai and Silverstein (2010).
Using the decomposition given in (B.14), the estimates (B.16) and (B.17) (which shows that all terms involving $\mathbf{B}_t(z)$ and $\mathbf{C}_t(z)$ are negligible) and the fact

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

we obtain

$$\begin{aligned}
& y_{\lfloor nt\rfloor}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}[\beta_{j,t}(z)d_{j,t}(z)] \\
=& \frac{z^2\underline{\tilde{s}}_t^2(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\Big[\operatorname{tr}\Big(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\Big)^{-1}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-1}\mathbf{T}_n \\
& \times\Big(z\mathbf{I} - \frac{\lfloor nt\rfloor - 1}{n} b_{j,t}(z)\mathbf{T}_n\Big)^{-1}\mathbf{T}_n\Big] \\
& + \frac{z^2\underline{\tilde{s}}_t^2(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor} b_{j,t}^2(z)\mathbb{E}\left[\operatorname{tr}\mathbf{A}_t(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{A}_t(z)\mathbf{T}_n\right] + o(1) \\
=& \frac{\underline{\tilde{s}}_t^2(z)}{n}\mathbb{E}\left[\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-3}\mathbf{T}_n^2\right] \\
& + \frac{z^4\underline{\tilde{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\operatorname{tr}\mathbf{A}_t(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{A}_t(z)\mathbf{T}_n\right] + o(1).
\end{aligned} \tag{B.18}$$

For the term $\mathbf{A}_t(z)$ in (B.15) (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 (B.18), respectively, yields

$$\begin{aligned}
&\frac{z^4\tilde{\underline{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\mathbb{E}\left[\operatorname{tr}\mathbf{A}_t(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{A}_t(z)\mathbf{T}_n\right]\\
=&\frac{z^4\tilde{\underline{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i,l\neq j}\mathbb{E}\Big[\operatorname{tr}\left(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{j,t}(z)\mathbf{T}_n\right)^{-1}\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\mathbf{D}_{i,j,t}^{-1}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\\
&\times\mathbf{D}_{l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star-n^{-1}\mathbf{T}_n\right)\left(z\mathbf{I}-\frac{\lfloor nt\rfloor-1}{n}b_{l,t}(z)\mathbf{T}_n\right)^{-1}\mathbf{T}_n\Big]\\
=&\frac{z^2\tilde{\underline{s}}_t^4(z)}{\lfloor nt\rfloor n}\sum_{j=1}^{\lfloor nt\rfloor}\sum_{i,l\neq j}A_{i,l,j}(z,t)+o(1), \qquad \text{(B.19)}
\end{aligned}$$

where

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

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 (B.19) vanishes asymptotically. For this purpose we use the formulas 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 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 (B.51))

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

which is shown in Lemma B.9 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\} \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).
\end{aligned}
\tag{B.20}
$$

Here we used for the last equality the fact

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

$$\leq K + \mathbb{E}\Big|\beta_{i,j,t}(z)\mathbf{r}_i^\star \mathbf{D}_{i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n + \mathbf{I}\right)^{-1}\mathbf{T}_n\left(\mathbf{D}_{i,j,t}^{-1}(z) - \beta_{i,j,t}(z)\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,j,t}^{-1}(z)\right)\mathbf{T}_n\mathbf{D}_{i,j,t}^{-1}(z)\mathbf{r}_i\Big|$$
$$\leq K.$$

Hence,

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

We now apply (B.13) for (B.20) and obtain

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

This implies (4.28) for the real case, namely,

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

B.6 Further auxiliary results

Lemma B.4. *We have uniformly in $n\in\mathbb{N}, t\in[t_0,1], z\in\mathcal{C}_n$*

$$\mathbb{E}||\mathbf{D}_t^{-1}(z)||^q \leq K, \tag{B.21}$$

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

$$||\mathbf{D}_t^{-1}(z)|| \leq K + n\varepsilon_n^{-1} I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) \leq \eta_{l,t}\}. \tag{B.22}$$

Proof of Lemma B.4. We restrict ourselves to the first assertion. Let first $z\in\mathcal{C}_u$, that is,

$z = x + iv_0$ for some $x \in [x_l, x_r]$. Then,

$$||\mathbf{D}_t^{-1}(z)|| = \frac{1}{\min(|\lambda_{\min}(\mathbf{B}_{n,t}) - z|, |\lambda_{\max}(\mathbf{B}_{n,t}) - z|)} \leq \frac{1}{v_0} = K.$$

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 $m > 0$

$$\mathbb{P}\left(||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) < \eta_{l,t}\right) = o\left(\lfloor nt \rfloor^{-m}\right) = o\left(n^{-m}\right), \tag{B.23}$$

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

$$\liminf_{n\to\infty} \lambda_{\min}(\mathbf{T}_n)(1 - \sqrt{y_t})^2 I_{(0,1)}(y_t) t$$

and x_l. We estimate

$$\begin{aligned}
\mathbb{E}||\mathbf{D}_t^{-1}(z)||^q \leq & K\mathbb{E}\left[||\mathbf{D}_t^{-1}(z)|| I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_{\min}(\mathbf{B}_{n,t}) \geq \eta_{l,t}\}\right]^q \\
& + K\mathbb{E}\left[||\mathbf{D}_t^{-1}(z)|| I\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) < \eta_{l,t}\}\right]^q \\
\leq & K + Kn^q \varepsilon_n^{-q} n^{-m} \leq K.
\end{aligned}$$

To derive a bound for the first term, 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

$$\begin{aligned}
& ||\mathbf{D}_t^{-1}(z)|| I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_{\min}(\mathbf{B}_{n,t}) \geq \eta_{l,t}\} \\
= & \frac{1}{\min(|\lambda_{\min}(\mathbf{B}_{n,t}) - (x_r + iv)|, |\lambda_{\max}(\mathbf{B}_{n,t}) - (x_r + iv)|)} I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_{\min}(\mathbf{B}_{n,t}) \geq \eta_{l,t}\} \\
\leq & \frac{1}{x_r - \lambda_{\max}(\mathbf{B}_{n,t})} I\{||\mathbf{B}_{n,t}|| \leq \eta_{r,t} \text{ and } \lambda_{\min}(\mathbf{B}_{n,t}) \geq \eta_{l,t}\} \\
\leq & \frac{1}{x_r - \eta_{r,t}} \leq \frac{1}{x_r - \limsup\limits_{n\to\infty} ||\mathbf{T}_n||(1 + \sqrt{y_{t_0}})^2} = K.
\end{aligned} \tag{B.24}$$

For the second summand, we conclude

$$\begin{aligned}
& ||\mathbf{D}_t^{-1}(z)|| I\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) < \eta_{l,t}\} \\
\leq & \frac{1}{\min(|\lambda_{\min}(\mathbf{B}_{n,t}) - z|, |\lambda_{\max}(\mathbf{B}_{n,t}) - z|)} I\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) < \eta_{l,t}\} \\
\leq & n\varepsilon_n^{-1} I\left\{||\mathbf{B}_{n,t}|| > \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) < \eta_{l,t}\right\}.
\end{aligned} \tag{B.25}$$

The bounds in (B.24) and (B.25) show that (B.22) holds true. The assertion in (B.21) follows by applying (B.23). □

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.

Lemma B.5. *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 4.3*

$$\mathbb{E}|M_n^1(z_1,t) - M_n^1(z_2,t)|^{2+\delta} \lesssim |z_1 - z_2|^{2+\delta}. \tag{B.26}$$

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

$$\mathbb{E}|Z_n^1(z,t_1,t_2)|^4 \lesssim \left(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\right)^4, \tag{B.27}$$

$$\mathbb{E}|Z_n^2(z,t_1,t_2)|^{4+\delta} \lesssim \left(\frac{\lfloor nt_2 \rfloor - \lfloor nt_1 \rfloor}{n}\right)^{2+\delta/2}, \tag{B.28}$$

where

$$M_n^1(z,t_1) - M_n^1(z,t_2) = Z_n^1(z,t_1,t_2) + Z_n^2(z,t_1,t_2), \tag{B.29}$$

and Z_n^1 and Z_n^2 are defined in (4.22) *and* (4.23)*, respectively.*

The proof of Lemma B.5 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 4.2) and the convergence of the non-random part $(\hat{M}_n^2)_{n\in\mathbb{N}}$ (Theorem 4.4), we need the stronger moment assumption from Theorem 2.1, namely

$$\sup_{i,j,n} \mathbb{E}|x_{ij}|^{12} < \infty, \tag{B.30}$$

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 $q \geq 2$

$$\mathbb{E}|\mathbf{r}_j^\star \mathbf{A} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}|^q \lesssim \begin{cases} (\operatorname{tr} \mathbf{A}\mathbf{A}^\star)^{q/2} n^{-(q\wedge 6)}, \\ ||\mathbf{A}||^q n^{-((q/2)\wedge 5)}. \end{cases} \tag{B.31}$$

Furthermore, combining (B.31) with arguments given in the proof of (9.9.6) in Bai and Silverstein (2010), we obtain the following lemma.

