

Rate of Convergence for Sparse Sample Covariance Matrices



F. Götze, A. Tikhomirov, and D. Timushev

Abstract This work is devoted to the estimation of the convergence rate of the empirical spectral distribution function (ESD) of sparse sample covariance matrices in the Kolmogorov metric. We consider the case with the sparsity $np_n \sim \log^\alpha n$, for some $\alpha > 1$ and assume that the moments of the matrix elements satisfy the condition $\mathbb{E} |X_{jk}|^{4+\delta} \leq C < \infty$, for some $\delta > 0$. We also obtain approximation estimates for the Stieltjes transform in the bulk.

Keywords Sparse sample covariance matrices · Marchenko–Pastur law · Rate of convergence · Stieltjes transform

1 Introduction

For the first time, random matrices arose in the work of Wishart [1] in mathematical statistics and later in the work of Wigner [2] for applications in nuclear physics. Wigner used the eigenvalues of random Hermitian matrices with centered independent identically distributed elements to describe the energy levels of heavy nuclei. Later, such matrices were called Wigner matrices. Wigner proved that the density of the empirical spectral distribution function of the eigenvalues of such matrices converges to the semicircle law when the dimension of the matrix increases to infinity.

The second important class of random matrices is an ensemble of sample covariance matrices. These are matrices of the form $\mathbf{X}\mathbf{X}^*$, where $\mathbf{X}$ is a $n \times m$ matrix with

F. Götze

Faculty of Mathematics, Bielefeld University, Bielefeld, Germany

e-mail: goetze@math.uni-bielefeld.de

A. Tikhomirov (✉) · D. Timushev

Institute of Physics and Mathematics, Komi Science Center of Ural Division of RAS, Syktyvkar, Russia

e-mail: tikhomirov@ipm.komisc.ru

D. Timushev

e-mail: timushev@ipm.komisc.ru

centered identically distributed independent elements. In the context of statistical analysis, the columns of matrix $\mathbf{X}$ are an m -dimensional sample of a n -dimensional data vector. With the size of matrix $\mathbf{X}$ increasing to infinity, so that the ratio n/m converges to some constant, the density of the limiting empirical spectral distribution of the matrix $\mathbf{X}\mathbf{X}^*$ was explicitly found by Marchenko and Pastur [3]. Sample covariance matrices play an important role in the problems of multivariate statistical analysis, in particular in the method of Principal components analysis (PCA). Random matrices of this type also appear in the theory of wireless communication. The spectral density of these matrices is used in calculations related to the capacity of a multiple-input multiple-output (MIMO) channel. This connection between the theory of random matrices and wireless communication was established by Telatar [4]. In this linear model, the element X_{ij} of the channel matrix $\mathbf{X}$ is the transmission coefficient from the j -th transmitter to the i -th receiver antenna. The resulting signal can be represented by the linear relationship $\mathbf{y} = \mathbf{X}\mathbf{r} + \mathbf{s}$, where $\mathbf{r}$ is the input signal and $\mathbf{s}$ is zero-mean Gaussian noise with some variance σ^2 . In the case of i.i.d. Gaussian input signals, the channel capacity is determined by the expression $(1/n) \mathbb{E} \log \det(\mathbf{I} + \sigma^{-2}\mathbf{X}\mathbf{X}^*)$. This application of the random matrix theory was one of the reasons of the explosive growth in this topic in the next decade. An important area of sample covariance matrices applications is Graph theory. In the case of a directed graph, its adjacency matrix is asymmetric and so instead of its eigenvalues, we have to consider singular ones, which naturally leads to sample covariance matrices. An example of such a graph is a bipartite random graph, the vertices of which can be divided into two groups, in each of which the vertices are not connected to each other. If the edges probability p_n of such graph tends to zero with the number of vertices increasing to infinity, we get a sparse sample covariance matrix. The behaviour of eigenvalues and eigenvectors of a sparse random matrix significantly depends on its sparsity and we cannot apply the results obtained for the case of non-sparse matrices. The case of sparse matrices has been considered in a number of works (see [5–10]). The spectral properties of sparse sample covariance matrices with the sparsity $np_n \sim n^\alpha$ were studied in [8]. In particular, a local law for the eigenvalue density was proved and the limiting distribution of the shifted, rescaled largest eigenvalue was found assuming that matrix elements satisfy the moments assumption $\mathbb{E}|X_{jk}|^q \leq (Cq)^{cq}$. Estimates for the Stieltjes transform were also obtained in the domain $\{z = u + iv : \lambda_-/2 \leq u \leq \lambda_+ + 1, 0 < v < 3\}$, where λ_- and λ_+ are left and right endpoints of the Marchenko-Pastur distribution support, respectively.

The present work is devoted to the estimation of the convergence rate of the empirical spectral distribution function (ESD) of sparse sample covariance matrices in the Kolmogorov metric. We consider the case with the sparsity $np_n \sim \log^\alpha n$, for some $\alpha > 1$ and assume that matrix elements moments satisfy the condition $\mathbb{E}|X_{jk}|^{4+\delta} \leq C < \infty$. It should be noted that in addition to the result on the rate of convergence, we also obtain approximation estimates for the Stieltjes transform in the support of the limiting distribution assuming some truncation condition (see Sect. 3).

2 Main Results

Let $m = m(n)$, $m \geq n$. Consider independent identically distributed zero-mean random variables X_{jk} , $1 \leq j \leq n$, $1 \leq k \leq m$ with $\mathbb{E} X_{jk} = 1$ and independent set of independent Bernoulli random variables ξ_{jk} , $1 \leq j \leq n$, $1 \leq k \leq m$ with $\mathbb{E} \xi_{jk} = p_n$. In addition suppose that $np_n \rightarrow \infty$ as $n \rightarrow \infty$. In what follows, for simplicity of notation, we will omit the index n in p_n if this does not cause confusion.

Observe the sequence of random matrices

$$\mathbf{X} = \frac{1}{\sqrt{mp_n}} (\xi_{jk} X_{jk})_{1 \leq j \leq n, 1 \leq k \leq m}.$$

Denote by $s_1 \geq \dots \geq s_n$ the singular values of $\mathbf{X}$ and define the empirical spectral distribution function (ESD) of the sample covariance matrix $\mathbf{W} = \mathbf{X}\mathbf{X}^*$:

$$F_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbb{I}\{s_j^2 \leq x\},$$

where $\mathbb{I}\{A\}$ stands for the event A indicator.

We are interested in the rate of the convergence of the distribution $F_n(x)$ to the Marchenko–Pastur distribution $G_y(x)$ with the density

$$g_y(x) = \frac{1}{2\pi yx} \sqrt{(x - a^2)(b^2 - x)} \mathbb{I}\{a^2 \leq x \leq b^2\},$$

where $y = y(n) = \frac{n}{m}$ and $a = 1 - \sqrt{y}$, $b = 1 + \sqrt{y}$.

We will follow the standard symmetrization technique (see [11]). To linearize the ESD we consider the symmetrized distribution function

$$\tilde{F}_n(x) = \frac{1 + \text{sign}(x)F_n(x^2)}{2},$$

and the symmetrized Marchenko–Pastur distribution function $\tilde{G}_y(x)$ with the density

$$\tilde{g}_y(x) = \frac{1}{2\pi y|x|} \sqrt{(x^2 - a^2)(b^2 - x^2)} \mathbb{I}\{a^2 \leq x^2 \leq b^2\}.$$

Then we have for the rate of the convergence of the ESD function to the Marchenko–Pastur distribution function in the Kolmogorov metric that

$$\Delta_n := \sup_x |F_n(x) - G_y(x)| = 2 \sup_x |\tilde{F}_n(x) - \tilde{G}_y(x)|.$$

In what follows we shall consider the symmetrized ESD only and we shall omit the symbol “ $\sim$ ” in the notation.

We state the following result.

Theorem 1 *Let $\mathbb{E} X_{jk} = 0$ and $\mathbb{E} |X_{jk}|^2 = 1$. Assume that*

$$\mathbb{E} |X_{jk}|^{4+\delta} \leq \mu_{4+\delta} < \infty,$$

for any $j, k \geq 1$ and for some $\delta > 0$. Suppose that there exists a positive constant B , such that

$$np_n \geq B \log^{\frac{2}{\varkappa}} n,$$

where $\varkappa = \frac{\delta}{2(4+\delta)}$. Then there exists a constant C depending on δ and $\mu_{4+\delta}$ such that

$$\Delta_n \leq \frac{C \log n}{np_n},$$

with high probability.

Remark 1 We say that some event A_n holds with high probability, if for any $Q > 0$, $\mathbb{P}\{A_n\} \geq 1 - n^{-Q}$, for sufficiently large n .

To prove Theorem 1, we will use the smoothing inequality developed in [11], Corollary 3.2. For the reader's convenience, we reproduce the statement.

For any x such that $|x| \in [1 - \sqrt{y}, 1 + \sqrt{y}]$, define the quantities

$$\beta = \beta(x) := \sqrt{y} - ||x| - 1|,$$

and

$$\mathbb{J}_\varepsilon = [1 - \sqrt{y} + \frac{1}{2}\varepsilon, 1 + \sqrt{y} - \frac{1}{2}\varepsilon], \text{ for any } 0 < \varepsilon < \sqrt{y}.$$

Lemma 1 (Smoothing inequality) *Suppose that $y < 1$. Let ε, v_0 be positive numbers such that $\varepsilon \leq \sqrt{y}/2$ and $v_0 < C\varepsilon^{3/2}$. Denote by $s_n(z)$ the Stieltjes transform of $F_n(x)$, and by $S_y(z)$ — the Stieltjes transform of $G_y(x)$. Note that $0 \leq \beta \leq \sqrt{y}$. For any $x : |x| \in \mathbb{J}_\varepsilon$, we have $\beta \geq \frac{1}{2}\varepsilon$. Then the following inequality holds:*

$$\begin{aligned} \Delta_n \leq & 2 \int_{-\infty}^{\infty} |s_n(u + iV) - S_y(u + iV)| du + C_1 v_0 + C_2 \varepsilon^{\frac{3}{2}} \\ & + 2 \sup_{|x| \in \mathbb{J}_\varepsilon} \int_{v'}^V |s_n(x + iu) - S_y(x + iu)| du, \end{aligned} \quad (1)$$

where $V > v_0$, $v' = \frac{v_0}{\sqrt{\beta}}$.

The last inequality implies that it is enough to estimate the proximity of the corresponding Stieltjes transforms in the domain

$$\mathcal{D}_\varepsilon := \{z = u + iv : |u| \in \mathbb{J}_\varepsilon, v_0/\sqrt{\beta} \leq v \leq V\},$$

where the exact values v_0, ε will be defined later. This will imply the bound for Δ_n .

The paper is organized as follows. In Sect. 3 we investigate the Stieltjes transform of the symmetrized ESD function in the domain $\mathcal{D}_\varepsilon$ assuming that the elements X_{jk} satisfy some truncation and moment assumption (C0) (will be defined below). We state Theorem 2 and prove several Corollaries. In Sect. 4.1 we show that the elements X_{jk} may satisfy the condition (C0). In Sect. 4.2 smoothing inequality (1) and Corollary 3 allow us to bound the quantity Δ_n . Sect. 5 is devoted to the proof of Theorem 2: in Sect. 5.1 we get the bound for the diagonal elements of resolvent matrix $\mathbf{R}(z)$ when $z \in \mathcal{D}_\varepsilon$ (the main difficulty is with $\text{Im } z$ close to real axis of the complex plane), and in Sect. 5.2 we estimate T_n . In the Appendix we state and prove some auxiliary results.

3 The Stieltjes Transforms Proximity

Note that the symmetrized empirical spectral distribution function (ESD) of the sample covariance matrix $\mathbf{W} = \mathbf{X}\mathbf{X}^*$ may be rewrite as

$$F_n(x) = \frac{1}{2n} \sum_{j=1}^n \left(\mathbb{I}\{s_j \leq x\} + \mathbb{I}\{-s_j \leq x\} \right).$$

By the definition of the Stieltjes transform we have

$$s_n(z) = \frac{1}{2n} \left(\sum_{j=1}^n \frac{1}{s_j - z} + \sum_{j=1}^n \frac{1}{-s_j - z} \right) = \frac{1}{n} \sum_{j=1}^n \frac{z}{s_j^2 - z^2}$$

and

$$S_y(z) = \frac{-(z^2 - ab) + \sqrt{(z^2 - a^2)(z^2 - b^2)}}{2yz}.$$

Using the inequality $|S_y(z)| \leq \frac{1}{\sqrt{y}}$ we obtain $\frac{1}{|yS_y(z)+z|} = \left| yS_y(z) - \frac{1-y}{z} \right| \leq 1 + 2\sqrt{y}$, for $z \in \mathcal{D}_\varepsilon$. Note that $F_n(x)$ is the ESD of the block matrix

$$\mathbf{V} = \begin{bmatrix} \mathbf{O}_n & \mathbf{X} \\ \mathbf{X}^* & \mathbf{O}_m \end{bmatrix},$$

where $\mathbf{O}_k$ is $k \times k$ matrix with zero elements.

