The local circular law III: general case

Jun Yin[*]

Department of Mathematics, University of Wisconsin-Madison
Madison, WI 53706-1388, USA jyin@math.wisc.edu

Abstract

In the first part [9] of this article series, Bourgade, Yau and the author of this paper proved a local version of the circular law up to the finest scale $N^{-1/2+\varepsilon}$ for non-Hermitian random matrices at any point $z \in \mathbb{C}$ with $||z| - 1| > c$ for any constant $c > 0$ independent of the size of the matrix. In the second part [10], they extended this result to include the edge case $|z| - 1 = o(1)$, under the main assumption that the third moments of the matrix elements vanish. (Without the vanishing third moment assumption, they proved that the circular law is valid near the spectral edge $|z| - 1 = o(1)$ up to scale $N^{-1/4+\varepsilon}$.) In this paper, we will remove this assumption, i.e. we prove a local version of the circular law up to the finest scale $N^{-1/2+\varepsilon}$ for non-Hermitian random matrices at any point $z \in \mathbb{C}$.

AMS Subject Classification (2010): 15B52, 82B44

Keywords: local circular law, universality.

1 Introduction and Main result

The circular law in random matrix theory describes the macroscopic limiting spectral measure of normalized non-Hermitian matrices with independent entries. Its origin goes beck to the work of Ginibre [18], who found the joint density of the eigenvalues of such Gaussian matrices. More precisely, for an $N \times N$ matrix with independent entries $\frac{1}{\sqrt{N}} z_{ij}$ such that z_{ij} is identically distributed according to the measure $\mu_g = \frac{1}{\pi} e^{-|z|^2} \mathrm{dA}(z)$ (dA denotes the Lebesgue measure on $\mathbb{C}$), its eigenvalues $\mu_1, \ldots, \mu_N$ have a probability density proportional to

$$\prod_{i<j} |\mu_i - \mu_j|^2 e^{-N \sum_k |\mu_k|^2} \tag{1.1}$$

with respect to the Lebesgue measure on $\mathbb{C}^N$. These random spectral measures define a determinantal point process with the explicit kernel (see [18])

$$K_N(z_1, z_2) = \frac{N}{\pi} e^{-\frac{N}{2}(|z_1|^2 + |z_2|^2)} \sum_{\ell=0}^{N-1} \frac{(N z_1 \overline{z_2})^\ell}{\ell!} \tag{1.2}$$

with respect to the Lebesgue measure on $\mathbb{C}$. This integrability property allowed Ginibre to derive the circular law for the eigenvalues, i.e., $\frac{1}{N} \rho_1^{(N)}$ converges to the uniform measure on the unit circle,

$$\frac{1}{\pi} \mathbb{1}_{|z|<1} \mathrm{dA}(z). \tag{1.3}$$

This limiting law also holds for real Gaussian entries [14], for which a more detailed analysis was performed in [8, 17, 28].

For non-Gaussian entries, Girko [19] argued that the macroscopic limiting spectrum is still given by (1.3). His main insight is commonly known as the *Hermitization technique*, which converts the convergence of complex empirical measures into the convergence of logarithmic transforms of a family of Hermitian matrices. If we denote the original non-Hermitian matrix by X and the eigenvalues of X by μ_j, then for any C^2 function F we have the identity

$$\frac{1}{N} \sum_{j=1}^{N} F(\mu_j) = \frac{1}{4\pi N} \int \Delta F(z) \operatorname{Tr} \log(X^* - z^*)(X - z) \mathrm{dA}(z). \tag{1.4}$$

Due to the logarithmic singularity at 0, it is clear that the small eigenvalues of the Hermitian matrix $(X^* - z^*)(X - z)$ play a special role. A key question is to estimate the small eigenvalues of $(X^* - z^*)(X - z)$, or in other words, the small singular values of $(X - z)$. This problem was not treated in [19], but the gap was remedied in a series of papers. First Bai [3] was able to treat the logarithmic singularity assuming bounded density and bounded high moments for the entries of the matrix (see also [4]). Lower bounds on the smallest singular values were given in Rudelson, Vershynin [26, 27], and subsequently Tao, Vu [30], Pan, Zhou [23] and Götze, Tikhomirov [20] weakened the moments and smoothness assumptions for the circular law, till the optimal L^2 assumption, under which the circular law was proved in [31]. On the other hand, Wood [33] showed that the circular law also holds for sparse random n by n matrices where each entry is nonzero with probability $n^{\alpha-1}$ where $0 < \alpha \leqslant 1$.

In the first part of this article [9], Bourgade, Yau and the author of this paper proved a local version of the circular law, up to the optimal scale $N^{-1/2+\varepsilon}$, in the bulk of the spectrum. In the second part [10], they

extended this result to include the edge case, under the assumption that the third moments of the matrix elements vanish. (Without the vanishing third moment assumption, they also proved that the circular law is valid near the spectral edge $|z| - 1 = \mathrm{o}(1)$ up to scale $N^{-1/4+\varepsilon}$.) This vanishing third moment condition is also the main assumption in Tao and Vu's work on local circular law [32]. In the current paper, we will remove this assumption, i.e. we prove a local version of the circular law up to the finest scale $N^{-1/2+\varepsilon}$ for non-Hermitian random matrices at any point $z \in \mathbb{C}$.

More precisely, we considered an $N \times N$ matrix X with independent real[1] centered entries with variance N^{-1}. Let μ_j, $j \in [\![1, N]\!]$ denote the eigenvalues of X. To state the local circular law, we first define the notion of *stochastic domination*.

Definition 1.1. *Let $W = W^{(N)}$ be a family of random variables and $\Psi = \Psi^{(N)}$ be a family of deterministic parameters. We say that W is* stochastically dominated by Ψ *if for any $\sigma > 0$ and $D > 0$ we have*

$$\mathbb{P}\Big[|W| > N^\sigma \Psi \Big] \;\leqslant\; N^{-D} \tag{1.5}$$

for sufficiently large N. We denote this stochastic domination property by

$$W \;\prec\; \Psi, \quad \text{or} \quad W = \mathrm{O}_\prec(\Psi).$$

Furthermore, Let $U^{(N)}$ be a possibly N-dependent parameter set. We say $W(u)$ is stochastically dominated *by $\Psi(u)$ uniformly in $u \in U^{(N)}$, if for any $\sigma > 0$ and $D > 0$ we have*

$$\sup_{u \in U^{(N)}} \mathbb{P}\Big[|W(u)| > N^\sigma \Psi(u) \Big] \;\leqslant\; N^{-D} \tag{1.6}$$

for uniformly sufficiently large N (may depends on σ and D).

Note: In the most cases of this paper, the $U^{(N)}$ is chosen as the product of the index sets $1 \leqslant i, j \leqslant N$ and some compact set in $\mathbb{C}^2$.