Lemma B.6. *Let $j, m \in \mathbb{N}_0$, $q \geq 2$ and $\mathbf{A}_l$, $l \in \{1, \ldots, m+1\}$ be $p \times p$ (random) matrices independent of $\mathbf{r}_j$ which obey for any $\tilde{q} \geq 2$*

$$\mathbb{E}||\mathbf{A}_l||^{\tilde{q}} < \infty, \ l \in \{1, \ldots, m+1\}.$$

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 $l \in \{1, \ldots, m+1\}$, $\tilde{q} \geq 2$

$$\mathbb{E}\left[\operatorname{tr} \mathbf{A}\mathbf{A}_l^\star\right]^{\tilde{q}} < \infty,$$

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 B.1. In fact, as the proof of Lemma B.6 reveals, we could impose a less restrictive condition on the spectral moments of $\mathbf{A}_l$, $l \in \{1, \ldots, m+1\}$. For our purpose, it is sufficient to state the previous lemma in this form, since, when applying Lemma B.6, 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 $\mathbf{D}_{j,t}^{-1}(z)$, while we will make use of the first assertion in 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 B.6. For $m = 0$, the assertion of the lemma follows directly from (B.31) for any $q \geq 2$. We continue the proof by an induction over the integer m for some fixed $q \geq 2$.

$$\begin{aligned}
&\mathbb{E}\Big|\Big(\prod_{k=1}^{m} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big) \left(\mathbf{r}_j^\star \mathbf{A}_{m+1} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m+1}\right) \Big|^q \\
\lesssim &\mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big) \left(\mathbf{r}_j^\star \mathbf{A}_{m} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m}\right) \left(\mathbf{r}_j^\star \mathbf{A}_{m+1} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m+1}\right) \Big|^q \\
&+ \mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big) n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m} \left(\mathbf{r}_j^\star \mathbf{A}_{m+1} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m+1}\right) \Big|^q \\
\leq &\Big(\mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big) \left(\mathbf{r}_j^\star \mathbf{A}_{m} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m}\right) \Big|^{2q} \mathbb{E}\left|\left(\mathbf{r}_j^\star \mathbf{A}_{m+1} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m+1}\right)\right|^{2q} \Big)^{\frac{1}{2}}
\end{aligned}$$

$$+ \mathbb{E}\Big|\Big(\prod_{k=1}^{m-1} \mathbf{r}_j^\star \mathbf{A}_k \mathbf{r}_j\Big) n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_m \left(\mathbf{r}_j^\star \mathbf{A}_{m+1} \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_{m+1}\right) \Big|^q.$$

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 (B.30) for $q \geq 2$

$$\mathbb{E}|\gamma_{j,t}(z)|^q \lesssim n^{-((q/2)\wedge 5)}. \tag{B.32}$$

We need an estimate for moments of complex martingale difference schemes. We refer to Lemma 2.1 in Li (2003), which is an 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 B.5.

Proof of Lemma B.5. In the following, we will often make use of the decompositions

$$\begin{aligned} \mathbf{D}_t^{-1}(z) &= \mathbf{D}_{j,t}^{-1}(z) - \beta_{j,t}(z)\mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t}^{-1}(z), \\ \beta_{j,t}(z) &= b_{j,t}(z) - \beta_{j,t}(z)b_{j,t}(z)\gamma_{j,t}(z). \end{aligned} \tag{B.33}$$

Observing the decomposition (B.29), our aim is to show the inequalities in (B.27) and (B.28), where we assume $t_2 > t_1$ w.l.o.g.

Step 1: *Analysis of* Z_n^2

Beginning with the proof of (B.28) 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

$$\begin{aligned} \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\right|^{4+\delta} \lesssim& \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\right|^{4+\delta} \\ &+ \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}(z)b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\gamma_{j,t_2}(z)\right|^{4+\delta} \\ \lesssim& \mathbb{E}\left|(\mathbb{E}_j - \mathbb{E}_{j-1})\left\{\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{j,t_2}^{-2}(z)\right\}\right|^{4+\delta} \\ &+ \mathbb{E}\left|\beta_{j,t_2}(z)b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\gamma_{j,t_2}(z)\right|^{4+\delta} \end{aligned} \tag{B.34}$$

$$\lesssim n^{-(2+\delta/2)}.$$

We should explain the bound for (B.34) 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 B.4). As a further preparation, we observe for $z \in \mathcal{C}_n, t \in [t_0, 1]$ from Lemma B.4

$$\begin{aligned} ||\mathbf{D}_t^{-1}(z)|| \lesssim & 1 + n\varepsilon_n^{-1} I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) \leq \eta_{l,t}\} \\ \leq & 1 + n^2 I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) \leq \eta_{l,t}\}, \end{aligned} \tag{B.35}$$

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

$$\begin{aligned} |\beta_{j,t}(z)| = & |1 - \mathbf{r}_j^\star \mathbf{D}_t^{-1}(z)\mathbf{r}_j| \leq 1 + |\mathbf{r}_j|^2 ||\mathbf{D}_t^{-1}(z)|| \\ \lesssim & 1 + |\mathbf{r}_j|^2 + n^3 I\{||\mathbf{B}_{n,t}|| \geq \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}) \leq \eta_{l,t}\}. \end{aligned} \tag{B.36}$$

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

$$|b_{j,t}(z)| \leq K. \tag{B.37}$$

Similarly to these bounds, using (B.22) in Lemma B.4 for the matrix $\mathbf{D}_{j,t}^{-1}(z)$, we get for any $m \geq 1$

$$\begin{aligned} |\gamma_{j,t}(z)| = & |\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z)\mathbf{r}_j - n^{-1}\mathbb{E}[\operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z)]| \lesssim |\mathbf{r}_j|^2 ||\mathbf{D}_{j,t}^{-1}(z)|| + \mathbb{E}||\mathbf{D}_{j,t}^{-1}(z)|| \\ \lesssim & |\mathbf{r}_j|^2 + |\mathbf{r}_j|^2 n\varepsilon_n^{-1} I\{||\mathbf{B}_{n,t}^{(-j)}|| \geq \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}^{(-j)}) \leq \eta_{l,t}\} \\ & + |\mathbf{r}_j|^2 n\varepsilon_n^{-1} \mathbb{P}\{||\mathbf{B}_{n,t}^{(-j)}|| \geq \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}^{(-j)}) \leq \eta_{l,t}\} \\ \leq & |\mathbf{r}_j|^2 + n^3 I\{||\mathbf{B}_{n,t}^{(-j)}|| \geq \eta_{r,t} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t}^{(-j)}) \leq \eta_{l,t}\} + o\left(n^{-m}\right), \end{aligned}$$

where we used the fact that for any $m > 0$

$$\begin{aligned} &\mathbb{P}\{||\mathbf{B}_{n,t_2}^{(-j)}|| \geq \eta_{r,t_2} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t_2}^{(-j)}) \leq \eta_{l,t_2}\} = o\left(n^{-m}\right), \\ &\mathbb{P}\{||\mathbf{B}_{n,t_2}|| \geq \eta_{r,t_2} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t_2}) \leq \eta_{l,t_2}\} = o\left(n^{-m}\right) \end{aligned} \tag{B.38}$$

and the notation

$$\mathbf{B}_{n,t}^{(-j)} = \mathbf{B}_{n,t} - \mathbf{r}_j \mathbf{r}_j^\star.$$