Let $\mathbf{R} = \mathbf{R}(z)$ be the resolvent matrix of $\mathbf{V}$: $\mathbf{R} = (\mathbf{V} - z\mathbf{I}_{n+m})^{-1}$. It is easy to see that

$$\mathbf{R} = \begin{bmatrix} z(\mathbf{X}\mathbf{X}^* - z^2\mathbf{I}_n)^{-1} & (\mathbf{X}\mathbf{X}^* - z^2\mathbf{I}_n)^{-1}\mathbf{X} \\ \mathbf{X}^*(\mathbf{X}\mathbf{X}^* - z^2\mathbf{I}_n)^{-1} & z(\mathbf{X}^*\mathbf{X} - z^2\mathbf{I}_m)^{-1} \end{bmatrix}.$$

This implies

$$s_n(z) = \frac{1}{n} \sum_{j=1}^n R_{jj} = \frac{1}{n} \sum_{l=1}^m R_{l+n, l+n} + \frac{m-n}{nz}. \quad (2)$$

Throughout this section the conditions (C0) are assumed to be fulfilled:

$$\mathbb{E}|X_{jk}|^{4+\delta} \leq \mu_{4+\delta} < \infty \text{ and } |X_{jk}| \leq C(np)^{\frac{1}{2}-\varkappa}, \text{ with } \varkappa = \frac{\delta}{2(4+\delta)}. \quad (\text{C0})$$

Let $s_0 > 1$ be some constant depending on δ , V be some positive constant. The exact values of these constants will be determined below. For any $0 < v \leq V$ we define k_v as $k_v = k_v(V) := \min\{l \geq 0 : s_0^l v \geq V\}$. Let

$$\Lambda_n = \Lambda_n(z) := s_n(z) - S_y(z),$$

where $z = u + iv$ and $v > 0$. Set the notation

$$b(z) = z - \frac{1-y}{z} + 2yS_y(z), \quad a_n(z) = \operatorname{Im} b(z) + \frac{\log(np) \log n}{np|b(z)|},$$

and

$$b_n(z) = z - \frac{1-y}{z} + 2yS_y(z) + y\Lambda_n(z) = b(z) + y\Lambda_n(z).$$

Note that, for $z \in \mathcal{D}_\varepsilon$

$$v + 2y \operatorname{Im} S_y(z) \leq \operatorname{Im} b(z) \leq |b(z)| \leq C(y, V),$$

where $C(y, V) = 2 + 3\sqrt{y} + V$. For given $\gamma > 0$ consider the event

$$\mathcal{Q}_\gamma(v) := \{|\Lambda_n(u + iv)| \leq \gamma a_n(u + iv), \text{ for all } u : |u| \in \mathbb{J}_\varepsilon\}$$

and the event $\widehat{\mathcal{Q}}_\gamma(v) = \bigcap_{l=0}^{k_v} \mathcal{Q}_\gamma(s_0^l v)$. For any γ there exists the constant $V = V(\gamma)$ such that

$$\Pr\{\widehat{\mathcal{Q}}_\gamma(V)\} = 1. \quad (3)$$

It may be $V = \sqrt{2/\gamma}$, for example. In what follows, we assume that γ and V are chosen so that equality (3) is satisfied. We will also assume that γ is chosen so that for all $0 < v \leq V$

$$\gamma \sqrt{y} a_n(z) \max\{1, (1 + 2\sqrt{y})\sqrt{y}\} < 1/4.$$

We denote $\mathbf{Q} = \widehat{\mathbf{Q}}_y(V)$. Summing the equations in (2), we get

$$s_n(z) = -\frac{1}{z + yS_y(z) - \frac{1-y}{z}}(1 + T_n - y\Lambda_n s_n(z)) = S_y(z)(1 + T_n - y\Lambda_n s_n(z)), \quad (4)$$

where

$$T_n = \frac{1}{n} \sum_{j=1}^n \varepsilon_j R_{jj},$$

where

$$\begin{aligned} \varepsilon_j &= \varepsilon_{j1} + \varepsilon_{j2} + \varepsilon_{j3}, \\ \varepsilon_{j1} &= \frac{1}{m} \sum_{l=1}^m R_{l+n, l+n} - \frac{1}{m} \sum_{l=1}^m R_{l+n, l+n}^{(j)}, \\ \varepsilon_{j2} &= \frac{1}{mp} \sum_{l=1}^m (X_{jl}^2 \xi_{jl} - p) R_{l+n, l+n}^{(j)}, \\ \varepsilon_{j3} &= \frac{1}{mp} \sum_{1 \leq l \neq k \leq m} X_{jl} X_{jk} \xi_{jl} \xi_{jk} R_{l+n, k+n}^{(j)}. \end{aligned}$$

Here by symbol $R_{l,k}^{(j)}$ we denote the elements of the resolvent matrix defined via matrix $\mathbf{X}$ with row j deleted.

In this section we formulate the following result.

Theorem 2 *Under the conditions (C0), for all $z \in \mathcal{D}_\varepsilon$ with $v_0 = Cn^{-1} \log^4 n$, $\varepsilon \geq \left(\frac{C \log n}{n}\right)^{2/3}$, for any $K \geq 1$ there exists constant $C = C(K)$ and $q \leq K \log n$, such that*

$$\mathbb{E} |T_n|^q \mathbb{I}\{\mathbf{Q}\} \leq C^q q^{\frac{q}{2}} (1 + \sqrt{nv} q^{q/2}) \beta_n^q,$$

where

$$\beta_n = \frac{a_n(z)}{nv} + \frac{1}{np}.$$

Corollary 1 *Under the conditions of Theorem 2, in the domain $\mathcal{D}_\varepsilon$, for any $K \geq 1$ there exists constant $C = C(K)$ and $q \leq K \log n$, such that*

$$\mathbb{E} |\Lambda_n|^q \mathbb{I}\{\mathbf{Q}\} \leq \frac{C^q q^q \beta_n^q}{|b(z)|^q}.$$

Proof From Theorem 2 it follows

$$\mathbb{E} |T_n|^q \mathbb{I}\{\mathbf{Q}\} \leq C^q q^q \beta_n^q.$$

It is straightforward to check that $|b_n(z)|\mathbb{I}(\mathcal{Q}) \geq c|b(z)|\mathbb{I}(\mathcal{Q})$ in the domain $\mathcal{D}_\varepsilon$. Thus it is enough to consider the inequality $|\Lambda_n| \leq \frac{C|T_n|}{|b(z)|}$. The Corollary is proved. $\square$

Corollary 2 *Under the conditions of Theorem 2, in the domain $\mathcal{D}_\varepsilon$ for any $Q > 1$ there exists a constant C depending on Q , such that*

$$\Pr \left\{ |\Lambda_n| > C \frac{\beta_n \log n}{|b(z)|}; \mathcal{Q} \right\} \leq Cn^{-Q}.$$

Proof For the proof it is enough to apply Markov inequality choosing $q \sim \log n$ and result of Corollary 1. $\square$

Corollary 3 *Under the conditions of Theorem 2, in the domain $\mathcal{D}_\varepsilon$ for any $Q > 1$ there exists a constant C depending on Q , such that*

$$\Pr \left\{ |\Lambda_n| > C \frac{\beta_n \log n}{|b(z)|} \right\} \leq Cn^{-Q}.$$

Proof For the proof of this Corollary we shall use the so called additive descent method. We partition interval $[v', V]$ by points

$$v_k = V - k\delta_n, \quad k = 1, \dots, K,$$

where $K = (V - v')/\delta_n$ and δ_n is choosing so that for any $k = 1, \dots, K$ and any $v_k \geq v \geq v_{k+1}$

$$|\Lambda_n(u + iv_k) - \Lambda_n(u + iv)| \leq \delta_n/v^2 \leq 1/n.$$

It is enough to put $\delta_n = 1/n^3$. Note that by Lemma 6, we have

$$\frac{\beta_n}{|b(z)|} \leq \frac{C|a_n(z)|}{nv|b(z)|} \leq \frac{1}{4}a_n(u + iv_1) \text{ and } \max \left\{ \frac{1}{np|b(z)|}, \frac{C}{np} \right\} \leq \frac{1}{4}a_n(u + iv_1).$$

These implies

$$\Pr\{|\Lambda_n(u + iv_1)| \geq a_n(u + iv_1)\} \leq Cn^{-Q}.$$

Applying Corollary 2, we get

$$\Pr \left\{ |\Lambda_n(u + iv_1)| \geq C \frac{\beta_n \log n}{|b(z)|}; \mathcal{Q} \right\} \leq Cn^{-Q}.$$

Repeating this for $k = 1, \dots, K$ and using the union bound, we get

$$\Pr \left\{ |\Lambda_n(u + iv)| > C \frac{\beta_n \log n}{|b(z)|}, \text{ for any } V \geq v \geq v' \right\} \leq Cn^{-Q}.$$

Thus Corollary 3 is proved. $\square$

Corollary 4 *Under the conditions of Theorem 2, in the domain $\mathcal{D}_\varepsilon$ we have, for $q \sim \log n$,*

$$\Pr \left\{ \bigcup_{1-\sqrt{y} \leq |u| \leq 1+\sqrt{y}} \left\{ |\Lambda_n(u + iv)| > K \frac{\beta_n \log n}{|b(z)|} \right\} \right\} \leq Cn^{-Q}.$$

Proof Note that, for $v \geq v_0$,

$$\left| \frac{\partial \Lambda_n(u + iv)}{\partial u} \right| \leq \frac{2}{v^2} \leq n^2.$$

Split interval $[1 - \sqrt{y}, 1 + \sqrt{y}]$ into $N = n^3$ intervals with endpoints u_k , $k = 1, \dots, N$ such that $|u_{k+1} - u_k| \leq \frac{C}{N}$ and for any $u_k \leq u_{k+1}$,

$$|\Lambda_n(u + iv) - \Lambda_n(u_k + iv)| \leq K \frac{\beta_n \log n}{2|b(z)|}.$$

Further we write

$$\Pr \left\{ \bigcup_{||u|-1| \leq \sqrt{y}} \left\{ |\Lambda_n(u + iv)| > K \frac{\beta_n \log n}{|b(z)|} \right\} \right\} \leq \Pr \left\{ \bigcup_k \left\{ |\Lambda_n(u_k + iv)| > K \frac{\beta_n \log n}{2|b(z)|} \right\} \right\}.$$

Applying the union bound, we get

$$\Pr \left\{ \bigcup_{||u|-1| \leq \sqrt{y}} \left\{ |\Lambda_n(u + iv)| > K \frac{\beta_n \log n}{|b(z)|} \right\} \right\} \leq \sum_{k=1}^N \Pr \left\{ |\Lambda_n(u_k + iv)| > K \frac{\beta_n \log n}{2|b(z)|} \right\}.$$

Taking into account that $N \leq Cn^3$ for $q \sim \log n$, and applying Corollary 3 we obtain the result. $\square$

4 The Proof of Theorem 1

In this section we assume only that

$$\max_{j,k \geq 1} \mathbb{E} |X_{jk}|^{4+\delta} = \mu_{4+\delta} < \infty.$$

Using the standard truncation procedure we show that the Kolmogorov distance between the empirical spectral distribution and the limiting one is of order $O((np)^{-1} \log n)$. This implies, the proximity of the eigenvalues to the corresponding quantiles of the limit distribution of order $O(\Delta_n)$ with probability close to 1.

4.1 Truncation

In the proof we use the standard truncation procedure (see [11], Sect. 5, for example). Introduce the truncated random values

$$\widehat{X}_{jk} = X_{jk} \mathbb{I}\{|X_{jk}| \leq C(np)^{\frac{1}{2}-\varkappa}\}, \text{ where } \varkappa = \frac{\delta}{2(4+\delta)},$$

the centered random values

$$\widetilde{X}_{jk} = \widehat{X}_{jk} - \mathbb{E} \widehat{X}_{jk},$$

and the normalized random values

$$\check{X}_{jk} = \sigma^{-1} \widetilde{X}_{jk}, \text{ where } \sigma^2 = \mathbb{E} |\widetilde{X}_{jk}|^2.$$

Denote by $\widehat{\mathbf{V}}, \widetilde{\mathbf{V}}, \check{\mathbf{V}}$ matrices obtained from $\mathbf{V}$ by replacing X_{jk} with $\widehat{X}_{jk}, \widetilde{X}_{jk}, \check{X}_{jk}$. Let $\widehat{F}_n(x), \widetilde{F}_n(x), \check{F}_n(x)$ be the symmetrized ESD functions of the corresponding matrices.

Lemma 2

$$\sup_x |F_n(x) - G_y(x)| \leq \sup_x |\check{F}_n(x) - G_y(x)| + \frac{C}{np}.$$

Proof We will successively estimate the values $\sup_x |\widehat{F}_n(x) - F_n(x)|, \sup_x |\widetilde{F}_n(x) - \widehat{F}_n(x)|$ and $\sup_x |\check{F}_n(x) - \widetilde{F}_n(x)|$.

First we note that

$$\Pr\{\widehat{X}_{jk} \neq X_{jk}\} \leq \frac{\mu_{4+\delta}}{(np)^2}. \quad (5)$$

Bai's rank inequality (see [12], Theorem A.43) gives

$$\sup_x |\widehat{F}_n(x) - F_n(x)| \leq \frac{\text{rank}\{\widehat{\mathbf{V}} - \mathbf{V}\}}{n}.$$

Since the rank does not exceed the number of non-zero elements of the matrix, we have

$$\sup_x |\widehat{F}_n(x) - F_n(x)| \leq \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^m \xi_{jk} \mathbb{I}\{|X_{jk}| \geq C(np)^{\frac{1}{2}-\varkappa}\} = \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^m \eta_{jk},$$

where $\eta_{jk} = \xi_{jk} \mathbb{I}\{|X_{jk}| \geq C(np)^{\frac{1}{2}-\varkappa}\}$ are Bernoulli distributed with $\mathbb{E} \eta_{jk} \leq 1/(pn^2)$. This implies

$$\begin{aligned}
\mathbb{E} \left[\sup_x |\widehat{F}_n(x) - F_n(x)| \right]^q &\leq \frac{C^q}{n^q} \left[\left(\sum_{j=1}^n \sum_{k=1}^m \mathbb{E} \eta_{jk} \right)^q + \left| \sum_{j=1}^n \sum_{k=1}^m (\eta_{jk} - \mathbb{E} \eta_{jk}) \right|^q \right] \\
&\leq \frac{C^q}{n^q p^q} + \frac{C^q q^{q/2}}{n^q} \left(\sum_{j=1}^n \sum_{k=1}^m \mathbb{E} (\eta_{jk} - \mathbb{E} \eta_{jk})^2 \right)^{q/2} + \frac{C^q q^q}{n^q} \sum_{j=1}^n \sum_{k=1}^m \mathbb{E} |\eta_{jk} - \mathbb{E} \eta_{jk}|^q.
\end{aligned}$$

Applying (5) and Rosenthal's inequality (see [11], Lemma 8.1), we obtain

$$\begin{aligned}
\mathbb{E} \left[\sup_x |\widehat{F}_n(x) - F_n(x)| \right]^q &\leq \frac{C^q \mu_{4+\delta}^q}{(np)^q} + \frac{C^q q^{\frac{q}{2}}}{n^q p^{\frac{q}{2}}} + \frac{C^q q^q}{n^q p} \\
&\leq \frac{C^q \max\{1, (qp)^{\frac{q}{2}}, (qp)^{q-1} q\}}{(np)^q}.
\end{aligned}$$

Using Chebyshev's inequality, we get

$$\Pr \left\{ \sup_x |\widehat{F}_n(x) - F_n(x)| \geq K \frac{C \max\{1, \sqrt{qp}, qp\}}{np} \right\} \leq \frac{1}{K^q}.$$

Choosing $q \sim \log n$ and $K \sim \text{const}$, we write

$$\Pr \left\{ \sup_x |\widehat{F}_n(x) - F_n(x)| \geq \frac{C \max\{1, p \log n\}}{np} \right\} \leq C n^{-Q}.$$

Consider the centered ESD function $\widehat{F}_n(x)$. Note that the elements X_{jk} are i.i.d., so the rank of the matrix $\widehat{\mathbf{V}} - \widetilde{\mathbf{V}}$ is 1 (since all elements of the matrix are the same constant $\mathbb{E} \widehat{X}_{jk}$). Therefore

$$\sup_x |\widehat{F}_n(x) - \widetilde{F}_n(x)| \leq \frac{C}{n}.$$

Consider now the normalized ESD function $\check{F}_n(x)$. It is easy to see that

$$\sup_x |\widetilde{F}_n(x) - G_y(x)| \leq \sup_x |\check{F}_n(x) - G_y(x)| + \sup_x |G_y(x\sigma) - G_y(x)|.$$

Note that the random variables $\check{X}_{jk}$ satisfy the condition (C0). Given that the density of the distribution function $G_y(x)$ is bounded by a constant, we obtain

$$\sup_x |\widetilde{F}_n(x) - G_y(x)| \leq \sup_x |\check{F}_n(x) - G_y(x)| + C|\sigma - 1|.$$

It can be shown, that

$$|1 - \sigma| \leq \frac{C}{(np)^{1+2\kappa}}.$$

Thus,

$$\sup_x |\widetilde{F}_n(x) - G_y(x)| \leq \sup_x |\check{F}_n(x) - G_y(x)| + \frac{C}{(np)^{1+2\kappa}}.$$

The lemma is proved. $\square$

4.2 The Proof of Theorem

Now we may consider the random variables satisfying condition (C0). Put $v_0 = Cn^{-1} \log^4 n$, $\varepsilon = \left(\frac{C \log n}{n}\right)^{2/3}$. We shall apply smoothing inequality (1) and the result of Corollary 3.