In this paper, as in [9], [10] and [32], we assume that the probability distributions of the matrix elements satisfy the following uniform subexponential decay property:

$$\sup_{(i,j) \in [\![1,N]\!]^2} \mathbb{P}\left(|\sqrt{N} X_{i,j}| > \lambda \right) \leqslant \vartheta^{-1} e^{-\lambda^\vartheta} \tag{1.7}$$

for some constant $\vartheta > 0$ independent of N. This condition can of course be weakened to an hypothesis of boundedness on sufficiently high moments, but the error estimates in the following Theorem would be weakened as well.

Note: most constants appearing in this work may depend on ϑ, but we will not emphasize this dependence in the proof.

Let $f : \mathbb{C} \to \mathbb{R}$ be a fixed smooth compactly supported function, and $f_{z_0}(\mu) = N^{2s} f(N^s(\mu - z_0))$, where z_0 depends on N, and s is a fixed scaling parameter in $[0, 1/2]$. Let D denote the unit disk. Theorem 2.2 of [9] and Theorem 1.2 of [10] assert that the following estimate holds: (Note: Here $\|f_{z_0}\|_1 = O(1)$)

$$\left(N^{-1} \sum_j f_{z_0}(\mu_j) - \frac{1}{\pi} \int_D f_{z_0}(z) \, \mathrm{dA}(z) \right) \;\prec\; N^{-1+2s}, \quad s \in (0, 1/2], \tag{1.8}$$

---

[1]For the sake of notational simplicity we do not consider complex entries in this paper, but the statements and proofs are similar.

[9] if $||z_0| - 1| > c$ for some $c > 0$ independent of N or [10] if the third moments of matrix entries vanish. This implies that the circular law holds after zooming up to scale $N^{-1/2+\varepsilon}$ ($\varepsilon > 0$) under these conditions. In particular, there are neither clusters of eigenvalues nor holes in the spectrum at such scales. We note that in [9] and [10], the scaling parameter was denoted as a, but the letter a will be used as a fixed index in this work.

We aim at understanding the circular law for any $z_0 \in \mathbb{C}$ without the vanishing third moment assumption. The following theorem is our main result.

Theorem 1.2. Local circular law: *Let X be an $N \times N$ matrix with independent centered entries of variances $1/N$. Suppose that the distributions of the matrix elements satisfy the subexponential decay property (1.7). Let f_{z_0} be defined as above (1.8) and D denote the unit disk. Then for any $s \in (0, 1/2]$ and any $z_0 \in \mathbb{C}$, we have*

$$\left(N^{-1} \sum_j f_{z_0}(\mu_j) - \frac{1}{\pi} \int_D f_{z_0}(z) \mathrm{dA}(z) \right) \prec N^{-1+2s}. \tag{1.9}$$

Notice that the main new assertion of (1.9) is for the case: $|z_0| - 1 = o(1)$ and the third moments not vanishing, since the other cases were proved in [9] and [10], stated in (1.8).

Remark: Shortly after the preprint [9] appeared, a version of local circular law (both in the bulk and near the edge) was proved by Tao and Vu [32] under the assumption that the first three moments of the matrix entries match a Gaussian distribution, i.e., the third moment vanish.

In the next section we will introduce our main strategy and improvements.

2 Proof of Theorem 1.2

Proof of Thm. 1.2: The bulk case of Thm. 1.2 was proved in Theorem 2.2 of [9]. Furthermore, it is easy to see that the results in Thm. 1.2 for $s = 1/2$ follow from the results in for $s < 1/2$. Hence in this proof, we can assume that

$$||z_0| - 1| = o(1), \quad s \in (0, 1/2)$$

In the edge case, our Thm. 1.2 was proved in the Thm 1.2 of [10] with the vanishing third moment assumption. Hence the goal of this paper is to improve the proof of Thm. 1.2 of [10]. One can easily check that in the proof of Thm. 1.2 of [10], the condition $\mathbb{E}X_{ij}^3 = 0$ was only used in the Lemma 2.13 of [10]. Therefore, we only need to prove a stronger version of Lemma 2.13 in [10], i.e., the one without vanishing third moment condition. More precisely, it only remains to prove the following lemma 2.2. (Here we use the same notations as in [10], except for the scaling parameter) $\qquad \square$

Before stating lemma 2.2, i.e., the stronger version of Theorem 1.2 of [10], we introduce some definitions and notations. First, we introduce the notation

$$Y := Y_z := X - zI$$

where I is the identity operator. In the following, we use the notation $A \sim B$ when $cB \leqslant |A| \leqslant c^{-1}B$, where $c > 0$ is independent of N. For any matrix M, we denote M^T as the transpose of M and M^* as the Hermitian conjugate. Usually we choose $z - z_0 \sim N^{-s}$, hence we define the scaled parameter ξ:

$$z = z_0 + N^{-s}\xi, \ i.e., \ \xi := N^s(z - z_0)$$

Define the Green function of $Y_z^* Y_z$ and its trace by, where $w \in \mathbb{C}$ and $\operatorname{Im} w > 0$,

$$G(w) := G(w, z) = (Y_z^* Y_z - w)^{-1}, \quad m(w) := m(w, z) = \frac{1}{N} \operatorname{Tr} G(w, z) = \frac{1}{N} \sum_{j=1}^{N} \frac{1}{\lambda_j(z) - w}. \tag{2.1}$$

Let $m_{\mathrm{c}} := m_{\mathrm{c}}(w, z)$ be the unique solution of

$$m_{\mathrm{c}}^{-1} = -w(1 + m_{\mathrm{c}}) + |z|^2 (1 + m_{\mathrm{c}})^{-1} \tag{2.2}$$

with positive imaginary part. As proved in [9] and [10], for some regions of (w, z) with high probability, $m(w, z)$ converges to $m_{\mathrm{c}}(w, z)$ pointwise, as $N \to \infty$. Let ρ_c be the measure whose Stieltjes transform is m_{c}. This measure is compactly supported and $\operatorname{supp} \rho_c = [\max\{0, \lambda_-\}, \lambda_+]$, where

$$\lambda_\pm := \lambda_\pm(z) := \frac{(\alpha \pm 3)^3}{8(\alpha \pm 1)}, \quad \alpha := \sqrt{1 + 8|z|^2}. \tag{2.3}$$

Note that λ_- has the same sign as $|z| - 1$. It is well-known that $\rho_{\mathrm{c}}(x, z)$ can be obtained from its Stieltjes transform $m_{\mathrm{c}}(x + i\eta, z)$ via

$$\rho_{\mathrm{c}}(x, z) = \frac{1}{\pi} \operatorname{Im} \lim_{\eta \to 0^+} m_{\mathrm{c}}(x + i\eta, z) = \frac{1}{\pi} \mathbb{1}_{x \in [\max\{0, \lambda_-\}, \lambda_+]} \operatorname{Im} \lim_{\eta \to 0^+} m_{\mathrm{c}}(x + i\eta, z).$$

(Some basic properties of m_c and ρ_c were discussed in section 2.2 of [10])