Using (B.31) and (B.32), 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 $\delta \geq 0$

$$\mathbb{E}\Big||\mathbf{r}_j|^{2q}\gamma_{j,t}(z)\Big|^{4+\delta} \lesssim n^{-(2+\delta/2)}. \tag{B.39}$$

Combining these inequalities, we conclude

$$\begin{aligned}
&\mathbb{E}\left|\beta_{j,t_2}(z)b_{j,t_2}(z)\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}(z)\mathbf{r}_j\gamma_{j,t_2}(z)\right|^{4+\delta} \\
\lesssim&\mathbb{E}\Big|\Big(1+|\mathbf{r}_j|^2+n^3 I\{||\mathbf{B}_{n,t_2}|| \geq \eta_{r,t_2} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t_2}) \leq \eta_{l,t_2}\}\Big)|\mathbf{r}_j|^2 \\
&\times\Big(1+n^2 I\{||\mathbf{B}_{n,t_2}^{(-j)}|| \geq \eta_{r,t_2} \text{ or } \lambda_{\min}(\mathbf{B}_{n,t_2}^{(-j)}) \leq \eta_{l,t_2}\}\Big)^2\gamma_{j,t_2}(z)\Big|^{4+\delta}.
\end{aligned} \tag{B.40}$$

The expectation in (B.40) can now be estimated by multiplying these terms out and using the inequalities (B.37) and (B.38).
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 (B.29), we show that (B.26) 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 (B.28). 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 (B.34) 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 (B.37)

$$\begin{aligned}
& \mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) b_{j,t}(z_1) b_{j,t}(z_2) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \right)^2 \right|^{2+\delta} \\
\lesssim & \mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \right)^2 \right|^{2+\delta} \\
= & \mathbb{E} \left| (\mathbb{E}_j - \mathbb{E}_{j-1}) \left[\left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j \right)^2 - \left(n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right)^2 \right] \right|^{2+\delta} \\
= & \mathbb{E} \Big| (\mathbb{E}_j - \mathbb{E}_{j-1}) \Big[\left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right) \\
& \times \left(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \mathbf{r}_j + n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{D}_{j,t}^{-1}(z_1) \mathbf{D}_{j,t}^{-1}(z_2) \right) \Big] \Big|^{2+\delta}
\end{aligned}$$

$$\lesssim \mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\Big[\big(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\big)\, \mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j\Big]\Big|^{2+\delta}$$
$$+\mathbb{E}\Big|(\mathbb{E}_j - \mathbb{E}_{j-1})\Big[\big(\mathbf{r}_j^\star \mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\big)\, n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{j,t}^{-1}(z_1)\mathbf{D}_{j,t}^{-1}(z_2)\Big]\Big|^{2+\delta}$$
$$\lesssim n^{-(1+\delta/2)},$$

where we used Lemma B.6 with $q = 2+\delta$ and $m = 1$ and Lemma B.4 for the last inequality. These considerations show that (B.26) holds true.

Step 3: *Analysis of* Z_n^1

Next, we show the estimate (B.27) for the term Z_n^1. Doing so, we will need condition (B.30) 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 $j \le \lfloor nt_1 \rfloor$

$$\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\\
&-\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\}.
\end{aligned}$$

Hence, using the identity (B.33), we obtain the representation

$$\begin{aligned}
&Z_n^1(z,t_1,t_2)\\
=&\sum_{j=1}^{\lfloor nt_1\rfloor}\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j-\mathbb{E}_{j-1})\Big\{-\beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j - \beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j\\
&+\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j - \beta_{j,t_1}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j\\
&-\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j + \beta_{j,t_1}\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_1}^{-1}\mathbf{r}_j\mathbf{r}_j^\star\mathbf{D}_{j,t_2}^{-2}\mathbf{r}_j\Big\}.
\end{aligned}$$

We use the substitutions

$$\mathbf{D}_{j,t_2}^{-1} = \mathbf{D}_{k,j,t_2}^{-1} - \beta_{k,j,t_2}\mathbf{D}_{k,j,t_2}^{-1}\mathbf{r}_k\mathbf{r}_k^\star\mathbf{D}_{k,j,t_2}^{-1} \tag{B.41}$$

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 $\gamma_{k,j,t}(z) = \mathbf{r}_k^\star\mathbf{D}_{k,j,t}^{-1}(z)\mathbf{r}_k - n^{-1}\mathbb{E}[\operatorname{tr}\mathbf{T}_n\mathbf{D}_{k,j,t}^{-1}(z)]$. 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, $q\in\mathbb{N}$, $q_{l_1}\in\mathbb{N}_0$, $l_1\in\{1,\ldots,q\}$, there exists an index $l_1\in\{1,\ldots,q\}$ 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 $l_2\in\{1,\ldots,q_{l_1}\}$, $l_1\in\{1,\ldots,q\}$ 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),

$$\mathbb{E}|Z_n^1(z,t_1,t_2)|^4 = \mathbb{E}\Big|\sum\sum_{j=1}^{\lfloor nt_1\rfloor}\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j-\mathbb{E}_{j-1})T_{j,k}\Big|^4 \lesssim n\sum\sum_{j=1}^{\lfloor nt_1\rfloor}\mathbb{E}\Big|\sum_{k=\lfloor nt_1\rfloor+1}^{\lfloor nt_2\rfloor}(\mathbb{E}_j-\mathbb{E}_{j-1})T_{j,k}\Big|^4,$$

in order to prove (B.27), it suffices to show that for $j\in\{1,\ldots,\lfloor nt_1\rfloor\}$ and $k\in\{\lfloor nt_1\rfloor+1,\ldots,\lfloor nt_2\rfloor\}$

$$\mathbb{E}\left|(\mathbb{E}_j-\mathbb{E}_{j-1})T_{j,k}\right|^4 \lesssim n^{-6}.$$

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

$$R_1 = \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j \gamma_{j,t_1}|^4, \tag{B.42}$$

$$R_2 = \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j|^4, \tag{B.43}$$

$$R_3 = \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \gamma_{j,t_2}|^4, \tag{B.44}$$

$$R_4 = \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{k,j,t_2}\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j \gamma_{k,j,t_2}|^4, \tag{B.45}$$

$$R_5 = \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{k,j,t_2}\beta_{j,t_2}\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j \gamma_{k,j,t_2}\gamma_{j,t_2}|^4. \tag{B.46}$$

For further investigations, we observe that

$$(\mathbb{E}_j - \mathbb{E}_{j-1}) \operatorname{tr}\Big(\prod_{l=1}^{q_1} \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_l\Big) = 0, \tag{B.47}$$

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 (B.42), 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 (B.34), in particular the bound in (B.39),

$$\begin{aligned} R_1 \lesssim & \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_2 \left(\mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 - n^{-1}\mathbf{T}_n \mathbf{A}_3\right) \mathbf{r}_j \gamma_{j,t_1}|^4 \\ & + n^{-4}\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_2 \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_j \gamma_{j,t_1}|^4 \\ = & n^{-4}\mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_1}\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_2 \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_j \gamma_{j,t_1}|^4 \lesssim n^{-6}. \end{aligned}$$

For (B.43), we have using Lemma B.6 and (B.47)

$$\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 \end{aligned}$$

$$
\begin{aligned}
&\lesssim n^{-4}\mathbb{E}\left(\operatorname{tr}\left(\mathbf{A}_1\mathbf{r}_k\left(\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k-n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_2\right)\mathbf{r}_k^\star\mathbf{A}_3\right)\left(\mathbf{A}_1\mathbf{r}_k\left(\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k-n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_2\right)\mathbf{r}_k^\star\mathbf{A}_3\right)^\star\right)^2+n^{-6}\\
&=n^{-4}\mathbb{E}\left(\operatorname{tr}\left(\mathbf{A}_1\mathbf{r}_k\left(\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k-n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_2\right)\mathbf{r}_k^\star\mathbf{A}_3\right)\left(\mathbf{A}_3^\star\mathbf{r}_k\overline{(\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k-n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_2)}\mathbf{r}_k^\star\mathbf{A}_1^\star\right)\right)^2+n^{-6}\\
&=n^{-4}\mathbb{E}|\left(\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k-n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_2\right)^2\mathbf{r}_k^\star\mathbf{A}_3\mathbf{A}_3^\star\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_1^\star\mathbf{A}_1\mathbf{r}_k|^2+n^{-6}\\
&\lesssim n^{-6}.
\end{aligned}
$$