The estimation of the first term in the smoothing inequality when $\text{Im } z = V$ is standard procedure since $|R_{jj}| < 1/V$. It is not difficult to obtain the estimation

$$\int_{-\infty}^{\infty} |s_n(u + iV) - S_y(u + iV)| du \leq C \left(\frac{1}{n} + \frac{1}{np} \right).$$

The main problem appears in the application of the second integral of the smoothing inequality. Using the bound from Corollary 3 we get

$$\begin{aligned} \sup_{|u| \in \mathbb{J}_\varepsilon} \int_{v'}^V |s_n(u + iv) - S_y(u + iv)| dv &\leq \sup_{|u| \in \mathbb{J}_\varepsilon} \int_{v'}^V \left(\frac{a_n(z)}{nv} + \frac{1}{np} \right) \frac{C \log n}{|b(z)|} dv \\ &\leq \int_{v'}^V \frac{C \log n}{nv} dv + \sup_{|u| \in \mathbb{J}_\varepsilon} \int_{v'}^V \frac{C \log^2 n \log(np)}{nvn p |b(z)|^2} dv + \sup_{|u| \in \mathbb{J}_\varepsilon} \int_{v'}^V \frac{C \log n}{np |b(z)|} dv \\ &\leq \frac{C \log^2 n}{n} + \frac{C \log^4 n}{npn\beta} + \frac{C \log n}{np} \leq \frac{C \log n}{np}. \end{aligned}$$

The Theorem is proved.

5 The Proof of Theorem 2

The main difficulty in proving the Theorem is to obtain an estimate for diagonal elements of resolvent $\mathbf{R}$ in the domain $\mathcal{D}_\varepsilon$.

5.1 Estimation of Resolvent Diagonal Elements

Let $\mathbb{T} = \{1, \dots, n\}$, $\mathbb{J} \subset \mathbb{T}$ and $\mathbb{T}^{(1)} = \{1, \dots, m\}$, $\mathbb{K} \subset \mathbb{T}^{(1)}$. Consider σ -algebras $\mathfrak{M}^{(\mathbb{J}, \mathbb{K})}$, generated by the elements of $\mathbf{X}$ with the exception of the rows with number in $\mathbb{J}$ and the columns with number in $\mathbb{K}$. We will write for brevity $\mathfrak{M}_j$ instead of $\mathfrak{M}^{(\mathbb{J} \cup \{j\}, \mathbb{K})}$ and $\mathfrak{M}_{l+n}$ instead of $\mathfrak{M}^{(\mathbb{J}, \mathbb{K} \cup \{l\})}$. By symbol $\mathbf{X}^{(\mathbb{J}, \mathbb{K})}$ we denote the matrix $\mathbf{X}$ which rows with numbers in $\mathbb{J}$ are deleted, and which columns with numbers in $\mathbb{K}$ are deleted too. In a similar way, we will denote all objects defined via $\mathbf{X}^{(\mathbb{J}, \mathbb{K})}$, such as the resolvent matrix $\mathbf{R}^{(\mathbb{J}, \mathbb{K})}$, the ESD Stieltjes transform $s_n^{(\mathbb{J}, \mathbb{K})}$, $\Lambda_n^{(\mathbb{J}, \mathbb{K})}$ and so on. The symbol $\mathbb{E}_j$ denotes the conditional expectation with respect to the σ -algebra $\mathfrak{M}_j$, and $\mathbb{E}_{l+n}$ — with respect to σ -algebra $\mathfrak{M}_{l+n}$. Let $\mathbb{J}^c = \mathbb{T} \setminus \mathbb{J}$, $\mathbb{K}^c = \mathbb{T}^{(1)} \setminus \mathbb{K}$. Consider the events

$$\mathcal{A}_{jk}(v, \mathbb{J}, \mathbb{K}; C) = \{|R_{jk}^{(\mathbb{J}, \mathbb{K})}(u + iv)| \leq C\}. \quad (6)$$

Set

$$\begin{aligned} \mathcal{A}^{(1)}(v, \mathbb{J}, \mathbb{K}) &= \cap_{j=1}^n \cap_{k=1}^m \mathcal{A}_{j,k}(v, \mathbb{J}, \mathbb{K}), \\ \mathcal{A}^{(2)}(v, \mathbb{J}, \mathbb{K}) &= \cap_{j=1}^m \cap_{k=1}^n \mathcal{A}_{j+n,k}(v, \mathbb{J}, \mathbb{K}), \\ \mathcal{A}^{(3)}(v, \mathbb{J}, \mathbb{K}) &= \cap_{j=1}^n \cap_{k=1}^m \mathcal{A}_{j,k+n}(v, \mathbb{J}, \mathbb{K}), \\ \mathcal{A}^{(4)}(v, \mathbb{J}, \mathbb{K}) &= \cap_{j=1}^m \cap_{k=1}^n \mathcal{A}_{j+n,k+n}(v, \mathbb{J}, \mathbb{K}). \end{aligned} \quad (7)$$

For all $\{j, k\} \in (\mathbb{J}^c \cup \mathbb{K}^c) \times (\mathbb{J}^c \cup \mathbb{K}^c)$ and u we have $|R_{jk}^{(\mathbb{J}, \mathbb{K})}| \leq \frac{1}{v}$. Introduce events

$$\widehat{\mathcal{Q}}_v^{(\mathbb{J}, \mathbb{K})}(v) := \bigcap_{l=0}^{k_v} \left\{ \left| \Lambda_n^{(\mathbb{J}, \mathbb{K})}(u + i s_0^l v) \right| \leq \gamma a_n(u, s_0^l v) + \frac{|\mathbb{J}| + |\mathbb{K}|}{n s_0^l v} \right\}.$$

It is easy to see that

$$\widehat{\mathcal{Q}}_v(v) \subset \widehat{\mathcal{Q}}_v^{(\mathbb{J}, \mathbb{K})}(v).$$

For the diagonal elements of $\mathbf{R}$ we can write

$$R_{jj}^{(\mathbb{J}, \mathbb{K})} = S_y(z) \left(1 - \varepsilon_j^{(\mathbb{J}, \mathbb{K})} R_{jj}^{(\mathbb{J}, \mathbb{K})} + y \Lambda_n^{(\mathbb{J}, \mathbb{K})} R_{jj}^{(\mathbb{J}, \mathbb{K})} \right), \text{ for } j \in \mathbb{J}^c, \quad (8)$$

$$R_{l+n, l+n}^{(\mathbb{J}, \mathbb{K})} = -\frac{1}{z + y S_y(z)} \left(1 - \varepsilon_{l+n}^{(\mathbb{J}, \mathbb{K})} R_{l+n, l+n}^{(\mathbb{J}, \mathbb{K})} + y \Lambda_n^{(\mathbb{J}, \mathbb{K})} R_{l+n, l+n}^{(\mathbb{J}, \mathbb{K})} \right), \text{ for } l \in \mathbb{K}^c. \quad (9)$$

Correction terms $\varepsilon_j^{(\mathbb{J}, \mathbb{K})}$ for $j \in \mathbb{J}^c$ and $\varepsilon_{l+n}^{(\mathbb{J}, \mathbb{K})}$ for $l \in \mathbb{K}^c$ are defined as

$$\begin{aligned}
\varepsilon_j^{(\mathbb{J}, \mathbb{K})} &= \varepsilon_{j1}^{(\mathbb{J}, \mathbb{K})} + \cdots + \varepsilon_{j3}^{(\mathbb{J}, \mathbb{K})}, \\
\varepsilon_{j1}^{(\mathbb{J}, \mathbb{K})} &= \frac{1}{m} \sum_{l=1}^m R_{l+n, l+n}^{(\mathbb{J}, \mathbb{K})} - \frac{1}{m} \sum_{l=1}^m R_{l+n, l+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})}, \\
\varepsilon_{j2}^{(\mathbb{J}, \mathbb{K})} &= \frac{1}{mp} \sum_{l=1}^m (X_{jl}^2 \xi_{jl} - p) R_{l+n, l+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})}, \\
\varepsilon_{j3}^{(\mathbb{J}, \mathbb{K})} &= \frac{1}{mp} \sum_{1 \leq l \neq k \leq m} X_{jl} X_{jk} \xi_{jl} \xi_{jk} R_{l+n, k+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})},
\end{aligned}$$

and

$$\begin{aligned}
\varepsilon_{l+n}^{(\mathbb{J}, \mathbb{K})} &= \varepsilon_{l+n,1}^{(\mathbb{J}, \mathbb{K})} + \cdots + \varepsilon_{l+n,3}^{(\mathbb{J}, \mathbb{K})}, \\
\varepsilon_{l+n,1}^{(\mathbb{J}, \mathbb{K})} &= \frac{1}{m} \sum_{j=1}^n R_{jj}^{(\mathbb{J}, \mathbb{K})} - \frac{1}{m} \sum_{j=1}^n R_{jj}^{(\mathbb{J}, \mathbb{K} \cup \{l+n\})}, \\
\varepsilon_{l+n,2}^{(\mathbb{J}, \mathbb{K})} &= \frac{1}{mp} \sum_{j=1}^n (X_{jl}^2 \xi_{jl} - p) R_{jj}^{(\mathbb{J}, \mathbb{K} \cup \{l+n\})}, \\
\varepsilon_{l+n,3}^{(\mathbb{J}, \mathbb{K})} &= \frac{1}{mp} \sum_{1 \leq j \neq k \leq n} X_{jl} X_{kl} \xi_{jl} \xi_{kl} R_{jk}^{(\mathbb{J}, \mathbb{K} \cup \{l+n\})}.
\end{aligned}$$

Theorem 3 *Under the conditions (C0), for any $0 < \gamma < \gamma_0$ and $u_0 > 0$ there exist constants $H = H(\delta, \mu_{4+\delta}, \gamma, u_0)$ and $C = C(\delta, \mu_{4+\delta}, \gamma, u_0)$ such that for all $1 \leq j \leq n$, $1 \leq k \leq m$ and for all $z \in \mathcal{D}_\varepsilon$, we have*

$$\begin{aligned}
\Pr \{ |R_{jk}| > H; \widehat{Q}_\gamma(v) \} &\leq Cn^{-c \log n}, \\
\Pr \{ \max\{|R_{j,k+n}|, |R_{j+n,k}|\} > H; \widehat{Q}_\gamma(v) \} &\leq Cn^{-c \log n}, \\
\Pr \{ |R_{j+n,k+n}| > H; \widehat{Q}_\gamma(v) \} &\leq Cn^{-c \log n}.
\end{aligned}$$

Proof In what follows, we put $Q := \widehat{Q}_\gamma(v)$.

Representation (8) and (9) yield that for all $\gamma \leq \gamma_0$ and all $\mathbb{J}, \mathbb{K}$ satisfying $|u_0|(|\mathbb{J}| + |\mathbb{K}|)/nv \leq 1/4$, the inequalities

$$|R_{jj}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{Q\} \leq 2|\varepsilon_j^{(\mathbb{J}, \mathbb{K})}| |R_{jj}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{Q\} + \frac{2}{\sqrt{y}} \quad (10)$$

and

$$|R_{l+n, l+n}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{Q\} \leq 2|\varepsilon_{l+n}^{(\mathbb{J}, \mathbb{K})}| |R_{l+n, l+n}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{Q\} + 1 + \frac{2}{\sqrt{y}} + \sqrt{y} \quad (11)$$

hold.

Consider the off-diagonal elements of the resolvent matrix. It can be shown that for $j \neq k \in \mathbb{J}^c$

$$R_{jk}^{(\mathbb{J}, \mathbb{K})} = R_{jj}^{(\mathbb{J}, \mathbb{K})} \left(-\frac{1}{\sqrt{mp}} \sum_{l=1}^m X_{jl} \xi_{jl} R_{l+n,k}^{(\mathbb{J} \cup \{j\}, \mathbb{K})} \right),$$

for $j \neq k \in \mathbb{K}^c$

$$R_{j+n,k+n}^{(\mathbb{J}, \mathbb{K})} = R_{j+n,j+n}^{(\mathbb{J}, \mathbb{K})} \left(-\frac{1}{\sqrt{mp}} \sum_{r=1}^n X_{rl} \xi_{rl} R_{k+n,r}^{(\mathbb{J}, \mathbb{K} \cup \{j+n\})} \right),$$

and

$$\begin{aligned} R_{j,k+n}^{(\mathbb{J}, \mathbb{K})} &= R_{jj}^{(\mathbb{J}, \mathbb{K})} \left(-\frac{1}{\sqrt{mp}} \sum_{r=1}^m X_{jr} \xi_{jr} R_{r+n,l+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})} \right), \\ R_{j+n,k}^{(\mathbb{J}, \mathbb{K})} &= R_{kk}^{(\mathbb{J}, \mathbb{K})} \left(-\frac{1}{\sqrt{mp}} \sum_{r=1}^m X_{kr} \xi_{kr} R_{r+n,j+n}^{(\mathbb{J} \cup \{k\}, \mathbb{K})} \right). \end{aligned}$$

Put

$$\begin{aligned} \varepsilon_{jk}^{(\mathbb{J}, \mathbb{K})} &= -\frac{1}{\sqrt{mp}} \sum_{l=1}^m X_{jl} \xi_{jl} R_{l+n,k}^{(\mathbb{J} \cup \{j\}, \mathbb{K})}, \quad \varepsilon_{j+n,k+n}^{(\mathbb{J}, \mathbb{K})} = -\frac{1}{\sqrt{mp}} \sum_{r=1}^n X_{rj} \xi_{rj} R_{r,k+n}^{(\mathbb{J}, \mathbb{K} \cup \{j+n\})}, \\ \varepsilon_{j+n,k}^{(\mathbb{J}, \mathbb{K})} &= -\frac{1}{\sqrt{mp}} \sum_{l=1}^m X_{kl} \xi_{kl} R_{l+n,j+n}^{(\mathbb{J} \cup \{k\}, \mathbb{K})}, \quad \varepsilon_{j,k+n}^{(\mathbb{J}, \mathbb{K})} = -\frac{1}{\sqrt{mp}} \sum_{l=1}^n X_{lk} \xi_{lk} R_{l+n,k+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})}. \end{aligned}$$

Then we get

$$\begin{aligned} R_{jk}^{(\mathbb{J}, \mathbb{K})} &= R_{jj}^{(\mathbb{J}, \mathbb{K})} \varepsilon_{jk}^{(\mathbb{J}, \mathbb{K})}, \quad R_{l+n,k+n}^{(\mathbb{J}, \mathbb{K})} = R_{l+n,l+n}^{(\mathbb{J}, \mathbb{K})} \varepsilon_{l+n,k+n}^{(\mathbb{J}, \mathbb{K})}, \\ R_{jl+n}^{(\mathbb{J}, \mathbb{K})} &= R_{jj}^{(\mathbb{J}, \mathbb{K})} \varepsilon_{j,k+n}^{(\mathbb{J}, \mathbb{K})}, \quad R_{k+n,j}^{(\mathbb{J}, \mathbb{K})} = R_{jj}^{(\mathbb{J}, \mathbb{K})} \varepsilon_{k+n,j}^{(\mathbb{J}, \mathbb{K})}. \end{aligned} \tag{12}$$