Definition 2.1. ϕ, χ, I and $Z_{X,\mathrm{c}}^{(f)}$
   Let $h(x)$ be a smooth increasing function supported on $[1, +\infty]$ with $h(x) = 1$ for $x \geqslant 2$ and $h(x) = 0$ for $x \leqslant 1$. For any $\varepsilon > 0$, define ϕ on $\mathbb{R}_+$ by (note: λ_+ depends on z)

$$\phi(x) := \phi_{\varepsilon, z}(x) := h(N^{2-2\varepsilon} x) (\log x) \left(1 - h\left(\frac{x}{2\lambda_+} \right) \right). \tag{2.4}$$

Let χ be a smooth cutoff function supported in $[-1, 1]$ with bounded derivatives and $\chi(y) = 1$ for $|y| \leqslant 1/2$. Recall dA denotes the Lebesgue measure on $\mathbb{C}$, for any fixed function g defined on $\mathbb{C}$, we define:

$$Z_{X,\mathrm{c}}^{(g)} := Z_{X,\mathrm{c}}^{(g)}(z_0, \varepsilon, s) := N \int \Delta g(\xi) \int_I \chi(\eta) \phi'(E) \operatorname{Re}(m(w) - m_{\mathrm{c}}(w)) \mathrm{d}E \mathrm{d}\eta \mathrm{dA}(\xi), \quad w = E + i\eta, \quad z = z_0 + N^{-s}\xi$$

and

$$I := I_\varepsilon := \left\{ w \in \mathbb{C} : N^{-1+\varepsilon} \sqrt{E} \leqslant \eta, \ E \geqslant N^{-2+2\varepsilon}, \ |w| \leqslant \varepsilon, \ w = E + i\eta \right\}. \tag{2.5}$$

   Note: the condition $E \geqslant N^{-2+2\varepsilon}$ was not in the definition of the I used in [10], but clearly this condition is implied by $\phi'(E) \neq 0$, i.e., our new I does not change the value of $Z_{X,\mathrm{c}}^{(g)}$. One can also easily check:

$$w \in I_\varepsilon \implies |w|^{1/2} \leqslant 2N^{1-\varepsilon} \eta \tag{2.6}$$

   With these notations and definitions, we claim the following main lemma. It is a stronger version of Lemma 2.13 in [10], i.e., the one without vanishing third moment condition.

Lemma 2.2. *Under the assumptions of Theorem 1.2, there exists a constant $C > 0$ such that for any small enough $\varepsilon > 0$ (independent of N), if $||z_0| - 1| \leqslant \varepsilon$, $s \in (0, 1/2)$, then*

$$Z_{X,c}^{(f)} \prec N^{C\varepsilon} c_f,$$

where c_f is a constant depending only on the function f.

As mentioned above, in the proof of Thm. 1.2 of [10], the vanishing third moment condition was only used in the Lemma 2.13 of [10]. Therefore with the improved Lemma (2.2), one can obtain our main result theorem 1.2 as in [10].

$\square$

In the next step, the lemma 2.2 will be reduced to lemma 2.4.

We note that the bounds proved in [10] for G_{ij}'s are not strong enough for our purpose in this paper. Unfortunately we noticed that it seems impossible to improve these bounds in general cases. On the other hand, we found that though the behaviors G's and $\mathcal{G}$'s are unstable in the region $|m| \leqslant (N\eta)^{-1}$, they are very stable in the region $|m| \gg (N\eta)^{-1}$ and many stronger bounds can be derived in this region. Therefore, in the following proof, we separate the $Z_{X,c}$ into two parts: the one comes for the region $|m| \leqslant (N\eta)^{-1}$ and the one comes for the region $|m| \gg (N\eta)^{-1}$. The first part can be easily bounded, since the m is small, so as its contribution to $Z_{X,c}$. For the second part, we will apply Green's function comparison method (which was first introduced in [15] for generalized Wigner matrix) and our new stronger bounds in the region $|m| \gg (N\eta)^{-1}$.

On the other hand, the old Green's function comparison method was not enough for our purpose, which is also the reason that in [10], the authors needed the extra assumption on the third moment of the matrix entries. In this work, we will introduce an improved Green's function comparison method, which provides an extra $N^{-1/2}$ factor than the previous method. This idea was motivated from the work in [6].

Definition 2.3. t_X *and* $A_X^{(f)}$
For $N \times N$ matrix X, we define

$$t_X := t_X(\varepsilon, w, z) := N^{-\varepsilon} N\eta \operatorname{Re} m$$

i.e.,

$$t_X := N^{-\varepsilon} \eta \operatorname{Re} \operatorname{Tr} \left((X^* - z^*)(X - z) - w \right)^{-1}, \quad \eta = \operatorname{Im} w$$

Now we extend the function h defined in Def. 2.1 to the whole real lane, i.e., $h(x) = h(-x)$, but still use the same notation $h(x)$. With these notations, we define:

$$A_X^{(f)} := A_X^{(f)}(z_0, \varepsilon, s) = N \int \Delta f(\xi) \int_I \chi(\eta) \phi'(E) \left(h(t_X) \operatorname{Re} m - \operatorname{Re} m_c \right) dEd\eta dA(\xi), \tag{2.7}$$

where $z = z_0 + N^{-s}\xi$, $w = E + i\eta$, $\phi = \phi_{\varepsilon,z}$ and $t_X = t_X(\varepsilon, w, z)$.

Note the only difference between $A_X^{(f)}$ and $Z_{X,c}^{(f)}$ is the $h(t_X)$ in front of $\operatorname{Re} m$. Then the difference of $A_X^{(f)}$ and $Z_{X,c}^{(f)}$ only comes from the region $h(t_X) \neq 1$, i.e, $|\operatorname{Re} m| \leqslant 2N^\varepsilon (N\eta)^{-1}$. Therefore, by the definitions of ϕ we have

$$|A_X^{(f)} - Z_{X,c}^{(f)}| \leqslant \int |\Delta f(\xi)| \int_I \chi(\eta) |\phi'(E)| \left(2N^\varepsilon (N\eta)^{-1} \right) dEd\eta dA(\xi) \leqslant N^{C\varepsilon} c_f \tag{2.8}$$

where we used $|(1 - h(t_X)) \operatorname{Re} m| \leqslant 2N^\varepsilon (N\eta)^{-1}$.

Proof of Lemma 2.2: With (2.8), it only remains to prove the following lemma.

Lemma 2.4. *Under the assumptions of Theorem 1.2, there exists a constant $C > 0$ such that for any small enough $\varepsilon > 0$ (independent of N), if $||z_0| - 1| \leqslant \varepsilon$ and $s \in (0, 1/2)$, then*

$$A_X^{(f)} \prec N^{C\varepsilon} c_f$$

where c_f is a constant depending only on the function f.

$\square$

In the next subsection, we will introduce the basic idea of proving Lemma 2.4. The rigorous proof will start from section 3.