Next, we have for the term R_4 defined in (B.45) by similar arguments as in the derivation of (B.34)

$$
\begin{aligned}
R_4 =&\mathbb{E}|(\mathbb{E}_j-\mathbb{E}_{j-1})\beta_{k,j,t_2}\{\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j-n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\}\gamma_{k,j,t_2}|^4\\
\lesssim& n^{-4}\mathbb{E}\Big[\left(\operatorname{tr}\left(\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)\left(\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)^\star\right)^2|\beta_{k,j,t_2}\gamma_{k,j,t_2}|^4\Big]\\
\lesssim& n^{-6},
\end{aligned}
$$

where we used the bound in Lemma B.6 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 (B.44), we first decompose using (B.41)

$$
\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)\\
=&\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 (B.44), 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

$$\begin{aligned}
R_{31} =& \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\gamma_{j,k,t_2}|^4, \qquad \text{(B.48)} \\
R_{32} =& \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\left(\mathbf{r}_j^\star\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)|^4, \\
R_{33} =& \mathbb{E}|(\mathbb{E}_j - \mathbb{E}_{j-1})\beta_{j,t_2}\beta_{k,j,t_2}\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\left(\mathbf{r}_j^\star\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\mathbf{r}_j - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_3\mathbf{r}_k\mathbf{r}_k^\star\mathbf{A}_3\right)\gamma_{k,j,t_2}|^4 \qquad \text{(B.49)}
\end{aligned}$$

and we used an analogue of the estimate (B.39) for the terms γ_{k,j,t_2} and $\tilde{\gamma}_{k,j,t_2}$ in the last step. The term R_{31} in (B.48) can be bounded using the bounds in (B.35), (B.36) and (B.38) 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 B.6

$$\begin{aligned}
R_{311} \lesssim & \mathbb{E}|\left(\mathbf{r}_k^\star\mathbf{A}_2\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_1\mathbf{r}_k - n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{A}_2\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_1\right)\gamma_{j,k,t_2}|^4 + n^{-4}\mathbb{E}|\mathbf{r}_j^\star\mathbf{A}_1\mathbf{T}_n\mathbf{A}_2\mathbf{r}_j\gamma_{j,k,t_2}|^4 \\
\lesssim & n^{-6} + n^{-4}\mathbb{E}\left[|\gamma_{j,k,t_2}|^4\left(\operatorname{tr}\left(\mathbf{A}_2\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_1\right)\left(\mathbf{A}_2\mathbf{r}_j\mathbf{r}_j^\star\mathbf{A}_1\right)^\star\right)^2\right] \\
\leq & n^{-6}.
\end{aligned}$$

Note that the term R_{33} defined in (B.49) can be bounded similarly. Similarly to R_{31} given in (B.48) we bound $|\beta_{j,t_2}|$ and get

$$R_{32} \lesssim R_{321} + R_{322} + o\left(n^{-l}\right),$$

where

$$R_{321} = \mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \right)|^4,$$
$$R_{322} = \mathbb{E}|\mathbf{r}_j^\star \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \right)|^4.$$

For the sake of brevity, we shall limit ourselves to investigating the summand R_{321}.

$$\begin{aligned} R_{321} \lesssim & \mathbb{E}| \left(\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \right) \left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \right)|^4 \\ & + n^{-4} \mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \left(\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \right)|^4 \\ \lesssim & n^{-6} + \left(\mathbb{E}|\mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_2 \mathbf{r}_j \mathbf{r}_j^\star \mathbf{A}_1|^8 \mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_3 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3|^8 \right)^{\frac{1}{2}} \\ \lesssim & n^{-6}. \end{aligned}$$

Finally, invoking Lemma B.6 and (B.39), we can show for the term R_5 defined in (B.46) that

$$R_5 \lesssim R_{51} + R_{52},$$

where

$$\begin{aligned} R_{51} = & \mathbb{E}|\beta_{k,j,t_2} \beta_{j,t_2} \left(\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \right) \gamma_{k,j,t_2} \gamma_{j,t_2}|^4 \\ \le & \left(\mathbb{E}|\mathbf{r}_j^\star \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{r}_j - n^{-1} \operatorname{tr} \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3|^8 \mathbb{E}|\beta_{k,j,t_2} \beta_{j,t_2} \gamma_{k,j,t_2} \gamma_{j,t_2}|^8 \right)^{\frac{1}{2}} \lesssim n^{-6}, \\ R_{52} = & n^{-4} \mathbb{E}|\beta_{k,j,t_2} \beta_{j,t_2} \mathbf{r}_k^\star \mathbf{A}_2 \mathbf{r}_k \mathbf{r}_k^\star \mathbf{A}_3 \mathbf{T}_n \mathbf{A}_1 \mathbf{r}_k \gamma_{k,j,t_2} \gamma_{j,t_2}|^4 \lesssim n^{-6}. \end{aligned}$$

Thus, the moment inequalities (B.26), (B.27) and (B.28) for M_n^1 hold true. □

For the proof of Lemma 4.2, we need the following identity, which can be proved similarly to (5.2) in Bai and Silverstein (1998).

Lemma B.7.

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

Lemma B.8. *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 B.8. 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

$$\lim_{n\to\infty} \tilde{\underline{s}}_{t_n}(z_n) = 0.$$

By choosing approriate subsequences, we assume without loss of generality that $(z_n)_{n\in\mathbb{N}}$ converges to limit in the closure of S and $(t_n)_{n\in\mathbb{N}}$ converges to a limit in $[t_0,1]$. From (2.5), we conclude

$$\lim_{n\to\infty} y \int \frac{\lambda \tilde{\underline{s}}_{t_n}(z_n)}{1+\lambda t_n \tilde{\underline{s}}_{t_n}(z_n)} dH(\lambda) = 1.$$

But, using the fact that H is compactly supported, we see that the expression above tends to 0. Thus, we get a contradiction. □

Lemma B.9. *In the real case, it holds for $i \neq l$ $(i,l \in \{1,\ldots,n\}\setminus\{j\})$*

$$\begin{aligned}
&\sup_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} \Big| \mathbb{E}\Big[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\beta_{l,i,j,t}(z)\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n \\
&\times \beta_{i,l,j,t}(z)\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\Big]\Big| \qquad \text{(B.50)}\\
=&o\left(n^{-1}\right).
\end{aligned}$$

Proof of Lemma B.9. 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

$$\beta_{j,t}(z) = \overline{\beta}_{j,t}(z) - \overline{\beta}^2_{j,t}(z)\hat{\gamma}_{j,t}(z) + \overline{\beta}^2_{j,t}(z)\beta_{j,t}(z)\hat{\gamma}^2_{j,t}(z), \qquad \text{(B.51)}$$

in order to replace $\beta_{l,i,j,t}(z)$ and $\beta_{i,l,j,t}(z)$. Note that $\mathbb{E}[||\mathbf{D}^{-1}_{l,i,j,t}(z)||] \leq K$ and $||\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}|| \leq K$ which follows from Lemma B.4 and Proposition B.1. By applying the triangle inequality, this gives us several summands for the mean in (B.50). More precisely, we can write

$$\begin{aligned}
&\Big| \mathbb{E}\Big[\operatorname{tr}\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\beta_{l,i,j,t}(z)\mathbf{D}^{-1}_{l,i,j,t}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}^{-1}_{l,i,j,t}(z)\left(t\tilde{\underline{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n \\
&\times \beta_{i,l,j,t}(z)\mathbf{D}^{-1}_{i,l,j,t}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}^{-1}_{i,l,j,t}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\Big]\Big| \\
\leq& \sum_{\zeta_1,\zeta_2} |\mathbb{E}[T(\zeta_1,\zeta_2)]|,
\end{aligned}$$

where $T(\zeta_1,\zeta_2)$ has the following form

$$\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star-n^{-1}\mathbf{T}_n\right)\zeta_1\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n$$
$$\times\zeta_2\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star-n^{-1}\mathbf{T}_n\right),$$