Inequalities (10) and (11) imply that

$$\Pr\{|R_{jj}| \mathbb{I}\{\mathbf{Q}\} > C\} \leq \Pr\left\{|\varepsilon_j| \mathbb{I}\{\mathbf{Q}\} > \frac{1}{4}\right\} \tag{13}$$

for $1 \leq j \leq n$ and $C > \frac{4}{\sqrt{y}}$, and

$$\Pr\{|R_{l+n,l+n}| \mathbb{I}\{\mathbf{Q}\} > C\} \leq \Pr\left\{|\varepsilon_{l+n}| \mathbb{I}\{\mathbf{Q}\} > \frac{1}{4}\right\} \tag{14}$$

for $1 \leq l \leq m$ and $C > 2(2 + 3\sqrt{y} + \frac{2}{\sqrt{y}})$. Relations (12) give

$$\Pr\{|R_{jk}|\mathbb{I}\{\mathbf{Q}\} > C\} \leq \Pr\{|R_{jj}|\mathbb{I}\{\mathbf{Q}\} > C\} + \Pr\{|\varepsilon_{jk}|\mathbb{I}\{\mathbf{Q}\} > 1\}$$

for $1 \leq j \neq k \leq n$, and

$$\Pr\{|R_{l+n,k+n}|\mathbb{I}\{\mathbf{Q}\} > C\} \leq \Pr\{|R_{l+n,l+n}|\mathbb{I}\{\mathbf{Q}\} > C\} + \Pr\{|\varepsilon_{l+n,k+n}|\mathbb{I}\{\mathbf{Q}\} > 1\}$$

for $1 \leq l \neq k \leq m$. Similarly we obtain

$$\Pr\{|R_{l,k+n}|\mathbb{I}\{\mathbf{Q}\} > C\} \leq \Pr\{|R_{l,l}|\mathbb{I}\{\mathbf{Q}\} > C\} + \Pr\{|\varepsilon_{l,k+n}|\mathbb{I}\{\mathbf{Q}\} > 1\}$$

and

$$\Pr\{|R_{l+n,k}|\mathbb{I}\{\mathbf{Q}\} > C\} \leq \Pr\{|R_{k,k}|\mathbb{I}\{\mathbf{Q}\} > C\} + \Pr\{|\varepsilon_{l+n,k}|\mathbb{I}\{\mathbf{Q}\} > 1\}.$$

By Rosenthal's inequality (see [11], Lemma 8.1) we find that

$$\mathbb{E}_j |\varepsilon_{jk}|^q \leq C^q \left(q^{\frac{q}{2}} (nv)^{-\frac{q}{2}} (\operatorname{Im} R_{kk}^{(j)})^{\frac{q}{2}} + q^q (np)^{-q\kappa-1} \frac{1}{n} \sum_{l=1}^m |R_{k,l+n}^{(j)}|^q \right)$$

for $1 \leq j \neq k \leq n$, and

$$\begin{aligned} \mathbb{E}_{j+n} |\varepsilon_{j+n,k+n}|^q &\leq C^q \left(q^{\frac{q}{2}} (nv)^{-\frac{q}{2}} (\operatorname{Im} R_{k+n,k+n}^{(j+n)})^{\frac{q}{2}} \right. \\ &\quad \left. + q^q (np)^{-q\kappa-1} \frac{1}{n} \sum_{r=1}^n |R_{k+n,r}^{(j+n)}|^q \right), \\ \mathbb{E}_j |\varepsilon_{j,k+n}|^q &\leq C^q \left(q^{\frac{q}{2}} (nv)^{-\frac{q}{2}} (\operatorname{Im} R_{k+n,k+n}^{(j+n)})^{\frac{q}{2}} \right. \\ &\quad \left. + q^q (np)^{-q\kappa-1} \frac{1}{n} \sum_{r=1}^n |R_{k+n,r+n}^{(j+n)}|^q \right), \\ \mathbb{E}_{j+n} |\varepsilon_{j+n,k}|^q &\leq C^q \left(q^{\frac{q}{2}} (nv)^{-\frac{q}{2}} (\operatorname{Im} R_{k+n,k+n}^{(j+n)})^{\frac{q}{2}} \right. \\ &\quad \left. + q^q (np)^{-q\kappa-1} \frac{1}{n} \sum_{r=1}^n |R_{k+n,r}^{(j+n)}|^q \right) \end{aligned}$$

for $1 \leq j \neq k \leq m$. Note that

$$\begin{aligned}
\Pr\{|\varepsilon_j^{(\mathbb{J}, \mathbb{K})}| > \frac{1}{4}; \mathcal{Q}\} &\leq \Pr\{\mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K})^c; \mathcal{Q}\} \\
&\quad + \Pr\{|\varepsilon_j^{(\mathbb{J}, \mathbb{K})}| > \frac{1}{4}; \mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K}); \mathcal{Q}\}, \\
\Pr\{|\varepsilon_{j+n}^{(\mathbb{J}, \mathbb{K})}| > \frac{1}{4}; \mathcal{Q}\} &\leq \Pr\{\mathcal{A}^{(1)}(sv, \mathbb{J}, \mathbb{K})^c; \mathcal{Q}\} \\
&\quad + \Pr\{|\varepsilon_{j+n}^{(\mathbb{J}, \mathbb{K})}| > 1/4; \mathcal{A}^{(1)}(sv, \mathbb{J}, \mathbb{K}); \mathcal{Q}\}, \\
\Pr\{|\varepsilon_{jk}^{(\mathbb{J}, \mathbb{K})}| > 1; \mathcal{Q}\} &\leq \Pr\{\mathcal{A}^{(2)}(sv, \mathbb{J}, \mathbb{K})^c; \mathcal{Q}\} \\
&\quad + \Pr\{|\varepsilon_{jk}^{(\mathbb{J}, \mathbb{K})}| > 1; \mathcal{A}^{(2)}(sv, \mathbb{J}, \mathbb{K}); \mathcal{Q}\}, \\
\Pr\{|\varepsilon_{l+n, k+n}^{(\mathbb{J}, \mathbb{K})}| > 1; \mathcal{Q}\} &\leq \Pr\{\mathcal{A}^{(3)}(sv, \mathbb{J}, \mathbb{K})^c; \mathcal{Q}\} \\
&\quad + \Pr\{|\varepsilon_{l+n, k+n}^{(\mathbb{J}, \mathbb{K})}| > 1; \mathcal{A}^{(3)}(sv, \mathbb{J}, \mathbb{K}); \mathcal{Q}\}, \\
\Pr\{|\varepsilon_{j+n, k}^{(\mathbb{J}, \mathbb{K})}| > 1; \mathcal{Q}\} &\leq \Pr\{\mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K})^c; \mathcal{Q}\} \\
&\quad + \Pr\{|\varepsilon_{j+n, k}^{(\mathbb{J}, \mathbb{K})}| > 1; \mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K}); \mathcal{Q}\}, \\
\Pr\{|\varepsilon_{k, j+n}^{(\mathbb{J}, \mathbb{K})}(v)| > 1; \mathcal{Q}\} &\leq \Pr\{\mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K})^c; \mathcal{Q}\} \\
&\quad + \Pr\{|\varepsilon_{k, l+n}^{(\mathbb{J}, \mathbb{K})}(v)| > 1; \mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K}); \mathcal{Q}\}.
\end{aligned}$$

Using Chebyshev's inequality we get

$$\Pr\{|\varepsilon_j^{(\mathbb{J}, \mathbb{K})}| > 1/4; \mathcal{Q}; \mathcal{A}^{(4)}\} \leq C^q \mathbb{E} \left(\mathbb{E}_j |\varepsilon_j|^q \right) \mathbb{I}\{\mathcal{Q}^{(\mathbb{J}, \mathbb{K})}\} \mathbb{I}\{\mathcal{A}^{(4)}\}.$$

Applying the triangle inequality, the results of Lemmas 5–8, the property of the multiplicative gradient descent of the resolvent, we arrive at the inequality

$$\begin{aligned}
\mathbb{E}_j \mathbb{I}\{\mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K})\} |\varepsilon_j|^q &\leq C^q \left[\frac{1}{(nv)^q} + \left(\frac{qs}{np} \right)^{\frac{q}{2}} + \frac{1}{np} \left(\frac{qs}{(np)^{2\kappa}} \right)^q \right. \\
&\quad \left. + \left(\frac{q^2}{nv} \right)^{\frac{q}{2}} + \frac{1}{np} \left(\frac{qs}{(np)^{2\kappa}} \right)^{\frac{q}{2}} \left(\frac{q^2}{nv} \right)^{\frac{q}{2}} + \left(\frac{q^2 s}{(np)^{2\kappa}} \right)^q \frac{1}{(np)^2} \right].
\end{aligned}$$

If we set $q \sim \log^2 n$, $nv > C \log^4 n$, $np > C \log^{\frac{2}{\kappa}} n$, we get

$$\mathbb{E}_j |\varepsilon_j|^q \mathbb{I}\{\mathcal{A}^{(4)}(sv, \mathbb{J} \cup \{j\}, \mathbb{K})\} \leq C n^{-c \log n}.$$

Moreover, the constant c can be made arbitrarily large. We may obtain similar estimates for the quantities ε_{l+n} , ε_{jk} , $\varepsilon_{j+n, k}$, $\varepsilon_{j, k+n}$, $\varepsilon_{j+n, k+n}$. Inequalities (13), (14) imply

$$\begin{aligned}
\Pr\{|R_{jj}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{\mathbf{Q}\} > C\} &\leq \Pr\{\mathcal{A}^{(4)}(sv, \mathbb{J} \cup \{j\}, \mathbb{K})^c\} + Cn^{-c \log n}, \\
\Pr\{|R_{l+n, l+n}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{\mathbf{Q}\} > C\} &\leq \Pr\{\mathcal{A}^{(1)}(sv, \mathbb{J}, \mathbb{K} \cup \{l\})^c\} + Cn^{-c \log n}, \\
\Pr\{|R_{jk}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{\mathbf{Q}\} > C\} &\leq \Pr\{\mathcal{A}^{(2)}(sv, \mathbb{J}, \mathbb{K} \cup \{l\})^c\} + Cn^{-c \log n}, \\
\Pr\{|R_{j+n, k}| \mathbb{I}\{\mathbf{Q}\} > C\} &\leq \Pr\{\mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K} \cup \{j\})^c\} + Cn^{-c \log n}, \\
\Pr\{|R_{k+n, j}| \mathbb{I}\{\mathbf{Q}\} > C\} &\leq \Pr\{\mathcal{A}^{(4)}(sv, \mathbb{J}, \mathbb{K} \cup \{j\})^c\} + Cn^{-c \log n}, \\
\Pr\{|R_{k+n, j+n}| \mathbb{I}\{\mathbf{Q}\} > C\} &\leq \Pr\{\mathcal{A}^{(3)}(sv, \mathbb{J}, \mathbb{K} \cup \{j\})^c\} + Cn^{-c \log n}.
\end{aligned}$$

The last inequalities give

$$\begin{aligned}
\max_{j, k \in \mathbb{J}^c \cup \mathbb{K}^c} \Pr\{|R_{j, k}^{(\mathbb{J}, \mathbb{K})}| \mathbb{I}\{\mathbf{Q}\} > C\} &\leq Cn^{-c \log n} \\
&+ \max_{j \in \mathbb{J}^c, k \in \mathbb{K}^c} \max\{\Pr\{\mathcal{A}^c(s_0 v, \mathbb{J} \cup \{j\}, \mathbb{K})\}, \Pr\{\mathcal{A}^c(s_0 v, \mathbb{J}, \mathbb{K} \cup \{k\})\}\}.
\end{aligned}$$

Note that $k_v \leq C \log n$ for $v \geq v_0 = n^{-1} \log^4 n$. So, choosing c large enough, we get

$$\Pr\{\mathcal{A}^c(v) \cap \mathbf{Q}\} \leq Cn^{-c \log n}.$$

This completes the proof of the theorem. $\square$

Corollary 5 *Under the conditions of Theorem 3, for any $v \geq v_0$ and $q \leq c \log n$ there exists a constant H such that*

$$\mathbb{E} |R_{jk}|^q \mathbb{I}\{\mathbf{Q}\} \leq H^q,$$

for $j, k \in \mathbb{T} \cup (\mathbb{T}^{(1)} + n)$.

Proof We may write

$$\mathbb{E} |R_{jk}|^q \mathbb{I}\{\mathbf{Q}\} \leq \mathbb{E} |R_{jk}|^q \mathbb{I}\{\mathbf{Q}\} \mathbb{I}\{\mathcal{A}(v)\} + \mathbb{E} |R_{jk}|^q \mathbb{I}\{\mathbf{Q}\} \mathbb{I}\{\mathcal{A}^c(v)\}.$$

Combine this inequality with $|R_{jk}| \leq v^{-1}$, we find that

$$\mathbb{E} |R_{jk}|^q \mathbb{I}\{\mathbf{Q}\} \leq C^q + v_0^{-q} \mathbb{E} \mathbb{I}\{\mathbf{Q}\} \mathbb{I}\{\mathcal{A}^c(v)\}.$$

Applying Theorem 3, we obtain what is required. $\square$

5.2 Estimation of T_n

Recall that

$$a := a_n(z) = \operatorname{Im} b(z) + \frac{\log(n) \log(np)}{np|b(z)|}.$$

Consider the smoothing of the indicator $h_\gamma(x)$:

$$h_\gamma(x, v) = \begin{cases} 1, & \text{for } |x| \leq \gamma a, \\ 1 - \frac{||x| - \gamma a|}{\gamma a}, & \text{for } \gamma a \leq |x| \leq 2\gamma a, \\ 0, & \text{for } |x| > 2\gamma a. \end{cases}$$

Note that

$$\mathbb{I}_{\widehat{\mathcal{Q}}_\gamma(v)} \leq h_\gamma(|\Lambda_n(u + iv)|, v) \leq \mathbb{I}_{\widehat{\mathcal{Q}}_{2\gamma}(v)},$$

where, as before,

$$\widehat{\mathcal{Q}}_\gamma(v) = \bigcap_{v=0}^{k_v} \{|\Lambda_n(u + is_0^v v)| \leq \gamma a_n(u, s_0^v v)\}.$$

We will estimate the value

$$D_n := \mathbb{E} |T_n|^q h_\gamma^q(|\Lambda_n|, v).$$

It is easy to see that

$$\mathbb{E} |T_n|^q \mathbb{I}\{Q\} \leq D_n.$$

To estimate D_n we use the approach developed in [13–16] that goes back to Stein's method. Let

$$\varphi(z) := \bar{z}|z|^{q-2}.$$

Set

$$\widehat{T}_n := T_n h_\gamma(|\Lambda_n|, v).$$

Then we can write

$$D_n := \mathbb{E} \widehat{T}_n \varphi(\widehat{T}_n).$$

The equality

$$T_n = 1 + \left(z - \frac{1-y}{z}\right) s_n(z) + y s_n^2(z) = b(z) \Lambda_n(z) + y \Lambda_n^2(z)$$

implies that for $z \in \mathcal{D}_\varepsilon$

$$|T_n| \mathbb{I}\{Q\} \leq C.$$

Consider

$$\mathcal{B} := \mathcal{A}^{(1)} \cap \mathcal{A}^{(2)} \cap \mathcal{A}^{(3)} \cap \mathcal{A}^{(4)},$$

where $\mathcal{A}^{(v)}$ were defined in (6) and (7). Then

$$D_n \leq \mathbb{E} |T_n|^q \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} + Cn^{-c \log n}.$$