2.1 Basic strategy of proving Lemma 2.4: Before we give the complete proof of this lemma, we introduce the basic idea and main improvement in the remainder of this section. Lemma 2.2 was proved in [10] under the vanishing third moment condition. With (2.8), that result implies that if X_{ij}'s are Gaussian variables, for all $1 \leqslant i, j \leqslant N$, then for any fixed $p \in 2\mathbb{N}$,

$$\mathbb{E}|A_{X,\mathrm{c}}^{(f)}|^p \prec N^{C\varepsilon p}, \qquad X_{ij} \sim \mathcal{N}(0, 1/N) \tag{2.9}$$

As one can see that $A_{X,c}^{(f)}$ is basically a linear functional of $m(w, z)$. Hence as in [10], we will apply the Green function comparison method to show that for sufficiently large N,

$$\mathbb{E}|A_X^{(f)}|^p \leqslant C\,\mathbb{E}|A_{X'}^{(f)}|^p + N^{C\varepsilon p}, \tag{2.10}$$

for any two different ensembles X and X' whose matrix elements satisfy the condition of Theorem 1.2. To complete the proof for lemma 2.4, we will choose X' to be the Ginibre ensemble, whose matrix elements are Gaussian variables. The X will be the general ensembles in lemma 2.4. Combining (2.9) and (2.10), with Markov inequality, one immediately obtains Lemma 2.4.

In applying the Green function comparison method, we estimate the expectation value of the functionals of Y, $G = (Y^*Y - w)^{-1}$ and $\mathcal{G} = (YY^* - w)^{-1}$, i.e., $\mathbb{E}F(Y, G, \mathcal{G})$. In [10] and most previous applications of Green function comparison method, one can only bound the expectation value of these functionals with their stochastically dominations. For example, in [10], for $i \neq j$ and $|w|^{1/2} \ll (N\eta)$, one has

$$|(YG)_{ij}| \prec 1$$

With this stochastically domain, the authors in [10] obtained that $|\mathbb{E}(YG)_{ij}| \leqslant N^{\sigma}$ for any $\sigma > 0$. In the present paper, under the condition $|\operatorname{Re} m| \gg (\eta N)^{-1}$, i.e., $h(t_X) > 0$, we will first show an improved bound: for $i \neq j$ and $|w|^{1/2} \ll (N\eta)$

$$|h(t_X)(YG)_{ij}| \prec \sqrt{\frac{|w|^{1/2}}{N\eta}}$$

Then using a new idea on Green's function comparison method, we will show that the expectation value of this term will obtain an extra factor $N^{-1/2}$, i.e.,

$$|\mathbb{E}\,h(t_X)(YG)_{ij}| \leqslant CN^{-1/2+\sigma}\sqrt{\frac{|w|^{1/2}}{N\eta}} \tag{2.11}$$

This extra factor $N^{-1/2}$ plays a key role in our new proof. A similar method was used in the [6].

Now we explain the basic idea of proving (2.11)-type bounds, i.e., where the extra $N^{-1/2}$ factor comes from. For simplicity we assume $X_{ij} \in \mathbb{R}$. Let $Y_z^{(i,i)}$ be the matrix obtained by removing i-th row and column of Y_z, and define

$$G^{(i,i)} := ((Y_z^{(i,i)})^* Y_z^{(i,i)} - w)^{-1}, \quad \mathcal{G}^{(i,i)} := (Y_z^{(i,i)} (Y_z^{(i,i)})^* - w)^{-1}$$

We write $h(t_X)(Y_z G)_{ij}$ as the polynomials of the i-th row/column of X: X_{ik}, X_{ki} $(1 \leqslant k \leqslant N)$, $G^{(i,i)}$ and $\mathcal{G}^{(i,i)}$, i.e.,

$$h(t_X)(Y_z G)_{ij} = P(\{X_{ik}\}_{k=1}^N, \{X_{ki}\}_{k=1}^N, G^{(i,i)}, \mathcal{G}^{(i,i)}) + \text{negligible error}$$

where P is a polynomial. By definition, X_{ik}, X_{ki} are independent of $G^{(i,i)}$ and $\mathcal{G}^{(i,i)}$. In this polynomial, we will show that the degrees of every monomials w.r.t. X_{ik} and X_{ki}'s are always odd numbers. Therefore, in taking the expectation value, with assumption $\mathbb{E} X_{ij} = 0$ and $|\mathbb{E} X_{ij}^k| \leqslant O(N^{-k/2})$, one will see an extra combination factor $N^{-1/2}$. The following simple example will show why the odd powers give an extra factor $N^{-1/2}$. Suppose we estimate $\mathbb{E} \sum_{kst} X_{ik} G_{kl}^{(i,i)} X_{is} G_{st}^{(i,i)} X_{it}$. Since X_{ik}, X_{ki} are independent of $G^{(i,i)}$ and $\mathcal{G}^{(i,i)}$, $\mathbb{E} X_{ij} = 0$ and $|\mathbb{E} X_{ij}^k| \leqslant O(N^{-k/2})$, the nonzero contributions only come from the terms where $k = s = t$, therefore

$$|\mathbb{E} \sum_{kst} X_{ik} G_{kl}^{(i,i)} X_{is} G_{st}^{(i,i)} X_{ti}| = |\mathbb{E} \sum_k X_{ik} G_{kl}^{(i,i)} X_{ik} G_{st}^{(i,i)} X_{ki}| \leqslant C N^{-1/2} \mathbb{E}(\max_{ab} |G_{ab}^{(i,i)}|)^2$$

On the other hand, without $\mathbb{E}$, this term can only be bounded without this $N^{-1/2}$ factor (with large deviation theory).

$$|\sum_{kst} X_{ik} G_{kl}^{(i,i)} X_{is} G_{st}^{(i,i)} X_{ti}| = |\sum_k X_{ik} G_{kl}^{(i,i)}||\sum_{st} X_{is} G_{st}^{(i,i)} X_{ti}| \leqslant (\log N)^C (\max_{ab} |G_{ab}^{(i,i)}|)^2$$

Note: one will not see this $N^{-1/2}$ factor if the degree is even number, e.g., $\mathbb{E} \sum X_{is} G_{st}^{(i,i)} X_{it}$. Based on this new idea, the main task of proving lemma 2.4 and (2.11)-type bounds is writing the functionals of Y_z's, G's and $\mathcal{G}$'s as the polynomials of X_{ik}, X_{ki} $(1 \leqslant k \leqslant N)$, $G^{(i,i)}$ and $\mathcal{G}^{(i,i)}$ for some $1 \leqslant i \leqslant N$, (up to negligible error) and counting the degree of each monomial.

3 PROOF OF LEMMA 2.4

In this section, we apply the Green's function comparison method to prove the Lemma 2.4. We will see the key input of proving Lemma 2.4 is the lemma 3.2. This new lemma is similar to (3.62)-(3.63) of [10], but without the third moments vanishing assumption. More precisely, the (3.62)-(3.63) of [10] is similar to the (3.4) of this work, and lemma 3.2 is the key step of proving (3.4). The proof of lemma 3.2 will start from section 4. In [10], the (3.62)-(3.63) can be easily proved by bounding the expectation value of these terms with their stochastically dominations. In this paper, as introduced in subsection 2.1, we will introduce a new comparison method to show that, for the contribution comes from X_{ij}'s third moment, their expectation values have an extra factor $N^{-1/2}$, i.e., lemma 3.2.

First of all, we state the following lemma. It will be used to estimate the expectation value of some random variables which are stochastically dominated, but not L_∞ bounded.