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

$$|\mathbb{E}[T(\zeta_1,\zeta_2)]|=o\left(n^{-1}\right). \tag{B.52}$$

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\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}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,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}_{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|\\
\le&K\Big|\mathbb{E}\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\Big|\\
&+Kn^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\Big|\\
&+Kn^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{T}_n\Big|\\
&+Kn^{-2}\Big|\mathbb{E}\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{T}_n\Big|\\
=&K(T_1+T_2+T_3)+o\left(n^{-1}\right),
\end{aligned}$$

where

$$\begin{aligned}
T_1=&\Big|\mathbb{E}\left[\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\right]\Big|,\\
T_2=&n^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n^2\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\Big|,\\
T_3=&n^{-1}\Big|\mathbb{E}\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{T}_n\Big|.
\end{aligned}$$

For the first summand, we obtain using (9.8.6) in Bai and Silverstein (2010) for the real case

$$\begin{aligned}
T_1\le&\Big|\mathbb{E}\Big[\big\{\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\\
&-n^{-1}\operatorname{tr}\mathbf{T}_n\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\big\}
\end{aligned}$$

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

$$= o\left(n^{-1}\right).$$

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 (B.52) holds true in the case $\zeta_1 = \overline{\beta}_{l,i,j,t}(z)$ and $\zeta_2 = -\overline{\beta}_{i,l,j,t}^2(z)\hat{\gamma}_{i,l,j,t}(z)$.
Finally, we consider the case $\zeta_1 = -\overline{\beta}_{l,i,j,t}^2(z)\hat{\gamma}_{l,i,j,t}(z)$ and $\zeta_2 = -\overline{\beta}_{i,l,j,t}^2(z)\hat{\gamma}_{i,l,j,t}(z)$. Note that $\overline{\beta}_{i,l,j,t}(z) = \overline{\beta}_{l,i,j,t}(z)$. We obtain (B.52), that is,

$$\begin{aligned}
&|\mathbb{E}[T(\zeta_1,\zeta_2)]| \\
=&\Big|\mathbb{E}\Big[\overline{\beta}_{i,j,l,t}^4(z)\operatorname{tr}\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\hat{\gamma}_{l,i,j,t}(z) \\
&\times \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\hat{\gamma}_{i,l,j,t}(z)\Big]\Big| \\
\le&\mathbb{E}^{\frac{1}{2}}|E_1|^2\mathbb{E}^{\frac{1}{4}}|E_2|^4\mathbb{E}^{\frac{1}{4}}|E_3|^4 = o\left(n^{-1}\right),
\end{aligned}$$

if

$$\mathbb{E}^{\frac{1}{2}}|E_1|^2 \le Kn^{-1}, \tag{B.53}$$

$$\mathbb{E}^{\frac{1}{4}}|E_2|^4 \le K, \tag{B.54}$$

$$\mathbb{E}^{\frac{1}{4}}|E_3|^4 = o(1), \tag{B.55}$$

where

$$\begin{aligned}
E_1 =&\overline{\beta}_{i,j,l,t}^4(z)\hat{\gamma}_{i,l,j,t}(z)\hat{\gamma}_{l,i,j,t}(z), \\
E_2 =&\mathbf{r}_l^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-1}\mathbf{T}_n\mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_i, \\
E_3 =&\operatorname{tr}\left\{\left(t\underline{\tilde{s}}_t(z)\mathbf{T}_n+\mathbf{I}\right)^{-2}\mathbf{T}_n\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}_n\right)\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\right\}.
\end{aligned}$$

We begin with a proof of (B.53). 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

$$\mathbb{E}|E_1|^2 \le Kn^{-2},$$

which proves (B.53). The estimate (B.54) can be proven similarly to Bai and Silverstein (2010), p. 290.
Finally, we will prove that (B.55) holds true. We obtain

$$\begin{aligned}
E_3 =&\mathbf{r}_i^\star\mathbf{D}_{i,l,j,t}^{-1}(z)\left(\mathbf{r}_l\mathbf{r}_l^\star - n^{-1}\mathbf{T}_n\right)\left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\left(\mathbf{r}_i\mathbf{r}_i^\star - n^{-1}\mathbf{T}\right)\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l \\
=&E_{31}+E_{32}+E_{33}+E_{34},
\end{aligned}$$

where

$$\begin{aligned}
E_{31} =& \mathbf{r}_i^\star \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l \\
E_{32} =& -n^{-1}\mathbf{r}_i^\star \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{r}_l\mathbf{r}_l^\star \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}^2\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l, \\
E_{33} =& -n^{-1}\mathbf{r}_i^\star \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{T}_n \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}\mathbf{r}_i\mathbf{r}_i^\star\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l, \\
E_{34} =& n^{-2}\mathbf{r}_i^\star \mathbf{D}_{i,l,j,t}^{-1}(z)\mathbf{T}_n \left(t\underline{\tilde{s}}_t(z)\mathbf{T}+\mathbf{I}\right)^{-2}\mathbf{T}^2\mathbf{D}_{l,i,j,t}^{-1}(z)\mathbf{r}_l.
\end{aligned}$$

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 (B.55). Hence, the proof of Lemma B.9 is finished. □

Lemma B.10. *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 B.10. 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 B.8 that $\operatorname{Re}(\underline{\tilde{s}}_t(z))$ is uniformly bounded away from 0. Thus, due to Theorem 4.5, 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 $K_1 > 0$

$$\operatorname{Re}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)^{-1} = \frac{\operatorname{Re}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)}{|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2} \geq K_1\operatorname{Im}^2(z).$$

For the imaginary part, we conclude for some $K_2 > 0$

$$\begin{aligned}
\operatorname{Im}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)^{-1} =& \frac{-\operatorname{Im}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)}{|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2} = \frac{1}{|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2} \operatorname{Im}\Big(\int \frac{-1}{\lambda - z} dF^{\underline{\mathbf{B}}_{n,t}}(\lambda)\Big) \\
=& \frac{1}{|\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]|^2} \int \frac{-\operatorname{Im}(z)}{|z-\lambda|^2} dF^{\underline{\mathbf{B}}_{n,t}}(\lambda) \geq K \mathbb{E} \int \frac{-\operatorname{Im}(z)}{|\lambda - z|^2} dF^{\underline{\mathbf{B}}_{n,t}}(\lambda) \\
\geq & - K_2 \operatorname{Im}(z).
\end{aligned}$$

By definition of $R_{n,t}(z)$, we have for all $n \geq N$

$$\begin{aligned}
\operatorname{Im}(R_{n,t}(z)) =& y_{\lfloor nt \rfloor} \lfloor nt \rfloor^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \operatorname{Im}\Big(\mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)] \left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)^{-1}\Big) \\
=& y_{\lfloor nt \rfloor} \lfloor nt \rfloor^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \operatorname{Im}\left(\mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\right) \operatorname{Re}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)^{-1} \\
& + y_{\lfloor nt \rfloor} \lfloor nt \rfloor^{-1} \sum_{j=1}^{\lfloor nt \rfloor} \operatorname{Re}\left(\mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\right) \operatorname{Im}\left(\mathbb{E}[\underline{\tilde{s}}_{n,t}(z)]\right)^{-1} \\
\geq & K_1 \operatorname{Im}^2(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big) \\
& - K_2 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Re}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big),
\end{aligned}$$

which implies that

$$\begin{aligned}
& \operatorname{Im}(R_{n,t}(z)) + \operatorname{Im}(z) \\
\geq & \operatorname{Im}(z) + K_1 \operatorname{Im}^2(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big) - K_2 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Re}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big) \\
\geq & \operatorname{Im}(z) \Big\{ 1 + K_1 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big) - K_2 \lfloor nt \rfloor^{-1} \operatorname{Re}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big) \Big\}.
\end{aligned}$$

The real and imaginary part of $\beta_{j,t}(z) d_{j,t}(z)$ might be negative, but, due to (4.29), we have for some $N \in \mathbb{N}$

$$\sup_{n \geq N} \sup_{\substack{z \in \mathcal{C}_n, \\ t \in [t_0, 1]}} \Big| K_1 \operatorname{Im}(z) \lfloor nt \rfloor^{-1} \operatorname{Im}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big) - K_2 \lfloor nt \rfloor^{-1} \operatorname{Re}\Big(y_{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{E}[\beta_{j,t}(z) d_{j,t}(z)]\Big) \Big| < 1$$

Thus, we conclude that

$$\inf_{n\geq N} \inf_{\substack{z\in\mathcal{C}_n,\\ t\in[t_0,1]}} \left(\mathrm{Im}(z) + \mathrm{Im}(R_{n,t}(z))\right) \geq 0.$$

□

C Details for the proof of Theorem 3.1

C.1 How to calculate mean and covariance in Theorem 2.1

The following result provides essential formulas for the calculation of the mean and covariance structure in Theorem 2.1 in the case $\mathbf{T}_n = \mathbf{I}$. It generalizes the formulas given in Proposition A.1 in Wang and Yao (2013) and Proposition 3.6 in Yao et al. (2015).