By the definition of T_n , we may rewrite the last inequality as

$$D_n := \frac{1}{n} \sum_{j=1}^n \mathbb{E} \varepsilon_j R_{jj} h_\gamma(|\Lambda_n|, v) \varphi(\widehat{T}_n) \mathbb{I}\{\mathcal{B}\} + Cn^{-c \log n}.$$

Set

$$D_n = D_n^{(1)} + D_n^{(2)} + Cn^{-c \log n}, \quad (15)$$

where

$$\begin{aligned} D_n^{(1)} &:= \frac{1}{n} \sum_{j=1}^n \mathbb{E} \varepsilon_{j1} R_{jj} h_\gamma(|\Lambda_n|, v) \varphi(\widehat{T}_n) \mathbb{I}\{\mathcal{B}\}, \\ D_n^{(2)} &:= \frac{1}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j R_{jj} h_\gamma(|\Lambda_n|, v) \varphi(\widehat{T}_n) \mathbb{I}\{\mathcal{B}\}, \\ \widehat{\varepsilon}_j &:= \varepsilon_{j2} + \varepsilon_{j3}. \end{aligned}$$

We have

$$\frac{1}{n} \sum_{j=1}^n \varepsilon_{j1} R_{jj} = \frac{1}{2n} s'_n(z) + \frac{s_n(z)}{2nz}$$

and this yields

$$\left| \frac{1}{n} \sum_{j=1}^n \varepsilon_{j1} R_{jj} \right| \leq \frac{C}{nv} \operatorname{Im} s_n(z) + \frac{C}{n}.$$

Using Hölder's inequality, we get

$$|D_n^{(1)}| \leq \frac{C a_n(z)}{nv} D_n^{\frac{q-1}{q}}. \quad (16)$$

Further, consider

$$\widehat{T}_n^{(j)} = \mathbb{E}_j \widehat{T}_n, \quad T_n^{(j)} = \mathbb{E}_j T_n, \quad \Lambda_n^{(j)} = \mathbb{E}_j \Lambda_n.$$

Represent $D_n^{(2)}$ in the form

$$D_n^{(2)} = D_n^{(21)} + \dots + D_n^{(24)}, \quad (17)$$

where

$$\begin{aligned} D_n^{(21)} &:= \frac{S_y(z)}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j h_\gamma(|\Lambda_n^{(j)}|, v) \varphi(\widehat{T}_n^{(j)}) \mathbb{I}\{\mathcal{B}\}, \\ D_n^{(22)} &:= \frac{1}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j (R_{jj} - S_y(z)) h_\gamma(|\Lambda_n^{(j)}|, v) \varphi(\widehat{T}_n^{(j)}) \mathbb{I}\{\mathcal{B}\}, \\ D_n^{(23)} &:= \frac{1}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j R_{jj} (h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)) \varphi(\widehat{T}_n^{(j)}) \mathbb{I}\{\mathcal{B}\}, \\ D_n^{(24)} &:= \frac{1}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j R_{jj} h_\gamma(|\Lambda_n|, v) (\varphi(\widehat{T}_n) - \varphi(\widehat{T}_n^{(j)})) \mathbb{I}\{\mathcal{B}\}. \end{aligned}$$

Since $E_j \widehat{\varepsilon}_j = 0$, we find

$$D_n^{(21)} = \frac{S_y(z)}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j h_\gamma(|\Lambda_n^{(j)}|, v) \varphi(\widehat{T}_n^{(j)}) \mathbb{I}\{\mathcal{B}^c\}.$$

From here it is easy to obtain that

$$|D_n^{(21)}| \leq C n^{-c \log n}. \quad (18)$$

Note that Eq. (4) implies

$$\Lambda_n = \frac{T_n}{b_n(z)},$$

where

$$b_n(z) = z - \frac{1-y}{z} + y S_y(z) + y s_n(z).$$

The last can also be rewritten as

$$b_n(z) = z - \frac{1-y}{z} + 2y S_y(z) + y \Lambda_n(z) = b(z) + y \Lambda_n(z).$$

5.2.1 Estimation of $D_n^{(22)}$

Using the representation of R_{jj} , we may write

$$D_n^{(22)} = \tilde{D}_n^{(22)} + \widehat{D}_n^{(22)},$$

where

$$\begin{aligned}\tilde{D}_n^{(22)} &:= \frac{S_y(z)}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j^2 R_{jj} h_\gamma(|\Lambda_n^{(j)}|, \nu) \varphi(\widehat{T}_n^{(j)}) \mathbb{I}\{\mathcal{B}\}, \\ \widehat{D}_n^{(22)} &:= \frac{y S_y(z)}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j \Lambda_n R_{jj} h_\gamma(|\Lambda_n^{(j)}|, \nu) \varphi(\widehat{T}_n^{(j)}) \mathbb{I}\{\mathcal{B}\}.\end{aligned}$$

By Hölder's inequality,

$$|\widehat{D}_n^{(22)}| \leq \frac{C}{n} \sum_{j=1}^n \mathbb{E}^{\frac{1}{q}} \left[\mathbb{E}_j |\widehat{\varepsilon}_j| |\Lambda_n| |R_{jj}| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right]^q D_n^{\frac{q-1}{q}}. \quad (19)$$

Further,

$$\mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| |R_{jj}| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right] \leq C \mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right].$$

It is obvious that

$$\begin{aligned}\mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right] \\ \leq \mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right] \mathbb{I}\{|b(z)| \leq \sqrt{|T_n|}\} \\ + \mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right] \mathbb{I}\{|b(z)| > \sqrt{|T_n|}\}.\end{aligned}$$

We have

$$\begin{aligned}|\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} &\leq |\Lambda_n| h_\gamma(|\Lambda_n|, \nu) \mathbb{I}\{\mathcal{B}\} \\ &+ |\Lambda_n| |h_\gamma(|\Lambda_n^{(j)}|, \nu) - h_\gamma(|\Lambda_n|, \nu)| \mathbb{I}\{\mathcal{B}\}.\end{aligned}$$

Since we have $|a_n| \leq |b(z)|$ for $z \in \mathcal{D}_\varepsilon$ in the last inequality, we can write

$$|b_n(z)| h(|\Lambda_n^{(j)}|, \nu) \geq \left(|b(z)| - \gamma a_n(z) - \frac{1}{nv} \right) \geq c |b(z)|$$

and therefore

$$|\Lambda_n| \leq \frac{C |T_n|}{|b(z)|}.$$

From here get

$$|\Lambda_n| h_\gamma(|\Lambda_n|, \nu) \mathbb{I}\{\mathcal{B}\} \mathbb{I}\{\sqrt{|T_n|} \leq |b(z)|\} \leq C \sqrt{|T_n|} h(|\Lambda_n|, \nu).$$

Further, we note that if $\sqrt{|T_n|} \geq |b(z)|$, then

$$\begin{aligned} |\Lambda_n| h_\gamma(|\Lambda_n|, \nu) &\leq \gamma a_n(z) h_\gamma(|\Lambda_n|, \nu) \leq C |b(z)| h_\gamma(|\Lambda_n|, \nu) \\ &\leq C \sqrt{|T_n|} h_\gamma(|\Lambda_n|, \nu). \end{aligned}$$

Using this we conclude that

$$\begin{aligned} \mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right] &\leq \mathbb{E}_j^{\frac{1}{2}} |\widehat{\varepsilon}_j|^2 \mathbb{I}\{|\Lambda_n^{(j)}| \leq C a_n(z)\} \mathbb{I}\{\mathcal{B}\} \mathbb{E}_j^{\frac{1}{2}} |\widehat{T}_n| \\ &+ \frac{C}{a_n(z)} \mathbb{E}_j^{\frac{1}{2}} |\widehat{\varepsilon}_j|^2 \mathbb{I}\{|\Lambda_n^{(j)}| \leq C a_n(z)\} \mathbb{I}\{\mathcal{B}\} \\ &\quad \times \mathbb{E}_j^{\frac{1}{2}} |\Lambda_n|^2 |\Lambda_n - \Lambda_n^{(j)}|^2 \mathbb{I}\left\{ \max\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq C \left(a_n(z) + \frac{1}{nv} \right) \right\}. \end{aligned}$$

Applying Lemmas 7 and 8, we obtain

$$\mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right] \leq C \left(\frac{(a_n(z))^{\frac{1}{2}}}{(nv)^{\frac{1}{2}}} + \frac{1}{(np)^{\frac{1}{2}}} \right) \left(\mathbb{E}_j^{\frac{1}{2}} |\widehat{T}_n| + \frac{a_n(z) + \frac{1}{nv}}{nv a_n(z)} \right).$$

Using that $a_n(z) \geq C/nv$ (see Lemma 3), we get

$$\mathbb{E}_j \left[|\widehat{\varepsilon}_j| |\Lambda_n| h_\gamma(|\Lambda_n^{(j)}|, \nu) \mathbb{I}\{\mathcal{B}\} \right] \leq C \left(\frac{(a_n(z))^{\frac{1}{2}}}{(nv)^{\frac{1}{2}}} + \frac{1}{(np)^{\frac{1}{2}}} \right) \left(\mathbb{E}_j^{\frac{1}{2}} |\widehat{T}_n| + \frac{1}{nv} \right). \quad (20)$$

Combining inequalities (19), (20), and $\frac{1}{nv} \leq C \sqrt{\frac{a_n(z)}{nv}}$, we get

$$|D_n^{(22)}| \leq C (\beta_n D_n^{\frac{q-1}{q}} + \sqrt{\beta_n} D_n^{\frac{2q-1}{2q}}). \quad (21)$$

5.2.2 Estimation of $D_n^{(23)}$

Note that

$$\begin{aligned} |h_\gamma(|\Lambda_n|, \nu) - |h_\gamma(|\Lambda_n^{(j)}|, \nu)| |R_{jj}| \mathbb{I}\{\mathcal{B}\} \\ \leq \frac{C}{a_n(z)} |\Lambda_n - \Lambda_n^{(j)}| \mathbb{I}\{\max\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq 2\gamma a_n(z)\} \mathbb{I}\{\mathcal{B}\}. \end{aligned}$$

Using Hölder's inequality and Cauchy's inequality, we get

$$D_n^{(23)} \leq \frac{C}{na_n(z)} \sum_{j=1}^n \mathbb{E}_j^{\frac{1}{q}} \{ [\mathbb{E}_j |\widehat{\varepsilon}_j|^2 \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\}]^{\frac{q}{2}} [\mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^2 \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\}]^{\frac{q}{2}} \} D_n^{\frac{q-1}{q}}.$$

Applying Lemmas 7, 8, 10, obtain

$$\begin{aligned} D_n^{(23)} &\leq Ca_n^{-1}(z) \left(\frac{C}{\sqrt{np}} + \frac{\sqrt{a_n(z)}}{\sqrt{nv}} \right) \frac{Ca_n^{\frac{1}{2}}(z)}{\sqrt{nv}} \left[\frac{a_n(z)}{nv} \right. \\ &\quad \left. + \frac{Ca_n^{\frac{1}{2}}(z)}{(nv)^{\frac{1}{2}}(np)^{\frac{1}{2}}} + \frac{C}{(nv)^{\frac{1}{2}}(np)^{\frac{1}{2}}} + n^{-c \log n} \right] D_n^{\frac{q-1}{q}}. \end{aligned}$$

Taking into account that $\text{Im } b(z) \geq \frac{C}{nv}$ ($a_n(z)nv \geq C$), the last gives the bound

$$D_n^{(23)} \leq C \left(\frac{a_n(z)}{nv} + \frac{C}{np} \right) D_n^{\frac{q-1}{q}} = \beta_n D_n^{\frac{q-1}{q}}.$$

5.2.3 Estimation of $D_n^{(24)}$

By Taylor's theorem, we have

$$D_n^{(24)} = \frac{1}{n} \sum_{j=1}^n \mathbb{E} \widehat{\varepsilon}_j R_{jj} h_\gamma(|\Lambda_n|, v) (\widehat{T}_n - \widehat{T}_n^{(j)}) \varphi'(\widehat{T}_n^{(j)}) + \tau(\widehat{T}_n - \widehat{T}_n^{(j)}) \mathbb{I}\{\mathcal{B}\},$$

where τ is uniform distributed on the interval $[0, 1]$ random variables independent on all other. Since $\mathbb{I}\{\mathcal{B}\} = 1$ yields $|R_{jj}| \leq C$, we find that

$$|D_n^{(24)}| \leq \frac{C}{n} \sum_{j=1}^n \mathbb{E} |\widehat{\varepsilon}_j| h_\gamma(|\Lambda_n|, v) |\widehat{T}_n - \widehat{T}_n^{(j)}| |\varphi'(\widehat{T}_n^{(j)}) + \tau(\widehat{T}_n - \widehat{T}_n^{(j)})| \mathbb{I}\{\mathcal{B}\}.$$

Taking into account the inequality

$$|\varphi'(\widehat{T}_n^{(j)}) + \tau(\widehat{T}_n - \widehat{T}_n^{(j)})| \leq Cq \left[|\widehat{T}_n^{(j)}|^{q-2} + q^{q-2} |\widehat{T}_n - \widehat{T}_n^{(j)}|^{q-2} \right],$$

obtain

$$\begin{aligned} |D_n^{(24)}| &\leq \frac{Cq}{n} \sum_{j=1}^n \mathbb{E} |\widehat{\varepsilon}_j| h_\gamma(|\Lambda_n|, v) |\widehat{T}_n - \widehat{T}_n^{(j)}| |\widehat{T}_n^{(j)}|^{q-2} \mathbb{I}\{\mathcal{B}\} \\ &\quad + \frac{Cq^{q-1}}{n} \sum_{j=1}^n \mathbb{E} |\widehat{\varepsilon}_j| h_\gamma(|\Lambda_n|, v) |\widehat{T}_n - \widehat{T}_n^{(j)}|^{q-1} \mathbb{I}\{\mathcal{B}\} =: \widehat{D}_n^{(24)} + \widetilde{D}_n^{(24)}. \end{aligned}$$

Applying Hölder's inequality, get

$$\widehat{D}_n^{(24)} \leq \frac{Cq}{n} \sum_{j=1}^n \mathbb{E}^{\frac{2}{q}} \left[\mathbb{E}_j \{ |\widehat{\varepsilon}_j| h_\gamma(|\Lambda_n|, \nu) |\widehat{T}_n - \widehat{T}_n^{(j)}| \mathbb{I}\{\mathcal{B}\} \} \right]^{\frac{q}{2}} \mathbb{E}^{\frac{q-2}{q}} |\widehat{T}_n^{(j)}|^q.$$