Lemma 3.1. *Let $v = v^{(N)}$ be a family of centered random variables with variance $1/N$, satisfying the sub exponential decay (1.7). Let $\widetilde{A} = \widetilde{A}^{(N)}$ and $A = A^{(N)}$ be families of random variables. Suppose $A \prec 1$, and*

$A = \sum_{n=0}^{C} A_n v^n$, where $|A_n| \leqslant N^C$ for some fixed constant $C > 0$. We also assume that $\widetilde{A}$ is independent of v and $|\widetilde{A}| \leqslant N^C$ for some $C > 0$. Then for any fixed $p \in \mathbb{N}$ and fixed (small) $\delta > 0$,

$$|\mathbb{E}\,\widetilde{A}A\,v^p| \leqslant (\mathbb{E}|\widetilde{A}|)N^{-p/2+\delta} + N^{-1/\delta}$$

for large enough N.

Note: Here A or A_i's may depend on v.

Proof of Lemma 3.1: By definition 1.1, the assumption $A \prec 1$, and the fact that v has sub exponential decay (1.7), for any fixed $\delta > 0$ and $D > 0$ there is a probability subset Ω such that $\mathbb{P}(\Omega) \geqslant 1 - N^{-D}$ and

$$|1_\Omega Av^p| \leqslant N^{-p/2+\delta}$$

Then

$$|\mathbb{E}\widetilde{A}Av^p| \leqslant (\mathbb{E}|\widetilde{A}|)N^{-p/2+\delta} + |\mathbb{E}_{\Omega^c}\widetilde{A}Av^p| \leqslant (\mathbb{E}|\widetilde{A}|)N^{-p/2+\delta} + O(N^{-D/2+2C})$$

for the second inequality, we used Cauchy Schwarz inequality. Choosing large enough D, we complete the proof of lemma 3.1.

$\square$

Because of this lemma, for any centered random variables v with variance $1/N$, satisfying the sub exponential decay (1.7), we define

$$\mathcal{M}_C(v) := \left\{ A : A = \sum_{n=0}^{C} A_n v^n, \ |A_n| \leqslant N^C \right\} \tag{3.1}$$

Now we return to prove Lemma 2.4.

Proof of Lemma 2.4: For simplicity, we assume that the matrix entries are real numbers. Let X and X' be two ensembles which satisfy the assumption of Theorem 1.2. To prove Lemma 2.4, as we explained in the beginning of subsection 2.1 (near (2.10)), one only needs to show that for any fixed small enough $\varepsilon > 0$, $s \in (0, 1/2)$, and $p \in 2\mathbb{N}$, if $||z_0| - 1| \leqslant \varepsilon$ then

$$\mathbb{E}|A_X^{(f)}|^p \leqslant C\,\mathbb{E}|A_{X'}^{(f)}|^p + N^{C\varepsilon p}, \tag{3.2}$$

for large enough N. For integer k, $0 \leqslant k \leqslant N^2$, define the following matrix X_k interpolating between X' and X:

$$X_k(i,j) = \begin{cases} X(i,j) & \text{if} \quad k \geqslant N(i-1)+j \\ X'(i,j) & \text{if} \quad k < N(i-1)+j \end{cases}.$$

Note that $X' = X_0$ and $X = X_{N^2}$. As one can see that the difference between X_k and X_{k-1} is just one matrix entry. We denote the index of this entry as $(a,b) := (a_k, b_k)$ $(a_k, b_k \in \mathbb{Z}, 1 \leqslant a_k, b_k \leqslant N)$, here $k = (a_k - 1)N + b_k$

Furthermore, we define $t_{X_{k-1}}, t_{X_k}, A_{X_{k-1}}^{(f)}, A_{X_k}^{(f)}$ with X_{k-1} and X_k, as in Def. 2.3. We are going to show that if this special matrix entry is in the diagonal line, i.e., $a = b$ then

$$\left| \mathbb{E}\left(A_{X_k}^{(f)}\right)^p - \mathbb{E}\left(A_{X_{k-1}}^{(f)}\right)^p \right| \leqslant N^{-3/2}\left(N^\varepsilon + 2\mathbb{E}\left(A_{X_{k-1}}^{(f)}\right)^p\right) \tag{3.3}$$

otherwise, i.e., $a \neq b$,

$$\left| \mathbb{E}\left(A_{X_k}^{(f)}\right)^p - \mathbb{E}\left(A_{X_{k-1}}^{(f)}\right)^p \right| \leqslant N^{-2}\left(N^\varepsilon + 2\mathbb{E}\left(A_{X_{k-1}}^{(f)}\right)^p\right) \tag{3.4}$$

for sufficiently large N (independent of k). Clearly, (3.3) and (3.4) imply (3.2).

We are going to compare the these functionals corresponding to X_k and X_{k-1} with a third one, corresponding to the matrix $\widetilde{Q}$ hereafter with deterministic (a, b) entry. We define the following $N \times N$ matrices (hereafter, $Y_\ell = X_\ell - zI$, $\ell = k$ or $k-1$):

$$v = v_{ab}\mathbf{e}_{ab} = X'(a, b)\mathbf{e}_{ab}, \tag{3.5}$$

$$u = u_{ab}\mathbf{e}_{ab} = X(a, b)\mathbf{e}_{ab}, \tag{3.6}$$

$$\widetilde{Q} = X_{k-1} - v = X_k - u, \tag{3.7}$$

$$Q = Y_{k-1} - v = Y_k - u, \tag{3.8}$$

$$R = (Q^*Q - wI)^{-1} \tag{3.9}$$

$$\mathcal{R} = (QQ^* - wI)^{-1} \tag{3.10}$$

$$S = (Y_{k-1}^*Y_{k-1} - wI)^{-1} \tag{3.11}$$

$$T = (Y_k^*Y_k - wI)^{-1} \tag{3.12}$$

Furthermore, we define $t_{\widetilde{Q}}$, $A_{\widetilde{Q}}^{(f)}$ with $\widetilde{Q}$, as in Def. 2.3. To prove (3.3) and (3.4), we will estimate $A_{X_{k-1}}^{(f)} - A_{X_{\widetilde{Q}}}^{(f)}$ and $A_{X_k}^{(f)} - A_{X_{\widetilde{Q}}}^{(f)}$.