Proposition C.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* (2.9). *For the random variable $\big(X(f_1,t_1), X(f_2,t)\big)_{t\in[t_0,1]}$ given in Theorem 2.1, we have the following formulas*

$$\begin{aligned}
\mathbb{E}[X(f_i,t)] =& \frac{1}{2\pi i} \lim_{r\searrow 1} \oint\limits_{|\xi|=1} f(t(1+h_t r\xi + h_t r^{-1}\xi^{-1} + h_t^2))\Big(\frac{\xi}{\xi^2 - r^{-2}} - \frac{1}{\xi}\Big) d\xi,\\
\operatorname{cov}(X(f_1,t_1), X(f_2,t_2)) =& \frac{1}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint\limits_{|\xi_1|=1} \oint\limits_{|\xi_2|=1} f_1(t_1(1+h_{t_1}r_1\xi_1 + h_{t_1}r_1^{-1}\xi_1^{-1} + h_{t_1}^2))\\
&\times \overline{f_2(t_2(1+h_{t_2}r_2\xi_2^{-1} + h_{t_2}r_2^{-1}\xi_2 + h_{t_2}^2))} \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1
\end{aligned}$$

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

$$\begin{aligned}
g_1(\xi_1,\xi_2) =& -\Big(h_1h_2r_1r_2\Big\{h_2^4r_1^2r_2^2t_2^2\xi_1^2\xi_2^2 + 2h_2^3r_1^2r_2t_2\xi_1^2\xi_2(r_2^2t_1 + t_2\xi_2^2)\\
&- 2h_1h_2r_1r_2t_1\xi_1\xi_2(r_2^2t_1(2+h_1r_1\xi_1) + r_1t_2\xi_1(h_1+2r_1\xi_1)\xi_2^2\Big\}\\
&+ h_1^2t_1\xi_2^2\Big\{r_2^2t_1(1+2h_1r_1\xi_1 + 3r_1^2\xi_1^2 + h_1^2r_1^2\xi_1^2 + 2h_1r_1^3\xi_1^3) + r_1^2t_2\xi_1^2\ (-1+r_1^2\xi_1^2)\xi_2^2\Big\}\\
&+ h_2^2\Big\{r_2^4t_1^2 - r_2^2t_1t_2(1+2h_1r_1\xi_1 - 3r_1^2\xi_1^2 + 2h_1^2r_1^2\xi_1^2 + 2h_1r_1^3\xi_1^3)\xi_2^2 + r_1^2t_2^2\xi_1^2\xi_2^4\Big\}\Big),\\
g_2(\xi_1,\xi_2) =& \big(h_2r_2 - h_1r_1\xi_1\xi_2\big)^2\big(h_2^2r_1r_2t_2\xi_1\xi_2 - h_1r_2t_1(1+h_1r_1\xi_1 + r_1^2\xi_1^2)\xi_2 + h_2r_1\xi_1(r_2^2t_1 + t_2\xi_2^2)\big)^2.
\end{aligned}$$

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.*

Proof of Proposition C.1. If suffices to consider the real case. Since $H = \delta_{\{1\}}$, we obtain from

Theorem 2.1 for $i \in \{1,2\}$

$$\mathbb{E}[X(f_i,t)] = -\frac{1}{2\pi i}\oint_{\mathcal{C}} f_i(z)\frac{ty\frac{\underline{\tilde{s}}_t^3(z)}{(t\underline{\tilde{s}}_t(z)+1)^3}}{\left(1-ty\frac{\underline{\tilde{s}}_t^2(z)}{(t\underline{\tilde{s}}_t(z)+1)^2}\right)^2}dz$$

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

where the contours $\mathcal{C},\mathcal{C}_1,\mathcal{C}_2$ enclose the interval given in (2.9) 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 (2.6) 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*

For calculation of the mean, we use a change of variables

$$z(\xi)=z=t(1+h_t r\xi+h_t r^{-1}\xi^{-1}+h_t^2),$$

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 (2.5), we have for $z\in\mathcal{C}$

$$\underline{\tilde{s}}_t(z) = -\frac{1}{t(1+h_t r\xi)}\ ,\ \frac{\underline{\tilde{s}}_t(z)}{t\underline{\tilde{s}}_t(z)+1} = -\frac{1}{th_t r\xi}\ ,\ dz = th_t(r-r^{-1}\xi^{-2})d\xi.$$

Thus, we can write for $i\in\{1,2\}$

$$\begin{aligned}\mathbb{E}[X(f_i,t)] &= \lim_{r\searrow 1}\frac{1}{2\pi i}\oint_{|\xi|=1} f_i(z(\xi))t^2h_t^2\frac{\left(\frac{1}{th_t r\xi}\right)^3}{\left(1-t^2h_t^2\left(\frac{1}{th_t r\xi}\right)^2\right)^2}th_t\left(r-r^{-1}\xi^{-2}\right)d\xi\\ &= \lim_{r\searrow 1}\frac{t}{2\pi i}\oint_{|\xi|=1} f_i(z(\xi))\frac{r^{-2}}{\xi t(\xi^2-r^{-2})}d\xi\\ &= -\lim_{r\searrow 1}\frac{t}{2\pi i}\oint_{|\xi|=1} f_i\left(z(\xi)\right)\left(\frac{1}{\xi t}-\frac{\xi}{t(\xi^2-r^{-2})}\right)d\xi\end{aligned}$$

$$= \lim_{r\searrow 1} \frac{1}{2\pi i} \oint_{|\xi|=1} f_i\left(z(\xi)\right)\Big(\frac{\xi}{(\xi^2 - r^{-2})} - \frac{1}{\xi}\Big)d\xi.$$

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 $t_2 \leq t_1$

$$\begin{aligned}
\operatorname{cov}(X(f_1,t_1),X(f_2,t_2)) = \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \frac{1}{2\pi^2} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} f_1(z_1(\xi_1))\overline{f_2(z_2(\xi_2))} \\
\times \sigma^2_{t_1,t_2}(z_1(\xi_1), \overline{z_2(\xi_2)}) t_1 h_{t_1}(r_1 - r_1^{-1}\xi_1^{-2}) t_2 h_{t_2}(r_2 - r_2^{-1}\xi_2^{-2})\overline{d\xi_2} d\xi_1 \\
= \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \frac{1}{2\pi^2} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} f_1(z_1(\xi_1))\overline{f_2(z_2(\xi_2^{-1}))} \\
\times \sigma^2_{t_1,t_2}(z_1(\xi_1), \overline{z_2(\xi_2^{-1})}) t_1 h_{t_1}(r_1 - r_1^{-1}\xi_1^{-2}) t_2 h_{t_2}(r_2 - r_2^{-1}\xi_2^{2}) d\xi_2 d\xi_1.
\end{aligned}$$

By proceeding similarly as for the mean and additionally using

$$\underline{\tilde{s}}_t(z)dz = \frac{h_t r}{t(1 + h_t r \xi)^2} d\xi,$$

we get by straightforward but tedious algebra the desired formula for the covariance (we partially used a computer algebra system). □