Jensen's inequality gives

$$\widehat{D}_n^{(24)} \leq \frac{Cq}{n} \sum_{j=1}^n \mathbb{E}^{\frac{2}{q}} \left[\mathbb{E}_j \{ |\widehat{\varepsilon}_j| h_\gamma(|\Lambda_n|, \nu) |\widehat{T}_n - \widehat{T}_n^{(j)}| \mathbb{I}\{\mathcal{B}\} \} \right]^{\frac{q}{2}} D_n^{\frac{q-2}{q}}.$$

To estimate $\widehat{D}_n^{(24)}$ we have to get bound for

$$V_j^{\frac{q}{2}} := \mathbb{E} \left[\mathbb{E}_j \{ |\widehat{\varepsilon}_j| h_\gamma(|\Lambda_n|, \nu) |\widehat{T}_n - \widehat{T}_n^{(j)}| \mathbb{I}\{\mathcal{B}\} \} \right]^{\frac{q}{2}}.$$

By Cauchy's inequality, we have

$$V_j^{\frac{q}{2}} \leq \mathbb{E}(V_j^{(1)})^{\frac{q}{4}} (V_j^{(2)})^{\frac{q}{4}}, \quad (22)$$

where

$$V_j^{(1)} := \mathbb{E}_j |\widehat{\varepsilon}_j|^2 \mathbb{I}\{\widehat{\mathcal{Q}}_{2\gamma}(\nu)\} \mathbb{I}\{\mathcal{B}\}, \quad V_j^{(2)} := \mathbb{E}_j |\widehat{T}_n - \widehat{T}_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, \nu) \mathbb{I}\{\mathcal{B}\}.$$

Estimation of $V_j^{(1)}$

Lemma 7 gives

$$\mathbb{E}_j |\varepsilon_{j2}|^2 \mathbb{I}\{\widehat{\mathcal{Q}}_{2\gamma}(\nu)\} \mathbb{I}\{\mathcal{B}\} \leq \frac{C}{np},$$

and Lemma 8, in its turn, gives

$$\mathbb{E}_j |\varepsilon_{j3}|^2 \mathbb{I}\{\widehat{\mathcal{Q}}_{2\gamma}(\nu)\} \mathbb{I}\{\mathcal{B}\} \leq \frac{C}{nv} a_n(z).$$

Summing up the obtained estimates, we arrive at the following inequality

$$V_j^{(1)} \leq \frac{C a_n(z)}{nv} + \frac{C}{np} =: \beta_n(z). \quad (23)$$

Estimation of $V_j^{(2)}$

To estimate $V_j^{(2)}$ we consider $\widehat{T}_n - \widehat{T}_n^{(j)}$. Since $\widehat{T}_n = T_n h_\gamma(|\Lambda_n|, \nu)$ and $\widehat{T}_n^{(j)} = \mathbb{E}_j \widehat{T}_n$, we have

$$\begin{aligned}\widehat{T}_n - \widehat{T}_n^{(j)} &= (T_n - T_n^{(j)})h_\gamma(|\Lambda_n|, v) - \mathbb{E}_j \left[(T_n - T_n^{(j)})h_\gamma(|\Lambda_n|, v) \right] \\ &\quad + T_n^{(j)} \left([h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)] - \mathbb{E}_j \left[h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v) \right] \right).\end{aligned}$$

Further, we note that

$$T_n = \Lambda_n b_n = \Lambda_n b(z) + y \Lambda_n^2.$$

Then

$$T_n - T_n^{(j)} = (\Lambda_n - \Lambda_n^{(j)})(b(z) + 2y\Lambda_n^{(j)}) + y(\Lambda_n - \Lambda_n^{(j)})^2 - y \mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})^2. \quad (24)$$

We get

$$\begin{aligned}\widehat{T}_n - \widehat{T}_n^{(j)} &= (b(z) + 2y\Lambda_n^{(j)}) \left[(\Lambda_n - \Lambda_n^{(j)})h_\gamma(|\Lambda_n|, v) - \mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})h_\gamma(|\Lambda_n|, v) \right] \\ &\quad + y \left[(\Lambda_n - \Lambda_n^{(j)})^2 - \mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})^2 \right] h_\gamma(|\Lambda_n|, v) \\ &\quad - y \mathbb{E}_j \left[(\Lambda_n - \Lambda_n^{(j)})^2 (h_\gamma(|\Lambda_n|, v) - \mathbb{E}_j h_\gamma(|\Lambda_n|, v)) \right] \\ &\quad - T_n^{(j)} \mathbb{E}_j \left[h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v) \right].\end{aligned} \quad (25)$$

Now return to estimation of $V_j^{(2)}$. Equality (25) implies

$$\begin{aligned}V_j^{(2)} &\leq 4|b(z)|^2 \mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^2 h_\gamma^4(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + 4|b(z)|^2 \left[\mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})h_\gamma(|\Lambda_n|, v) \right]^2 \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + 8y^2 \mathbb{E}_j |\Lambda_n^{(j)}|^2 |\Lambda_n - \Lambda_n^{(j)}|^2 h_\gamma^4(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + 8y^2 \left[\mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})h_\gamma(|\Lambda_n|, v) \right]^2 \mathbb{E}_j |\Lambda_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + 2y^2 \left[\mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^4 h_\gamma^4(|\Lambda_n|, v) \right] \mathbb{I}\{\mathcal{B}\} \\ &\quad + 2y^2 \left[\mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})^2 h_\gamma(|\Lambda_n|, v) \right]^2 \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + y^2 \left[\mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})^2 h_\gamma(|\Lambda_n|, v) \right]^2 \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + y^2 \left[\mathbb{E}_j(\Lambda_n - \Lambda_n^{(j)})^2 \right]^2 \left[\mathbb{E}_j h_\gamma(|\Lambda_n|, v) \right]^2 \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + |T_n^{(j)}|^2 \mathbb{E}_j \left[h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v) \right]^2 h_\gamma^2(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \\ &\quad + |T_n^{(j)}|^2 \left[\mathbb{E}_j (h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)) \right]^2 \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\}.\end{aligned}$$

We can rewrite it as

$$V_j^{(2)} \leq A_1 + A_2 + A_3 + A_4,$$

$$\begin{aligned}
A_1 &= C|b(z)|^2 \mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v) (h_\gamma^2(|\Lambda_n|, v) + \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v)) \mathbb{I}\{\mathcal{B}\}, \\
A_2 &= C \mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v) (|\Lambda_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v) + \mathbb{E}_j |\Lambda_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v)) \mathbb{I}\{\mathcal{B}\}, \\
A_3 &= C \mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^4 h_\gamma^2(|\Lambda_n|, v) (h_\gamma^2(|\Lambda_n|, v) + \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v)) \mathbb{I}\{\mathcal{B}\}, \\
A_4 &= C |T_n^{(j)}|^2 \mathbb{E}_j |h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)|^2 (h_\gamma^2(|\Lambda_n|, v) + \mathbb{E}_j h_\gamma^2(|\Lambda_n|, v)) \mathbb{I}\{\mathcal{B}\}.
\end{aligned}$$

Applying $h_\gamma(|\Lambda_n|, v) \leq 1$ and Lemma 10, we find that

$$A_1 \leq C|b(z)|^2 \frac{a_n(z)}{nv} \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{1}{(nv)^2} \right)$$

and

$$A_2 \leq C \frac{a_n^3(z)}{nv} \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{1}{(nv)^2} \right).$$

Note that

$$A_3 \leq \frac{C}{n^2 v^2} \mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v).$$

Combining Lemma 10 and $a_n(z) \geq \frac{1}{nv}$, get

$$A_3 \leq \frac{C a_n(z)}{(nv)^3} \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{1}{(nv)^2} \right).$$

Applying the triangle inequality, we obtain

$$\begin{aligned}
A_4 &= C \mathbb{E}_j |h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)|^2 (|T_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v) + \mathbb{E}_j |T_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v)) \\
&\leq C \mathbb{E}_j |h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)|^2 (\mathbb{E}_j |\widehat{T}_n|^2 + \mathbb{E}_j |T_n - T_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v)) \\
&\quad + C \mathbb{E}_j |h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)|^2 (|\widehat{T}_n|^2 + |T_n - T_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, v)).
\end{aligned}$$

Further, since

$$\begin{aligned}
|h_\gamma(|\Lambda_n|, v) - h_\gamma(|\Lambda_n^{(j)}|, v)| &\leq \frac{C}{\gamma a_n(z)} |\Lambda_n - \Lambda_n^{(j)}| \\
&\quad \times \mathbb{I}\{\min\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq 2\gamma a_n(z)\}, \quad (26)
\end{aligned}$$

we may write

$$\begin{aligned}
A_4 &\leq \frac{C}{a_n^2(z)} \mathbb{E}_j(|\widehat{T}_n|^2) + \mathbb{E}_j(|\widehat{T}_n|^2)|\Lambda_n - \Lambda_n^{(j)}|^2 \mathbb{I}\{\min\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq Ca_n(z)\} \\
&\quad + \frac{C}{a_n^2(z)} \mathbb{E}_j|\Lambda_n - \Lambda_n^{(j)}|^2 \mathbb{I}\{\min\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq 2\gamma a_n(z)\} \\
&\quad \times (h_\gamma^2(|\Lambda_n|, \nu)|T_n - T_n^{(j)}|^2 + \mathbb{E}_j|T_n - T_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, \nu)).
\end{aligned}$$

By representation (24),

$$|T_n - T_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, \nu) \leq (|b(z)|^2 + a_n^2(z))|\Lambda_n - \Lambda_n^{(j)}|^2 h_\gamma^2(|\Lambda_n|, \nu).$$

Given that $|T_n| \leq C|\Lambda_n||b_n(z)|$, arrive at the inequality

$$A_4 \leq C(|b(z)|^2 + a_n^2(z)) \mathbb{E}_j|\Lambda_n - \Lambda_n^{(j)}|^2 \mathbb{I}\{\min\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq Ca_n(z)\}.$$

Applying Lemma 10, obtain

$$A_4 \leq C(|b(z)|^2 + a_n^2(z)) \frac{Ca_n(z)}{nv} \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{C}{(nv)^2} \right).$$

Combining the estimates obtained for $A_1, \dots, A_4$ we conclude that

$$V_j^{(2)} \leq C(|b(z)|^2 + a_n^2(z)) \frac{Ca_n(z)}{nv} \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{C}{(nv)^2} \right). \quad (27)$$

Inequalities (22), (23) and (27) imply the bound

$$\begin{aligned}
V_j^{\frac{q}{2}} &\leq C^q \beta_n^{\frac{q}{4}}(z) \\
&\quad \times \left(|b(z)|^2 + a_n^2(z) \right)^{\frac{q}{4}} \left(\frac{a_n(z)}{nv} \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{C}{(nv)^2} \right) \right)^{\frac{q}{4}}.
\end{aligned} \quad (28)$$

Note that

$$\widehat{D}_n^{(24)} \leq Cq \left(\frac{1}{n} \sum_{j=1}^n V_j \right) D_n^{\frac{q-2}{q}}.$$

Then inequality (28) yields

$$\begin{aligned}
\widehat{D}_n^{(24)} &\leq Cq \beta_n^{\frac{1}{2}}(z) \left[\left(|b(z)|^2 + a_n^2(z) \right)^{\frac{1}{2}} \left(\frac{a_n(z)}{nv} \left(\frac{a_n^2(z)}{(nv)^2} \right. \right. \right. \\
&\quad \left. \left. \left. + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{C}{(nv)^2} \right) \right)^{\frac{1}{2}} \right] D_n^{\frac{q-2}{q}}.
\end{aligned}$$

Using that $|b(z)| \leq Ca_n(z)$ (see Lemma 4) we may rewrite it as

$$\widehat{D}_n^{(24)} \leq Cq\beta_n^2 D_n^{\frac{q-2}{q}}. \quad (29)$$

5.2.4 Estimation of $\widetilde{D}_n^{(24)}$

Recall that

$$\widetilde{D}_n^{(24)} = \frac{C^q q^{q-1}}{n} \sum_{j=1}^n \mathbb{E} |\widehat{\varepsilon}_j| |\widehat{T}_n - \widehat{T}_n^{(j)}|^{q-1} h_\gamma(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\}.$$

Using (25) and (24), we get

$$|\widehat{T}_n - \widehat{T}_n^{(j)}| \leq \frac{C}{nv} (a_n(z) + |b(z)|) + \frac{C}{n^2 v^2} + |T_n^{(j)}| |h(|\Lambda_n|, v) - h(|\Lambda_n^{(j)}|, v)|.$$

Further, (26) gives

$$\begin{aligned} |\widehat{T}_n - \widehat{T}_n^{(j)}| &\leq \frac{C}{nv} (a_n(z) + |b(z)|) + \frac{C}{n^2 v^2} \\ &\quad + \frac{C}{a_n(z)} |T_n^{(j)}| |\Lambda_n - \Lambda_n^{(j)}| \mathbb{I}\{|\Lambda_n| \leq Ca_n(z)\}. \end{aligned}$$

Applying the triangle inequality, obtain

$$\begin{aligned} |\widehat{T}_n - \widehat{T}_n^{(j)}| &\leq \frac{C}{nv} (a_n(z) + |b(z)|) + \frac{C}{n^2 v^2} \\ &\quad + \frac{C}{a_n(z)} |T_n| |\Lambda_n - \Lambda_n^{(j)}| \mathbb{I}\{\max\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq Ca_n(z)\} \\ &\quad + \frac{C}{a_n(z)} |T_n - T_n^{(j)}| |\Lambda_n - \Lambda_n^{(j)}| \mathbb{I}\{\min\{|\Lambda_n|, |\Lambda_n^{(j)}|\} \leq Ca_n(z)\}. \end{aligned}$$

Inequality (24), the relations $|T_n| \leq C|\Lambda_n||b(z)|$ and $\frac{1}{nv} \leq Ca_n(z)$ imply

$$|\widehat{T}_n - \widehat{T}_n^{(j)}| \leq \frac{C}{nv} (a_n(z) + |b(z)|).$$

The last gives

$$\widetilde{D}_n^{(24)} \leq \frac{C^q q^{q-1} a_n^{q-2}(z)}{(nv)^{q-2}} \frac{1}{n} \sum_{j=1}^n \mathbb{E} |\widehat{\varepsilon}_j| |\widehat{T}_n - \widehat{T}_n^{(j)}| h_\gamma(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\}.$$

Taking into account

$$\mathbb{E} |\widehat{\varepsilon}_j| |\widehat{T}_n - \widehat{T}_n^{(j)}| h_\gamma(|\Lambda_n|, v) \mathbb{I}\{\mathcal{B}\} \leq \mathbb{E}(V_j^{(1)} V_j^{(2)})^{\frac{1}{2}}$$

and inequalities (23), (27), we get

$$\begin{aligned} \widetilde{D}_n^{(24)} &\leq \frac{C^q q^{q-1} a_n^{q-1}(z)}{(nv)^{q-2}} \beta_n^{\frac{1}{2}} \left(\frac{a_n(z)}{nv} \right. \\ &\quad \times \left. \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{C}{(nv)^2} \right) \right)^{\frac{1}{2}}. \end{aligned}$$

We rewrite it as

$$\widetilde{D}_n^{(24)} \leq C^q q^{q-1} \beta_n^q \sqrt{nv}. \quad (30)$$

Combining relations (15), (16), (17), (18), (21), (29), (30), we find that

$$D_n \leq C n^{-c \log n} + C \beta_n D_n^{\frac{q-1}{q}} + C \beta_n^{\frac{1}{2}} D_n^{\frac{2q-1}{2q}} + C \beta_n^2 q D_n^{\frac{q-2}{q}} + C^q q^{q-1} \beta_n^q \sqrt{nv}.$$