First we introduce the notations

$$m_S = \frac{1}{N}\operatorname{Tr} S, \quad m_R = \frac{1}{N}\operatorname{Tr} R, \quad m_T = \frac{1}{N}\operatorname{Tr} T$$

We note: with Cauchy's interlace theorem, for some $C > 0$,

$$|m_S - m_R| \leqslant C(N\eta)^{-1}, \quad \eta = \operatorname{Im} w \tag{3.13}$$

holds for any w and z. It implies

$$|A_{X_{k-1}}^{(f)} - A_{\widetilde{Q}}^{(f)}| \leqslant C \tag{3.14}$$

To estimate $A_{X_{k-1}}^{(f)} - A_{\widetilde{Q}}^{(f)}$, from (2.7), we have

$$A_{X_{k-1}}^{(f)} - A_{\widetilde{Q}}^{(f)} = N \int \Delta f(\xi) \int_I \chi(\eta)\phi'(E)\Big(h(t_{X_{k-1}})\operatorname{Re} m_S - h(t_{\widetilde{Q}})\operatorname{Re} m_R\Big)\mathrm{d}E\mathrm{d}\eta\mathrm{d}A(\xi), \tag{3.15}$$

where $z = z_0 + N^{-s}\xi$, $w = E + i\eta$, $\phi = \phi_{\varepsilon,z}$. Recall $t_{X_{k-1}}$ and $t_{\widetilde{Q}}$ are defined with m_S and m_R respectively. Applying Taylor's expansion on the term $h(t_{X_{k-1}})\operatorname{Re} m_S - h(t_{\widetilde{Q}})\operatorname{Re} m_R$ in (3.15) and letting $h^{(k)}$ be the kth derivative of h, we have

$$h(t_{X_{k-1}})\operatorname{Re} m_S - h(t_{\widetilde{Q}})\operatorname{Re} m_R = \sum_{n=1}^{3} B_n(\widetilde{Q})\left(\operatorname{Re} m_S - \operatorname{Re} m_R\right)^n + B_4(X_{k-1}, \widetilde{Q})\left(\operatorname{Re} m_S - \operatorname{Re} m_R\right)^4 \tag{3.16}$$

where $B_n(\widetilde{Q})$ ($1 \leqslant n \leqslant 3$) and $B_4(X_{k-1}, \widetilde{Q})$ are defined as

$$B_n(\widetilde{Q}) := \frac{1}{n!}(N^{1-\varepsilon}\eta)^{(n-1)}\left(nh^{(n-1)}(t_{\widetilde{Q}}) + h^{(n)}(t_{\widetilde{Q}})t_{\widetilde{Q}}\right) \tag{3.17}$$

$$B_4(X_{k-1}, \widetilde{Q}) := \frac{1}{24}(N^{1-\varepsilon}\eta)^3\left(4h^{(3)}(\zeta) + h^{(4)}(\zeta)\zeta\right)$$

10

where ζ is between $t_{X_{k-1}}$ and $t_{\widetilde{Q}}$, and only depends on $t_{X_{k-1}}$, $t_{\widetilde{Q}}$ and h. As one can see that B_1, B_2 and B_3 are independent of v_{ab}. For the definition of B's, we note that if $n \geqslant 1$, then

$$h^{(n)}(x) \neq 0 \implies x \sim 1$$

Therefore, with $|h| \leqslant 1$, we obtain the following uniform bounds for B's:

$$|B_n| \leqslant (N^{1-\varepsilon}\eta)^{(n-1)}, \quad 1 \leqslant n \leqslant 4 \tag{3.18}$$

To estimate the $m_S - m_R$ in (3.16), we study the difference between m_S and m_R in the parameter set:

$$\left\{ (k, z, w) \in \mathbb{Z} \times \mathbb{C}^2 \ : \ 0 \leqslant k \leqslant N^2, \ ||z| - 1| \leqslant 2\varepsilon, \ w \in I_\varepsilon \right\} \tag{3.19}$$

Recall in (3.59) of [2] and the discussion below (3.61) of [2], it was proved that with the notations:

$$P_1(\widetilde{Q}) := \frac{1}{N} \operatorname{Re} \left(-2(QR^2)_{ab} \right) \tag{3.20}$$

$$P_2(\widetilde{Q}) := \frac{1}{N} \operatorname{Re} \left(w\mathcal{R}_{aa}(R^2)_{bb} + 2(QR^2)_{ab}(RQ^*)_{ba} + (QR^2Q^*)_{aa}R_{bb} \right)$$

$$P_3(\widetilde{Q}) :=$$

$$\frac{1}{N} \operatorname{Re} \left(-2(RQ^*)_{ba}^2(QR^2)_{ab} - 2(RQ^*)_{ba}(QR^2Q^*)_{aa}R_{bb} - 2(RQ^*)_{ba}w\mathcal{R}_{aa}(R^2)_{bb} - 2w\mathcal{R}_{aa}R_{bb}(QR^2)_{ab} \right)$$

the difference between $\operatorname{Re} m_S$ and $\operatorname{Re} m_R$, i.e., $(\frac{1}{N} \operatorname{Re} \operatorname{Tr} S - \frac{1}{N} \operatorname{Re} \operatorname{Tr} R)$ can be written as (recall $v_{ab} = X'(a,b)$)

$$\operatorname{Re} m_S - \operatorname{Re} m_R = \sum_{n=1}^{3} P_n(\widetilde{Q}) \cdot (v_{ab})^3 + P_4(X_{k-1}, \widetilde{Q}) \cdot (v_{ab})^4, \tag{3.21}$$

where $P_4(X_{k-1}, \widetilde{Q})$ depends on X_{k-1} and $\widetilde{Q}$, and the P's can be bounded as

$$P_1(\widetilde{Q}), \ P_2(\widetilde{Q}), \ P_3(\widetilde{Q}), \ P_4(X_{k-1}, \widetilde{Q}) \prec (N\eta)^{-1}, \tag{3.22}$$

uniformly for (k, z, w) in (3.19). In [2], the uniformness was not emphasized, but it can be easily checked. From (3.7)-(3.10) and the definition of $P_{1,2,3}(\widetilde{Q})$, we can see that $P_{1,2,3}(\widetilde{Q})$ only depend on $\widetilde{Q}$ and they are independent of v_{ab}.

Now we collect some simple bounds on P_i's. For L_∞ norm, by definition, it is easy to prove that the following inequalities always hold:

$$\|S\|, \quad \|\mathcal{R}\|, \quad \|R\|, \quad \|R^2\|, \quad \|QR\|, \quad \|QR^2\|, \quad \|QR^2Q^*\| \leqslant N^C$$

for any (k, z, w) in (3.19) and some fixed constant $C > 0$. Then with the definition in (3.20), we also have that for any (k, z, w) in (3.19) and some constant $C > 0$

$$P_1, P_2, P_3 = O(N^C). \tag{3.23}$$

Expanding S around R with the fact: $S = (R^{-1} + (Y_k^*Y_k - Q^*Q))^{-1}$, we obtain that for any fixed $m \in \mathbb{N}$

$$S - R = \sum_{n=1}^{m} \left(-R(Y_k^*Y_k - Q^*Q) \right)^n R + \left(-R(Y_k^*Y_k - Q^*Q) \right)^{m+1} S \tag{3.24}$$

Let $m = 5$ in (3.24). Now we take $\frac{1}{N}\operatorname{Re}\operatorname{Tr}$ on the both sides of (3.24) and compare it with (3.21). Since $Y_k^*Y_k - Q^*Q = v_{ab}(\mathbf{e}_{ba}Q) + v_{ab}(Q^*\mathbf{e}_{ab}) + v_{ab}^2\mathbf{e}_{aa}$, we can see that for $1 \leqslant l \leqslant 3$, the $P_l(Q)$ is the coefficient of the $(v_{ab})^l$ term in the r.h.s. of $\frac{1}{N}\operatorname{Re}\operatorname{Tr}$ (3.24) and

$$P_4(X_{k-1}, \widetilde{Q}) \in \mathcal{M}_C(v_{ab}) \tag{3.25}$$

Similarly, using this expansion ($m = 5$), and the fact:

$$\partial_w R = R^2 = O(N^C), \quad \partial_z R = R(Q + Q^*)R = O(N^C),$$

and $\partial_w S$, $\partial_z S = O(N^C)$, we can improve (3.22) to the following one:

$$\max_{(k,z,w)\in(3.19)} N\eta\left(|P_1(\widetilde{Q})| + |P_2(\widetilde{Q})| + |P_3(\widetilde{Q})| + |P_4(X_{k-1}, \widetilde{Q})|\right) \prec 1 \tag{3.26}$$

We note: this statement shows that (3.22) can hold for different $(k, z, w) \in$ (3.19) with the same probability subset.