C.2 Proof of (5.1) and (5.2)

Recall that $f_1(x) = x$ and $f_2(x) = x^2$. 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)

$$\int f_2(x) d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = \int x^2 d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = \Big(\frac{\lfloor nt \rfloor}{n}\Big)^2 \left(1 + y_{\lfloor nt \rfloor}\right) = \frac{\lfloor nt \rfloor}{n}\Big(\frac{\lfloor nt \rfloor}{n} + y_n\Big).$$

We calculate the quantities given in Proposition C.1 by using the residue theorem. We find for the real case

$$\begin{aligned}
\mathbb{E}[X(f_1,t)] =& \frac{t}{2\pi i}\lim_{r\searrow 1}\oint_{|\xi|=1}\frac{\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi}{\xi}\Big(\frac{\xi}{\xi^2-r^{-2}}-\frac{1}{\xi}\Big)d\xi \\
=& \frac{t}{2\pi i}\lim_{r\searrow 1}\oint_{|\xi|=1}\frac{\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi}{(\xi-r^{-1})(\xi+r^{-1})}d\xi-\frac{t}{2\pi i}\oint_{|\xi|=1}\frac{\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi}{\xi^2}d\xi \qquad \text{(C.1)} \\
=& \lim_{r\searrow 1} t\frac{\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi}{\xi+r^{-1}}\Big|_{\xi=r^{-1}}+\lim_{r\searrow 1} t\frac{\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi}{\xi-r^{-1}}\Big|_{\xi=-r^{-1}} \\
&-t\frac{\partial}{\partial\xi}\left(\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi\right)\Big|_{\xi=0} \\
=& t\lim_{r\searrow 1}\frac{2r^{-1}+2h_t^2r^{-1}}{2r^{-1}}-t(1+h_t^2)=0.
\end{aligned}$$

Note that $\xi=\pm r^{-1}$ are poles of order 1 for the first integrand in (C.1), since $r>1$, while $\xi=0$ is a pole of order 2 for the second integrand in (C.1). 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

$$\begin{aligned}
I_1 =& \frac{t^2}{2\pi i}\lim_{r\searrow 1}\oint_{|\xi|=1}\frac{\left(\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi\right)^2}{\xi\left(\xi-r^{-1}\right)\left(\xi+r^{-1}\right)}d\xi, \\
I_2 =& \frac{t^2}{2\pi i}\lim_{r\searrow 1}\oint_{|\xi|=1}\frac{\left(\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi\right)^2}{\xi^3}d\xi.
\end{aligned}$$

The integrand in I_1 has poles which are all of order 1 at the points $0,r^{-1},-r^{-1}$. Thus, using the residue theorem,

$$\begin{aligned}
I_1 =& t^2\lim_{r\searrow 1}\frac{\left(\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi\right)^2}{\left(\xi-r^{-1}\right)\left(\xi+r^{-1}\right)}\Big|_{\xi=0}+t^2\lim_{r\searrow 1}\frac{\left(\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi\right)^2}{\xi\left(\xi+r^{-1}\right)}\Big|_{\xi=r^{-1}} \\
&+t^2\lim_{r\searrow 1}\frac{\left(\xi+h_tr\xi^2+h_tr^{-1}+h_t^2\xi\right)^2}{\xi\left(\xi-r^{-1}\right)}\Big|_{\xi=-r^{-1}} \\
=& -t^2h_t^2+\frac{t^2(1+h_t)^4}{2}+\frac{t^2(1-h_t)^4}{2}=-th^2+\frac{t^2(1+h_t)^4}{2}+\frac{t^2(1-h_t)^4}{2}.
\end{aligned}$$

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 $t_2 \le t_1$

$$\begin{aligned}
\operatorname{cov}(X(f_1,t_1), X(f_1,t_2)) =& \frac{1}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} t_1 \left(1 + h_{t_1}\xi_1 + h_{t_1} r_1^{-1}\xi_1^{-1} + h_{t_1}^2\right) t_2 \\
& \times \left(1 + h_{t_2}\xi_2 + h_{t_2} r_2^{-1}\xi_2^{-1} + h_{t_2}^2\right) \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1 \\
=& -\frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \frac{h_{t_1}t_1t_2(h_1 + r_1\xi_1 + h_1^2 r_1\xi_1 + h_1 r_1^2\xi_1^2)}{r_1^2 r_2^2 \xi_1^3} d\xi_1 \\
=& -\frac{(2\pi i)^2}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \frac{\partial^2}{\partial^2\xi_1} \frac{h_{t_1}t_1t_2(h_1 + r_1\xi_1 + h_1^2 r_1\xi_1 + h_1 r_1^2\xi_1^2)}{r_1^2 r_2^2}\Big|_{\xi_1=0} \\
=& 2 \lim_{\substack{r_1>r_2,\\ r_1,r_2\searrow 1}} \frac{h_1^2 t_2}{r_2^2} = 2h_1^2 t_2 = 2yt_2,
\end{aligned}$$

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 $(t_2 \le t_1)$

$$\begin{aligned}
&\operatorname{cov}(X(f_2,t_1), X(f_2,t_2)) \\
=& \frac{1}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} t_1^2 \left(1 + h_{t_1}\xi_1 + h_{t_1} r_1^{-1}\xi_1^{-1} + h_{t_1}^2\right)^2 t_2^2 \\
& \times \left(1 + h_{t_2}\xi_2 + h_{t_2} r_2^{-1}\xi_2^{-1} + h_{t_2}^2\right)^2 \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1 \\
=& \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \frac{2h_1t_1t_2^2}{r_1^4 r_2^4 \xi_1^5} \left(h_1 + r_1 x + h_1^2 r_1 x + h_1 r_1^2 \xi_1^2\right)^2 \\
& \times \left(-h_1t_1 - h_1^2 r_1 t_1 \xi_1 - r_1 r_2^2 t_1 \xi_1 - h_2^2 r_1 r_2^2 t_1 \xi_1 + h_2^2 r_1 t_2 \xi_1 + h_1^2 r_1^3 t_1 \xi_1^3 - h_2^2 r_1^3 t_2 \xi_1^3\right) d\xi_1 \qquad \text{(C.2)} \\
=& \frac{(2\pi i)^2}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \frac{1}{(5-1)!} \frac{\partial^4}{\partial^4\xi_1} \Big\{ \frac{2h_1t_1t_2^2}{r_1^4 r_2^4} \left(h_1 + r_1\xi_1 + h_1^2 r_1\xi_1 + h_1 r_1^2 \xi_1^2\right)^2 \\
& \times \left(-h_1t_1 - h_1^2 r_1 t_1 \xi_1 - r_1 r_2^2 t_1 \xi_1 - h_2^2 r_1 r_2^2 t_1 \xi_1 + h_2^2 r_1 t_2 \xi_1 + h_1^2 r_1^3 t_1 \xi_1^3 - h_2^2 r_1^3 t_2 \xi_1^3\right) d\xi_1 \Big\}\Big|_{\xi_1=0} \\
=& 4t_2 y \left\{2t_1t_2 + [t_2 + 2(t_1+t_2)]\, y + 2y^2\right\}, \ t_2 \le t_1.
\end{aligned}$$

Note that $\xi_1 = 0$ is a pole of order 5 for the integrand in (C.2) 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 $t_2 \le t_1$

$$\begin{aligned}
&\operatorname{cov}(X(f_1,t_1), X(f_2,t_2)) \\
=&\frac{1}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \oint_{|\xi_2|=1} t_1 \left(1 + h_{t_1}\xi_1 + h_{t_1} r_1^{-1}\xi_1^{-1} + h_{t_1}^2\right) t_2^2 \\
&\qquad \times \left(1 + h_{t_2}\xi_2 + h_{t_2} r_2^{-1}\xi_2^{-1} + h_{t_2}^2\right)^2 \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1 && \text{(C.3)}\\
=&\frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \frac{1}{r_1^3 r_2^4 \xi_1^4} 2h_1 t_2^2 (h_1 + r_1\xi_1 + h_1^2 r_1 \xi_1 + h_1 r_1^2 \xi_1^2) \\
&\times \left\{-h_1 t_1 + h_1^2 r_1 t_1 \xi_1(-1 + r_1^2\xi_1^2) - r_1\xi_1 \left[(1+h_2^2) r_2^2 t_1 + h_2^2 t_2(-1 + r_1^2\xi_1^2)\right]\right\} d\xi_1 && \text{(C.4)}\\
=&4t_2 y(t_2 + y). && \text{(C.5)}
\end{aligned}$$