Applying Young's inequality, we conclude that

$$D_n \leq C^q q^{\frac{q}{2}} (1 + \sqrt{nv} q^{q/2}) \beta_n^q.$$

6 Appendix

Lemma 3 For any $z \in \mathcal{D}_\varepsilon$ we have

$$|b(z)| \leq \sqrt{2} \operatorname{Im} b(z).$$

Proof Note that

$$b(z) = \frac{1}{2y} \sqrt{\left(z - \frac{a^2}{z}\right) \left(z - \frac{b^2}{z}\right)}.$$

It is enough to prove that

$$\operatorname{Re} \left\{ \left(z - \frac{a^2}{z}\right) \left(z - \frac{b^2}{z}\right) \right\} \leq 0. \quad (31)$$

We have

$$\operatorname{Re} \left\{ \left(z - \frac{a^2}{z}\right) \left(z - \frac{b^2}{z}\right) \right\} = (u^2 - v^2) \left(1 + \frac{(1-y)^2}{|z|^4}\right) - 2(1+y).$$

It may be rewritten as

$$\operatorname{Re} \left\{ \left(z - \frac{a^2}{z} \right) \left(z - \frac{b^2}{z} \right) \right\} = |z|^2 \left(1 + \frac{(1-y)^2}{|z|^4} \right) - 2(1+y) - 2v^2 \left(1 + \frac{(1-y)^2}{|z|^4} \right).$$

It is straightforward to check that for $1 - \sqrt{y} \leq |z| \leq 1 + \sqrt{y}$

$$|z|^2 \left(1 + \frac{(1-y)^2}{|z|^4} \right) - 2(1+y) \leq 0.$$

This implies that for $1 - \sqrt{y} \leq |z| \leq 1 + \sqrt{y}$ inequality (31) holds. Since $1 - \sqrt{y} \leq |z|$ for all $z \in \mathcal{D}_\varepsilon$, to prove (31) for $z \in \mathcal{D}_\varepsilon$ is enough to consider $|z| \geq 1 + \sqrt{y}$. For this case we have

$$\operatorname{Re} \left\{ \left(z - \frac{a^2}{z} \right) \left(z - \frac{b^2}{z} \right) \right\} \leq u^2 - v^2 + (1 - \sqrt{y})^2 - 2(1+y) \leq u^2 - (1 + \sqrt{y})^2 \leq 0.$$

Thus Lemma is proved. $\square$

Lemma 4 For $1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y}$ the inequality

$$|b(z)| \leq C a_n(z)$$

holds.

Proof Note that

$$b(z) = z - \frac{1-y}{z} + 2yS_y(z) = \sqrt{\left(z - \frac{1-y}{z} \right)^2 - 4y}$$

and

$$a_n(z) = \operatorname{Im} \left\{ \sqrt{\left(z - \frac{1-y}{z} \right)^2 - 4y} \right\} + \frac{1}{nv} + \frac{1}{np}.$$

It is easy to show that for $1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y}$

$$\operatorname{Re} \left\{ \left(z - \frac{1-y}{z} \right)^2 - 4y \right\} \leq 0.$$

Indeed,

$$\operatorname{Re} \left\{ \left(z - \frac{1-y}{z} \right)^2 - 4y \right\} \leq u^2 + \frac{(1-y)^2}{u^2} - 2(1+y).$$

The last expression is not positive for $1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y}$. From the negativity of the real part it follows that

$$\operatorname{Im} \left\{ \sqrt{\left(z - \frac{1-y}{z}\right)^2 - 4y} \right\} \geq \frac{1}{\sqrt{2}} \left| \sqrt{\left(z - \frac{1-y}{z}\right)^2 - 4y} \right|$$

This implies the required. The Lemma is proved. $\square$

Lemma 5 *Under the conditions of Theorem 3, for $j \in \mathbb{J}^c$ and $l \in \mathbb{K}^c$ we have*

$$\max \{ |\varepsilon_{j1}^{(\mathbb{J}, \mathbb{K})}|, |\varepsilon_{l+n,1}^{(\mathbb{J}, \mathbb{K})}| \} \leq \frac{C}{nv}.$$

Proof For simplicity, we consider only the case $\mathbb{J} = \emptyset$ and $\mathbb{K} = \emptyset$. Note that

$$\begin{aligned} \varepsilon_{j1} &= \frac{1}{2m} \left(\left(\operatorname{Tr} \mathbf{R} - \frac{m-n}{z} \right) - \left(\operatorname{Tr} \mathbf{R}^{(j)} - \frac{m-n-1}{|z|} \right) \right) \\ &= \frac{1}{2m} \left(\operatorname{Tr} \mathbf{R} - \operatorname{Tr} \mathbf{R}^{(j)} \right) - \frac{1}{2mz}. \end{aligned}$$

Applying Schur's formula, we get

$$|\varepsilon_{j1}| \leq \frac{1}{nv}.$$

The second inequality is proved similarly. $\square$

Lemma 6 *For any $z \in \mathcal{D}_\varepsilon$*

$$\lim_{n \rightarrow \infty} (nvb(z)) = \infty.$$

Proof First we note that

$$b(z) = \frac{1}{z} \sqrt{(z-a)(z+a)(z-b)(z+b)}.$$

It is straightforward to check that, for $z \in \mathcal{D}_\varepsilon$

$$|b(z)| \geq C \sqrt{\min\{|1 - \sqrt{y} - z|, |1 - \sqrt{y} + z|, |1 + \sqrt{y} - z|, |1 + \sqrt{y} + z|\}} \geq C\sqrt{\beta + v}.$$

Further,

$$nv|b(z)| \geq C(nv_0/\sqrt{\beta})\sqrt{\beta + v} \geq nv_0 \rightarrow \infty.$$

Thus Lemma is proved. $\square$

Lemma 7 *Under the conditions of Theorem 3, for all $j \in \mathbb{J}^c$ the inequalities*

$$\mathbb{E}_j |\varepsilon_{j2}^{(\mathbb{J}, \mathbb{K})}|^2 \leq \frac{\mu_4}{np} \frac{1}{n} \sum_{l=1}^m |R_{l+n, l+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})}|^2$$

and

$$\mathbb{E}_{l+n} |\varepsilon_{l+n,2}^{(\mathbb{J},\mathbb{K})}|^2 \leq \frac{\mu_4}{np} \frac{1}{n} \sum_{j=1}^n |R_{jj}^{(\mathbb{J},\mathbb{K} \cup \{l\})}|^q.$$

are valid. In addition for $q > 2$ we have

$$\mathbb{E}_j |\varepsilon_{j2}^{(\mathbb{J},\mathbb{K})}|^q \leq C^q \left(\frac{q^{\frac{q}{2}}}{(np)^{\frac{q}{2}}} + \frac{q^q}{(np)^{2q\kappa+1}} \right) \frac{1}{n} \sum_{l=1}^m |R_{l+n,l+n}^{(\mathbb{J} \cup \{j\},\mathbb{K})}|^q$$

and

$$\mathbb{E}_{l+n} |\varepsilon_{l+n,2}^{(\mathbb{J},\mathbb{K})}|^q \leq C^q \left(\frac{q^{\frac{q}{2}}}{(np)^{\frac{q}{2}}} + \frac{q^q}{(np)^{2q\kappa+1}} \right) \frac{1}{n} \sum_{j=1}^n |R_{jj}^{(\mathbb{J},\mathbb{K} \cup \{l\})}|^q$$

for $l \in \mathbb{K}^c$.

Proof For simplicity, consider the case $\mathbb{J} = \emptyset$ and $\mathbb{K} = \emptyset$. The first two inequalities are obvious. We consider only $q > 2$. Applying Rosenthal's inequality (see [11], Lemma 8.1), obtain for any $q > 2$

$$\begin{aligned} \mathbb{E}_j |\varepsilon_{j2}|^q &= \frac{1}{(mp)^q} \mathbb{E}_j \left| \sum_{l=1}^m (X_{jl}^2 \xi_{jl} - p) R_{l+n,l+n}^{(j)} \right|^q \\ &\leq \frac{C^q}{(mp)^q} \left[q^{\frac{q}{2}} \left(\sum_{l=1}^m \mathbb{E}_j |X_{jl}^2 \xi_{jl} - p|^2 |R_{l+n,l+n}^{(j)}|^2 \right)^{\frac{q}{2}} \right. \\ &\quad \left. + q^q \sum_{l=1}^m \mathbb{E}_j |X_{jl}^2 \xi_{jl} - p|^q |R_{l+n,l+n}^{(j)}|^q \right] \\ &\leq \frac{C^q}{(mp)^{\frac{q}{2}}} \left[(q\mu_4)^{\frac{q}{2}} \left(\frac{1}{m} \sum_{l=1}^m |R_{l+n,l+n}^{(j)}|^2 \right)^{\frac{q}{2}} + \frac{q^q}{(mp)^{\frac{q}{2}}} \tilde{\mu}_{2q} \sum_{l=1}^m |R_{l+n,l+n}^{(j)}|^q \right] \\ &\leq C^q \left[\frac{(q\mu_4)^{\frac{q}{2}}}{(mp)^{\frac{q}{2}}} + \frac{mq^q}{(mp)^q} \tilde{\mu}_{2q} \right] \frac{1}{m} \sum_{l=1}^m |R_{l+n,l+n}^{(j)}|^q. \end{aligned} \quad (32)$$

Recall that

$$\tilde{\mu}_r = \mathbb{E} |X_{jk} \xi_{jk}|^r$$

and by the conditions of the Theorem

$$\tilde{\mu}_{2q} \leq C^q p (np)^{q-2q\kappa-2} \mu_{4+\delta}.$$

Substituting the last inequality in (32), we get

$$\mathbb{E}_j |\varepsilon_{j2}|^q \leq C^q \left[\frac{q^{\frac{q}{2}}}{(mp)^{\frac{q}{2}}} + \frac{q^q}{(mp)^{2q\kappa+1}} \right] \frac{1}{m} \sum_{l=1}^m |R_{l+n,l+n}^{(j)}|^q.$$

The second inequality can be proved similarly. $\square$

Lemma 8 *Under the conditions of Theorem 3 for all $j \in \mathbb{T}_{\mathbb{J}}$, $l \in \mathbb{T}_{\mathbb{K}}^1$*

$$\mathbb{E}_j |\varepsilon_{j3}^{(\mathbb{J}, \mathbb{K})}|^2 \leq \frac{C \operatorname{Im} s_n^{(\mathbb{J} \cup \{j\}, \mathbb{K})}}{nv}, \quad \mathbb{E}_{l+n} |\varepsilon_{l+n,3}^{(\mathbb{J}, \mathbb{K})}|^2 \leq \frac{C \operatorname{Im} s_n^{(\mathbb{J}, \mathbb{K} \cup \{l\})}}{nv}$$

are valid. In addition for $q > 2$ we have

$$\begin{aligned} \mathbb{E}_j |\varepsilon_{j3}^{(\mathbb{J}, \mathbb{K})}|^q &\leq C^q \left(q^q (nv)^{-\frac{q}{2}} \left(\frac{C \operatorname{Im} s_n^{(\mathbb{J} \cup \{j\}, \mathbb{K})}}{nv} \right)^{\frac{q}{2}} \right. \\ &\quad \left. + q^{\frac{3q}{2}} (nv)^{-\frac{q}{2}} (np)^{-q\kappa-1} \frac{1}{n} \sum_{l=1}^m \left(\operatorname{Im} R_{l+n,l+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})} \right)^{\frac{q}{2}} \right. \\ &\quad \left. + q^{2q} (np)^{-2q\kappa} \frac{1}{n^2} \sum_{l=1}^m \sum_{k=1}^m \left| R_{l+n,k+n}^{(\mathbb{J} \cup \{j\}, \mathbb{K})} \right|^q \right), \end{aligned}$$

$$\begin{aligned} \mathbb{E}_{l+n} |\varepsilon_{l+n,3}^{(\mathbb{J}, \mathbb{K})}|^q &\leq C^q \left(q^q (nv)^{-\frac{q}{2}} \left(\frac{C \operatorname{Im} s_n^{(\mathbb{J} \cup \{j\}, \mathbb{K})}}{nv} \right)^{\frac{q}{2}} \right. \\ &\quad \left. + q^{\frac{3q}{2}} (nv)^{-\frac{q}{2}} (np)^{-q\kappa-1} \frac{1}{n} \sum_{j=1}^m \left(\operatorname{Im} R_{jj}^{(\mathbb{J}, \mathbb{K} \cup \{l+n\})} \right)^q \right. \\ &\quad \left. + q^{2q} (np)^{-2q\kappa} n^{-2} \sum_{j=1}^n \sum_{k=1}^n \left| R_{kj}^{(\mathbb{J}, \mathbb{K} \cup \{l+n\})} \right|^q \right). \end{aligned}$$

Proof It suffices to apply the inequality from Corollary 1 of [17]. $\square$

Lemma 9 *Under the conditions of the theorem, the bound*

$$\mathbb{E}_j |R_{jj} - \mathbb{E}_j R_{jj}|^2 \mathbb{I}\{\mathcal{Q}\} \mathbb{I}\{\mathcal{B}\} \leq C \left(\frac{a_n(z)}{nv} + \frac{A_0^2(z)}{np} \right) + C n^{-c \log n}$$

is valid.