Inserting (3.16) and (3.21) into (3.15), we write $A_{X_{k-1}}^{(f)} - A_Q^{(f)}$ as a polynomial of v_{ab} as follows.

$$A_{X_{k-1}}^{(f)} - A_{\widetilde{Q}}^{(f)} = \mathscr{P}_1(\widetilde{Q}) \cdot v_{ab} + \mathscr{P}_2(\widetilde{Q}) \cdot (v_{ab})^2 + \mathscr{P}_3(\widetilde{Q}) \cdot (v_{ab})^3 + \mathscr{P}_4(X_{k-1}, \widetilde{Q}) \cdot (v_{ab})^4. \tag{3.27}$$

where

$$\mathscr{P}_1(\widetilde{Q}) := N\int \Delta f(\xi) \int_I (B_1 P_1)\chi(\eta)\phi'(E)\mathrm{d}E\mathrm{d}\eta\mathrm{d}A(\xi) \tag{3.28}$$

$$\mathscr{P}_2(\widetilde{Q}) := N\int \Delta f(\xi) \int_I (B_1 P_2 + B_2 P_1^2)\chi(\eta)\phi'(E)\mathrm{d}E\mathrm{d}\eta\mathrm{d}A(\xi)$$

$$\mathscr{P}_3(\widetilde{Q}) := N\int \Delta f(\xi) \int_I \left(B_1 P_3 + 2B_2 P_1 P_2 + B_3 P_1^3\right)\chi(\eta)\phi'(E)\mathrm{d}E\mathrm{d}\eta\mathrm{d}A(\xi)$$

$$\mathscr{P}_4(X_{k-1}, \widetilde{Q}) := N\int \Delta f(\xi) \int_I \left(\sum_n B_n \sum_{\sum_j i_j \geqslant 4} (v_{ab})^{(\sum_j i_j)-4} \prod_{j=1}^n P_{i_j}\right)\chi(\eta)\phi'(E)\mathrm{d}E\mathrm{d}\eta\mathrm{d}A(\xi)$$

where $B_n = B_n(\widetilde{Q})$, $P_n = P_n(\widetilde{Q})$ ($1 \leqslant n \leqslant 3$), $B_4 = B_4(X_{k-1}, \widetilde{Q})$ and $P_4 = P_4(X_{k-1}, \widetilde{Q})$. We note: $\mathscr{P}_1(\widetilde{Q})$, $\mathscr{P}_2(\widetilde{Q})$ and $\mathscr{P}_3(\widetilde{Q})$ are independent of v_{ab}.

Replacing X_{k-1} with X_k, with the same method, we obtain (Here v_{ab} is replaced with u_{ab})

$$A_{X_k}^{(f)} - A_{\widetilde{Q}}^{(f)} = \mathscr{P}_1(\widetilde{Q})u_{ab} + \mathscr{P}_2(\widetilde{Q})u_{ab}^2 + \mathscr{P}_3(\widetilde{Q})u_{ab}^3 + \mathscr{P}_4(X_k, \widetilde{Q})u_{ab}^4 \tag{3.29}$$

From (3.18) and (3.26), it is easy to check that $\mathscr{P}_1$, $\mathscr{P}_2$ and $\mathscr{P}_3 \prec 1$ uniformly hold for $1 \leqslant k \leqslant N^2$. For L_∞ bound, with (3.23), they are bounded by N^C for some C. Similarly, we can obtain that $\mathscr{P}_4 \prec 1$. With (3.25), we have $\mathscr{P}_4(X_{k-1}, \widetilde{Q}) \in \mathcal{M}_C(v_{ab})$ and $\mathscr{P}_4(X_k, \widetilde{Q}) \in \mathcal{M}_C(u_{ab})$. So far, we proved

$$\mathscr{P}_{1,2,3,4} \prec 1, \quad \mathscr{P}_{1,2,3}(\widetilde{Q}) \leqslant N^C, \quad \mathscr{P}_4(X_{k-1}, \widetilde{Q}) \in \mathcal{M}_C(v_{ab}), \quad \mathscr{P}_4(X_k, \widetilde{Q}) \in \mathcal{M}_C(u_{ab}) \tag{3.30}$$

uniformly hold for $1 \leqslant k \leqslant N^2$.

12

Now we return to prove (3.3) and (3.4). First we write

$$(A_{X_{k-1}}^{(f)})^p - (A_{X_k}^{(f)})^p = \sum_{j=0}^{p-1} \binom{p}{j} (A_{\widetilde{Q}}^{(f)})^j \left((A_{X_{k-1}}^{(f)} - A_{\widetilde{Q}}^{(f)})^{p-j} - (A_{X_k}^{(f)} - A_{\widetilde{Q}}^{(f)})^{p-j} \right).$$

We insert the (3.27) and (3.29) into the r.h.s. and write it in the following form

$$(A_{X_{k-1}}^{(f)})^p - (A_{X_k}^{(f)})^p = \sum_{m=1}^{4p} (\mathcal{A}_m v_{ab}^m - \mathcal{B}_m u_{ab}^m) \tag{3.31}$$

where $\mathcal{A}_m$ only contains $A_{\widetilde{Q}}^{(f)}$, $\mathscr{P}_{1,2,3}(\widetilde{Q})$, $\mathscr{P}_4(X_{k-1}, \widetilde{Q})$, and $\mathcal{B}_m$ only contains $A_{\widetilde{Q}}^{(f)}$, $\mathscr{P}_{1,2,3}(\widetilde{Q})$, $\mathscr{P}_4(X_k, \widetilde{Q})$ For example,

$$\mathcal{A}_3 = \mathcal{B}_3 = C_{p,3}(A_{\widetilde{Q}}^{(f)})^{p-3}\mathscr{P}_1^3(\widetilde{Q}) + C_{p,2}(A_{\widetilde{Q}}^{(f)})^{p-2}\mathscr{P}_1(\widetilde{Q})\mathscr{P}_2(\widetilde{Q}) + C_{p,1}(A_{\widetilde{Q}}^{(f)})^{p-1}\mathscr{P}_3(\widetilde{Q}) \tag{3.32}$$

where $C_{p,n}$ ($1 \leqslant n \leqslant 3$) are constants only depends on p. Since the first two moments of v_{ab} and u_{ab} coincide, u_{ab}, v_{ab} are independent of $\widetilde{Q}$, and $\mathcal{A}_1 = \mathcal{B}_1$, $\mathcal{A}_2 = \mathcal{B}_2$ only contain $A_{\widetilde{Q}}^{(f)}$, $\mathscr{P}_{1,2,3}(\widetilde{Q})$, we have

$$\mathbb{E}(A_{X_{k-1}}^{(f)})^p - \mathbb{E}(A_{X_k}^{(f)})^p = \sum_{m=3}^{4p} (\mathcal{A}_m v_{ab}^m - \mathcal{B}_m u_{ab}^m)$$