After simplifying the integrand in (C.3) with a computer algebra program, we see that it has a pole at $\xi_2 = 0$ of order 2. Note that the pole at

$$\xi_2 = \frac{h_2 r_2}{h_1 r_1 \xi_1}$$

is not relevant for an application of the residue theorem, since

$$\left|\frac{h_2 r_2}{h_1 r_1 \xi_1}\right| = \left|\frac{t_1 r_2}{r_1 t_2 \xi_1}\right| = \frac{t_1 r_2}{t_2 r_1} > 1.$$

The integrand in (C.4) has a pole at $\xi_1 = 0$ of order 4. Similarly, we have, again for $t_2 \le t_1$,

$$\operatorname{cov}(X(f_2,t_1), X(f_1,t_2)) = 4t_2 y(t_1 + y). \tag{C.6}$$

By combining (C.5) and (C.6), we have for $t_1, t_2 \in [t_0, 1]$

$$\operatorname{cov}(X(f_1,t_1), X(f_2,t_2)) = 4\min(t_1,t_2) y(t_2 + y).$$

C.3 Proof of Corollary 2.1

We apply Theorem 2.1 for the choice $h(x) = \log(x)$. Note that, as $y \ge t_0$, the interval in (2.9) 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

$$\int \log x d\tilde{F}^{y_{\lfloor nt \rfloor}}(x) = \int \log x dF^{y_{\lfloor nt \rfloor}}\left(\frac{n}{\lfloor nt \rfloor} x\right) = \int \log x dF^{y_{\lfloor nt \rfloor}}(x) + \log\left(\frac{\lfloor nt \rfloor}{n}\right)$$

$$
\begin{aligned}
&=\Big(-1+\frac{y_{\lfloor nt\rfloor}-1}{y_{\lfloor nt\rfloor}}\log(1-y_{\lfloor nt\rfloor})\Big)+\log\Big(\frac{\lfloor nt\rfloor}{n}\Big)\\
&=-1-\frac{1}{y_{\lfloor nt\rfloor}}\log(1-y_{\lfloor nt\rfloor})+\log\Big(\frac{\lfloor nt\rfloor}{n}-y_n\Big),
\end{aligned}
$$

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 C.1, we have for the mean of the limiting process $\mathbb{D}$ in the real case

$$
\mathbb{E}[\mathbb{D}(t)]=I_1+I_2, \tag{C.7}
$$

where

$$
\begin{aligned}
I_1=&\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1}\log(t(1+h_tr\xi+h_tr^{-1}\xi^{-1}+h_t^2))\frac{\xi}{\xi^2-r^{-2}}d\xi\\
=&\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1}\log(t|1+h_t\xi|^2)\frac{\xi}{\xi^2-r^{-2}}d\xi, \qquad\qquad (C.8)\\
I_2=&-\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1}\log(t(1+h_tr\xi+h_tr^{-1}\xi^{-1}+h_t^2))\frac{1}{\xi}d\xi\\
=&-\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1}\log(t|1+h_t\xi|^2)\frac{1}{\xi}d\xi
\end{aligned}
$$

(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

$$
\begin{aligned}
I_{11}=&\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1}\log(t(1+h_t\xi))\frac{\xi}{(\xi-r^{-1})(\xi+r^{-1})}d\xi,\\
=&\frac{1}{2}\lim_{r\searrow 1}\left\{\log\left(t\left(1+h_tr^{-1}\right)\right)+\log\left(t\left(1-h_tr^{-1}\right)\right)\right\}=\frac{1}{2}\log\left(t^2\left(1-h_t^2\right)\right),\\
I_{12}=&\frac{1}{2\pi i}\lim_{r\searrow 1}\oint\limits_{|\xi|=1}\log(t(1+h_t\xi^{-1}))\frac{\xi}{(\xi-r^{-1})(\xi+r^{-1})}d\xi
\end{aligned}
$$

$$=\frac{1}{2\pi i}\lim_{r\searrow 1}\oint_{|z|=1}\log(t(1+h_t z))\frac{r^2}{z(z-r)(r+z)}dz$$

$$=\lim_{r\searrow 1}\log(t(1+h_t z))\frac{r^2}{(z-r)(z+r)}\Big|_{z=0}=-\log(t).$$

These calculations imply

$$I_1=\frac{1}{2}\log\left(t^2\left(1-h_t^2\right)\right)-\log(t). \tag{C.9}$$

The quantity I_2 in (C.7) can be determined similarly using the decomposition

$$I_2=I_{21}+I_{22},$$

where

$$I_{21}=-\frac{1}{2\pi i}\lim_{r\searrow 1}\oint_{|\xi|=1}\log(t(1+h_t\xi))\frac{1}{\xi}d\xi=-\log t$$

$$I_{22}=-\frac{1}{2\pi i}\lim_{r\searrow 1}\oint_{|\xi|=1}\log(t(1+h_t\xi^{-1}))\frac{1}{\xi}d\xi=\log t.$$

This gives $I_2=0$, and by (C.9) and (C.7), 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 (C.8), we obtain for $t_2\le t_1$

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

$$=\frac{1}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint_{|\xi_1|=1}\oint_{|\xi_2|=1}\log(t_1|1+h_{t_1}\xi_1|^2)\log(t_2|1+h_{t_2}\xi_2|^2)\frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)}d\xi_2 d\xi_1$$

$$=I_3+I_4,$$

where (note that $|1+h_{t_2}\xi_2|^2=(1+h_{t_2}\xi_2)(1+h_{t_2}\xi_2^{-1})$)

$$I_3=\frac{1}{2\pi^2}\lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}}\oint_{|\xi_1|=1}\log(t_1|1+h_{t_1}\xi_1|^2)\oint_{|\xi_2|=1}\log(t_2(1+h_{t_2}\xi_2))\frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)}d\xi_2 d\xi_1,$$

$$I_4 = \frac{1}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \log(t_1|1+h_{t_1}\xi_1|^2) \oint_{|\xi_2|=1} \log(t_2(1+h_{t_2}\xi_2^{-1})) \frac{g_1(\xi_1,\xi_2)}{g_2(\xi_1,\xi_2)} d\xi_2 d\xi_1.$$

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

$$I_4 = \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \log(t_1|1+h_{t_1}\xi_1|^2) \frac{h_1 r_1}{r_2 + h_1 r_1 \xi_1} d\xi_1 = I_{41} + I_{42},$$

where

$$\begin{aligned} I_{41} =& \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \log(t_1(1+h_{t_1}\xi_1)) \frac{h_1 r_1}{r_2 + h_1 r_1 \xi_1} d\xi_1, \\ I_{42} =& \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|\xi_1|=1} \log(t_1(1+h_{t_1}\xi_1^{-1})) \frac{h_1 r_1}{r_2 + h_1 r_1 \xi_1} d\xi_1. \end{aligned}$$

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

$$\begin{aligned} I_{42} =& - \frac{2\pi i}{2\pi^2} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \oint_{|z_1|=1} \log(t_1(1+h_{t_1}z_1)) \frac{h_{t_1} r_1}{h_{t_1} r_1 z_1 + r_2 z_1^2} dz_1 \\ =& - \frac{(2\pi i)^2}{2\pi} \lim_{\substack{r_2>r_1,\\ r_1,r_2\searrow 1}} \Big\{ -\log(t_1) + \log\Big(t_1\Big(1 - \frac{h_{t_1}^2 r_1}{r_2}\Big)\Big)\Big\} \\ =& -2\log(1-h_{t_1}^2). \end{aligned}$$

Finally, we obtain for $t_2 \leq t_1$

$$\operatorname{cov}(\mathbb{D}(t_1), \mathbb{D}(t_2)) = I_3 + I_4 = -2\log(1-h_{t_1}^2) = -2\log(1-y_{t_1}).$$