Proof Consider the equality

$$R_{jj} = - \frac{1}{z - \frac{1-y}{z} + y s_n^{(j)}(z)} \left(1 + \widehat{\varepsilon}_j R_{jj} \right).$$

It implies

$$R_{jj} - \mathbb{E}_j R_{jj} = -\frac{1}{z - \frac{1-y}{z} + y s_n^{(j)}(z)} \left(\widehat{\varepsilon}_j R_{jj} - \mathbb{E}_j \widehat{\varepsilon}_j R_{jj} \right).$$

It is well-known that

$$y S_y(z) + z - \frac{1-y}{z} = \frac{1}{S_y(z)}.$$

Consequently,

$$\left| y S_y(z) + z - \frac{1-y}{z} \right| \geq \sqrt{y}.$$

Further, note that for a sufficiently small γ there is a constant H such that

$$\left| \frac{1}{z - \frac{1-y}{z} + y s_n^{(j)}(z)} \right| \mathbb{I}\{Q\} \leq H \mathbb{I}\{Q\}.$$

Hence,

$$\begin{aligned} \mathbb{E}_j |R_{jj} - \mathbb{E}_j R_{jj}|^2 \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} &\leq H^2 \left(\mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} \right. \\ &\quad \left. + \mathbb{E}_j \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} \mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \right). \end{aligned}$$

It is easy to see that

$$\mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} \leq C \mathbb{E}_j |\widehat{\varepsilon}_j|^2 \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} \leq C \left(\frac{1}{nv} a_n(z) + \frac{C}{np} \right).$$

Introduce events

$$Q^{(j)} = \left\{ |\Lambda_n^{(j)}| \leq 2\gamma a_n(z) + \frac{1}{nv} \right\}.$$

It is obvious that

$$\mathbb{I}\{Q\} \leq \mathbb{I}\{Q\} \mathbb{I}\{Q^{(j)}\}.$$

Consequently,

$$\mathbb{E}_j \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} \mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \leq \mathbb{E}_j \mathbb{I}\{Q\} \mathbb{I}\{\mathcal{B}\} \mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{Q^{(j)}\}.$$

Further, consider $\widetilde{Q} = \{|\Lambda_n| \leq 2\gamma a_n(z)\}$. We have

$$\mathbb{I}\{Q^{(j)}\} \leq \mathbb{I}\{\widetilde{Q}\}.$$

Then it follows that

$$\mathbb{E}_j \mathbb{I}\{\mathbf{Q}\} \mathbb{I}\{\mathcal{B}\} \mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \leq \mathbb{E}_j \mathbb{I}\{\mathbf{Q}\} \mathbb{I}\{\mathcal{B}\} \mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{\tilde{\mathbf{Q}}\}.$$

Next, the inequality

$$\mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{\tilde{\mathbf{Q}}\} \leq \mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{\tilde{\mathbf{Q}}\} \mathbb{I}\{\tilde{\mathcal{B}}\} + \mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{\tilde{\mathbf{Q}}\} \mathbb{I}\{\tilde{\mathcal{B}}^c\} \quad (33)$$

holds. By the conditions C0 and the inequality $|R_{jj}| \leq v_0^{-1}$, we obtain the bound

$$\mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{\tilde{\mathbf{Q}}\} \mathbb{I}\{\tilde{\mathcal{B}}^c\} \leq C n^{-c \log n}.$$

For the first term on the right side of (33) get

$$\mathbb{E}_j |\widehat{\varepsilon}_j|^2 |R_{jj}|^2 \mathbb{I}\{\tilde{\mathbf{Q}}\} \mathbb{I}\{\tilde{\mathcal{B}}\} \leq C \left(\frac{a_n(z)}{nv} + \frac{C}{np} \right).$$

This completes the proof. $\square$

Lemma 10 *Under the conditions of the theorem, we have*

$$\mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^2 \mathbb{I}\{\mathbf{Q}\} \mathbb{I}\{\mathcal{B}\} \leq \frac{C a_n(z)}{nv} \left(\frac{a_n^2(z)}{(nv)^2} + \frac{a_n(z)}{(np)(nv)} + \frac{1}{np(nv)} + \frac{C}{(nv)^2} \right) + n^{-c \log n}.$$

Proof Set $\widehat{\Lambda}_n^{(j)} = s_n^{(j)}(z) - S_y(z)$. By the Schur complement formula,

$$\Lambda_n - \widehat{\Lambda}_n^{(j)} = \frac{1}{2n} \left(1 + \frac{1}{np} \sum_{l,k=1}^m X_{jl} X_{jk} \xi_{jl} \xi_{jk} [R^{(j)}]_{k+n,l+n}^2 \right) R_{jj}.$$

Since $\widehat{\Lambda}_n^{(j)}$ is measurable with respect to $\mathfrak{M}^{(j)}$, we may write

$$\Lambda_n - \Lambda_n^{(j)} = (\Lambda_n - \widehat{\Lambda}_n^{(j)}) - \mathbb{E}_j \{\Lambda_n - \widehat{\Lambda}_n^{(j)}\}.$$

Introduce the notation

$$\eta_{j1} = \frac{1}{np} \sum_{l=1}^m (X_{jl}^2 \xi_{jl} - p) [\mathbf{R}^{(j)^2}]_{l+n,l+n}, \quad \eta_{j2} = \frac{1}{np} \sum_{l=1}^m X_{jl} X_{jk} \xi_{jl} \xi_{jk} [\mathbf{R}^{(j)^2}]_{k+n,l+n}.$$

In these notation

$$\begin{aligned}\Lambda_n - \Lambda_n^{(j)} &= \frac{1}{n} \left(1 + \frac{1}{n} \sum_{l=1}^m [\mathbf{R}^{(j)^2}]_{l+n, l+n} \right) (R_{jj} - \mathbb{E}_j R_{jj}) \\ &\quad + \frac{1}{n} (\eta_{j1} + \eta_{j2}) R_{jj} - \frac{1}{n} \mathbb{E}_j (\eta_{j1} + \eta_{j2}) R_{jj}.\end{aligned}$$

Note that

$$\mathbb{E}_j |\eta_{j1}|^2 \mathbb{I}\{\mathcal{Q}\} \mathbb{I}\{\mathcal{B}\} \leq \frac{Cp}{n^2 p^2} \sum_{l=1}^m |[\mathbf{R}^{(j)^2}]_{l+n, l+n}|^2 \mathbb{I}\{\mathcal{Q}^{(j)}\} \mathbb{I}\{\mathcal{B}^{(j)}\}.$$

Since

$$|[\mathbf{R}^{(j)^2}]_{l+n, l+n}| \leq \sum_{k=1}^m |R^{(j)}_{l+n, k+n}|^2 \leq \frac{C}{v} \text{Im } R^{(j)}_{l+n, l+n}.$$

Theorem 3 gives

$$\mathbb{E}_j |\eta_{j1}|^2 \mathbb{I}\{\mathcal{Q}\} \mathbb{I}\{\mathcal{B}\} \leq \frac{C}{npv^2} \frac{1}{n} \sum_{l=1}^m (\text{Im } R^{(j)}_{l+n, l+n})^2 \mathbb{I}\{\mathcal{Q}^{(j)}\} \mathbb{I}\{\mathcal{B}^{(j)}\} \leq \frac{Ca_n(z)}{npv^2}.$$

Similarly, for the second moment of η_{j2} we have the estimate

$$\begin{aligned}\mathbb{E}_j |\eta_{j2}|^2 \mathbb{I}\{\mathcal{Q}\} \mathbb{I}\{\mathcal{B}\} &\leq \frac{Cp^2}{n^2 p^2} \sum_{l,k=1}^m |[\mathbf{R}^{(j)^2}]^2_{l+n, k+n}|^2 \mathbb{I}\{\mathcal{Q}^{(j)}\} \mathbb{I}\{\mathcal{B}^{(j)}\} \\ &\leq \frac{C}{n^2} \text{Tr } |\mathbf{R}|^4 \mathbb{I}\{\mathcal{Q}^{(j)}\} \mathbb{I}\{\mathcal{B}^{(j)}\} \leq \frac{C}{nv^3} a_n(z).\end{aligned}$$

From the above estimates and Lemma 9 we conclude that

$$\begin{aligned}\mathbb{E}_j |\Lambda_n - \Lambda_n^{(j)}|^2 \mathbb{I}\{\mathcal{Q}\} \mathbb{I}\{\mathcal{B}\} &\leq \frac{Ca_n^2(z)(z)}{n^2 v^2} \mathbb{E}_j |R_{jj} - \mathbb{E}_j R_{jj}|^2 \mathbb{I}\{\mathcal{Q}\} \mathbb{I}\{\mathcal{B}\} \\ &\quad + \left(\frac{a_n(z)}{np(nv)^2} + \frac{Ca_n(z)}{(nv)^3} \right) \leq \frac{Ca_n(z)}{(nv)^2} \left(\frac{C}{np} + \frac{C}{(nv)} \right) + n^{-c \log n}.\end{aligned}$$

The lemma is proved. $\square$

Lemma 11 *There is an absolute constant $C > 0$ such that for any $z = u + iv$ satisfying*

$1 - \sqrt{y} - v \leq |u| \leq 1 + \sqrt{y} + v$ and $v > 0$ the inequality

$$|\Lambda_n| \leq C \min \left\{ \frac{|T_n|}{|b(z)|}, \sqrt{|T_n|} \right\} \quad (34)$$

is valid.

Proof Make the change of variables by setting

$$w = \frac{1}{\sqrt{y}} \left(z - \frac{1-y}{z} \right), \quad z = \frac{w\sqrt{y} + \sqrt{yw^2 + 4(1-y)}}{2},$$

and

$$\tilde{S}(w) = \sqrt{y} S_y(z), \quad \tilde{s}_n(w) = \sqrt{y} s_n(z).$$

In this notation, we can rewrite the main equation in the form

$$1 + w\tilde{s}_n(w) + \tilde{s}_n^2(w) = T_n.$$

It is easy to see that

$$\Lambda_n = \frac{1}{\sqrt{y}} (\tilde{s}_n(z) - \tilde{S}(w)).$$

Now it suffices to repeat the proof of Lemma B.1 from [18]. Note that this Lemma implies (34) for w satisfying $|\operatorname{Re} w| \leq 2 + \operatorname{Im} w$. From this we conclude that inequality (34) holds for $z = u + iv$ such that $1 - \sqrt{y} \leq |u| \leq 1 + \sqrt{y}$. The Lemma is proved. $\square$

Lemma 12 *Under conditions of Theorem 3 we have*

$$\Pr \left\{ |\Lambda_n(u + iV)| \leq C |S_y(u + iV)|^2 \left(\frac{1}{n} + \frac{1}{np} \right), \text{ for all } -\infty < u < \infty \right\} \geq 1 - Cn^{-Q}$$

Proof First we note that $|R_{jk}| \leq 1/V$, for $v = V$ for $j, l = 1, \dots, n$,

This implies that

$$|b_n(u + iV)| \geq \frac{1}{|S_y(u + iV)|} - \frac{1}{V}.$$

We may choose V so large that

$$|b_n(u + iV)| \geq \frac{2}{|S_y(z)|}.$$

From here it follows

$$\begin{aligned} |\Lambda_n| &\leq \frac{C|S_y(z)|}{n} \left| \sum_{j=1}^n \varepsilon_j R_{jj} \right| \leq \frac{C|S_y(z)|^2}{n} \left| \sum_{j=1}^n \varepsilon_j \right| \\ &\quad + \frac{C|S_y(z)|}{n} \left(\sum_{j=1}^n |\varepsilon_j| |R_{jj} - S_y(z)| \right). \end{aligned}$$

Using Eq. (9), we get

$$|\Lambda_n| \leq C|S_y(z)|^2 \left| \frac{1}{n} \sum_{j=1}^n \varepsilon_j \right| + C|S_y(z)|^2 \left(\frac{1}{n} \sum_{j=1}^n |\varepsilon_j|^2 \right).$$

From this inequality it follows that

$$\mathbb{E} |\Lambda_n|^q \leq C^q |S_y(z)|^{2q} \left(\mathbb{E} \left| \frac{1}{n} \sum_{j=1}^n \varepsilon_j \right|^q + \frac{1}{n} \sum_{j=1}^n \mathbb{E} |\varepsilon_j|^{2q} \right). \quad (35)$$

Applying Lemmas 5, 7, 8, we get

$$\frac{1}{n} \sum_{j=1}^n \mathbb{E} |\varepsilon_j|^{2q} \leq \frac{C^q}{(np)^q}.$$

To estimate the first term in the right hand side of (35), we use triangle inequality

$$\mathbb{E} \left| \frac{1}{n} \sum_{j=1}^n \varepsilon_j \right|^q \leq C^q \left(\mathbb{E} \left| \frac{1}{n} \sum_{j=1}^n \varepsilon_{j1} \right|^q + \mathbb{E} \left| \frac{1}{n} \sum_{j=1}^n \varepsilon_{j2} \right|^q + \mathbb{E} \left| \frac{1}{n} \sum_{j=1}^n \varepsilon_{j3} \right|^q \right) =: S_1 + S_2 + S_3.$$

By Lemma 5

$$\mathbb{E} \left| \frac{1}{n} \sum_{j=1}^n \varepsilon_{j1} \right|^q \leq \frac{C^q}{n^q}.$$

To estimate the second and the third terms we use the inequality of [17, Theorem1].

We get

$$S_2 \leq \frac{C^q q^q}{(np)^q}, \quad S_3 \leq \frac{C^q q^q}{n^q}.$$

It is enough to repeat the proof of [17, Theorem 3]. We omit here the details. Thus Lemma 12 is proved. $\square$

References

1. Wishart, J.: The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* **20A**(1/2), 32–52 (1928)
2. Wigner, E.P.: Characteristic vectors of bordered matrices with infinite dimensions. *Ann. Math.* **62**(3), 548–564 (1955)
3. Marchenko, V.A., Pastur, L.A.: Distribution of eigenvalues for some sets of random matrices. *Mat. Sb. (N.S.)* **72**(4), 507–536 (1967)
4. Telatar, E.: Capacity of multi-antenna Gaussian channels. *Eur. Trans. Telecomm.* **10**(6), 585–595 (1999)
5. Erdős, L., Knowles, A., Yau, H.-T., Yin, J.: Spectral statistics of Erdős-Rényi graphs I: Local semicircle law. *Ann. Probab.* **41**(3B), 2279–2375 (2013)
6. Erdős, L., Knowles, A., Yau, H.-T., Yin, J.: Spectral statistics of Erdős-Rényi graphs II: eigenvalue spacing and the extreme eigenvalues. *Commun. Math. Phys.* **314**(3), 587–640 (2012)
7. Lee, J.O., Schnelli, K.: Tracy-Widom distribution for the largest eigenvalue of real sample covariance matrices with general population. *Ann. Appl. Probab.* **26**(6), 3786–3839 (2016)
8. Hwang, J.Y., Lee, J.O., Schnelli, K.: Local law and Tracy-Widom limit for sparse sample covariance matrices. *Ann. Appl. Probab.* **29**(5), 3006–3036 (2019)
9. Hwang, J.Y., Lee, J.O., Yang, W.: Local law and Tracy-Widom limit for sparse stochastic block models. *Bernoulli* **26**(3), 2400–2435 (2020)
10. Lee, J.O., Schnelli, K.: Local law and Tracy-Widom limit for sparse random matrices. *Probab. Theory Relat. Fields* **171**, 543–616 (2018)
11. Götze, F., Tikhomirov, A.N.: Rate of Convergence of the Expected Spectral Distribution Function to the Marchenko–Pastur Law (2014). <https://arxiv.org/abs/1412.6284>
12. Bay, Z.D., Silverstein, J.W.: *Spectral Analysis of Large Dimensional Random Matrices*, 2nd edn. Springer Series in Statistics, New York, NY, USA (2010)
13. Götze, F., Tikhomirov, A.N.: Optimal bounds for convergence of expected spectral distributions to the semi-circular law. *Probab. Theory Relat. Fields* **165**, 163–233 (2016)
14. Götze, F., Naumov, A.A., Tikhomirov, A.N.: On the local semicircular law for Wigner ensembles. *Bernoulli* **24**(3), 2358–2400 (2018)
15. Götze, F., Naumov, A.A., Tikhomirov, A.N.: Local semicircle law under moment conditions: the stieltjes transform, rigidity, and delocalization. *Theory Probab. Appl.* **62**(1), 58–83 (2018)
16. Götze, F., Naumov, A.A., Tikhomirov, A.N.: Local semicircle law under fourth moment condition. *J. Theor. Probab.* **33**, 1327–1362 (2020)
17. Götze, F., Naumov, A.A., Tikhomirov, A.N.: Moment inequalities for linear and nonlinear statistics. *Theory Probab. Appl.* **65**(1), 1–16 (2020)
18. Götze, F., Naumov, A.A., Tikhomirov, A.N.: Local Semicircle Law under Moment Conditions. Part I: The Stieltjes Transform (2016). <https://arxiv.org/abs/1510.07350>