Recall the definition of $\mathcal{A}_m$ and $\mathcal{B}_m$ from (3.31), for the terms $m \geqslant 4$, using (3.30) and Lemma 3.1, we get

$$|\mathbb{E} \sum_{m=4}^{4p} (\mathcal{A}_m v_{ab}^m - \mathcal{B}_m u_{ab}^m)| \leqslant \sum_{j=0}^{p-1} \mathbb{E} \left| (A_{\widetilde{Q}}^{(f)})^j \right| O_{\prec}(N^{-2}) + N^{-2} \leqslant N^{-2} \left(O_{\prec}(1) + \mathbb{E}|A_{\widetilde{Q}}^{(f)}|^p \right).$$

Therefore, with $\mathcal{A}_3 = \mathcal{B}_3$,

$$\left| \mathbb{E}\left(A_{X_{k-1}}^{(f)} \right)^p - \mathbb{E}\left(A_{X_k}^{(f)} \right)^p \right| \leqslant N^{-2} \left(O_{\prec}(1) + \mathbb{E}|A_{\widetilde{Q}}^{(f)}|^p \right) + |\mathbb{E}\mathcal{A}_3| \left(|\mathbb{E}v_{ab}^3| + |\mathbb{E}u_{ab}^3| \right), \tag{3.33}$$

Similarly, using (3.30), (3.32), $A_{\widetilde{Q}}^{(f)} = O(N^C)$ and Lemma 3.1, we have

$$|\mathbb{E}\mathcal{A}_3| \left(|\mathbb{E}v_{ab}^3| + |\mathbb{E}u_{ab}^3| \right) \leqslant N^{-3/2} \left(O_{\prec}(1) + \mathbb{E}|A_{\widetilde{Q}}^{(f)}|^p \right) \tag{3.34}$$

As in [2]-(3.64), using Hölder's inequality and the bound (3.14), we have

$$\mathbb{E}|A_{\widetilde{Q}}^{(f)}|^p \leqslant \mathbb{E}\left((A_{X_{k-1}}^{(f)})^p \right) + \sum_{j=1}^p \binom{j}{p} \mathbb{E}\left(\left| A_{X_{k-1}}^{(f)} \right|^{p-j} |A_{X_{k-1}}^{(f)} - A_{\widetilde{Q}}^{(f)}|^j \right) \tag{3.35}$$

$$\leqslant \mathbb{E}\left((A_{X_{k-1}}^{(f)})^p \right) + \sum_{j=1}^p \binom{j}{p} \mathbb{E}\left(\left| A_{X_{k-1}}^{(f)} \right|^p \right)^{\frac{p-j}{p}} \mathbb{E}\left(\left| A_{X_{k-1}}^{(f)} - A_{\widetilde{Q}}^{(f)} \right|^p \right)^{\frac{j}{p}},$$

$$\leqslant \left(O_{\prec}(1) + 2\mathbb{E}|A_{X_{k-1}}^{(f)}|^p \right),$$

Then combining (3.33)-(3.35), we obtain (3.3). (Note: $p \in 2\mathbb{Z}$.)

To prove (3.4), we claim the following lemma, which provides the stronger bound on the expectation value of the r.h.s. of (3.32).

Lemma 3.2. *Assume $1 \leqslant a \neq b \leqslant N$. Let X be defined as in Theorem 1.2, except that $X_{ab} = 0$. For any fixed small enough $\varepsilon > 0$, if $||z_0| - 1| \leqslant \varepsilon$ and $s \in (0, 1/2)$, define $A_X^{(f)}$, $P_i(X)$, $B_i(X)$, $\mathscr{P}_i(X)$, $i = 1, 2, 3$ as in (2.7), (3.20), (3.17) and (3.28). (More precisely $\widetilde{Q}$, Q, R in (3.20) and (3.17) will be replaced with X, $Y = X - zI$ and $(Y^*Y - wI)$ respectively.) Then*

$$\left| \mathbb{E}(A_X^{(f)})^{p-3} \mathscr{P}_1^3(X) \right| + \left| \mathbb{E}(A_X^{(f)})^{p-2} \mathscr{P}_1(X) \mathscr{P}_2(X) \right| + \left| \mathbb{E}(A_X^{(f)})^{p-1} \mathscr{P}_3(X) \right| \prec N^{-1/2} \left(O_{\prec}(1) + \mathbb{E}|A_X^{(f)}|^p \right) \tag{3.36}$$

uniformly for (a, b).

We return to prove (3.33) and prove lemma (3.2) in the next section. Inserting this lemma and (3.32) into (3.33), as in (3.34), we obtain that if $a \neq b$, then

$$|\mathbb{E}\mathcal{A}| \leqslant N^{-2} \left(O_{\prec}(1) + \mathbb{E}|A_Q^{(f)}|^p \right) \tag{3.37}$$

Together with (3.33) and (3.35), we obtain (3.4). Clearly, (3.3) and (3.4) imply (3.2), and we complete the proof of lemma 2.4 and lemma 2.2.

$\square$

4 PROOF OF LEMMA 3.2

Lemma 3.2 bounds the expectation values of some polynomials of $A_X^{(f)}$ and $\mathscr{P}_{1,2,3}(X)$. Roughly speaking Lemma 3.2 shows that the expectation value of these polynomials are much less than their stochastic domination by a factor $N^{-1/2}$. (Note: a and b appear in the definitions of $P_{1,2,3}$ and $B_{1,2,3}$. The $\mathscr{P}_{1,2,3}$ are defined with $P_{1,2,3}$ and $B_{1,2,3}$.) As introduced in the second part of subsection 2.1 (below (2.11)), the main strategy of showing this extra factor is

- writing them as the polynomials (up to negligible error) of X_{ak}'s, X_{ka}'s ($1 \leqslant k \leqslant N$), $G^{(a,a)}$ and $\mathcal{G}^{(a,a)}$, which are defined as

$$G^{(a,a)} := ((Y_z^{(a,a)})^* Y_z^{(a,a)} - w)^{-1}, \quad \mathcal{G}^{(a,a)} := (Y_z^{(a,a)}(Y_z^{(a,a)})^* - w)^{-1}$$

and $Y^{(a,a)} := Y_z^{(a,a)}$ is the matrix obtained by removing the i−th row and column of Y_z.

- showing the degrees of the monomials of X_{ak}'s and X_{ka}'s in above polynomials are always odd (except for X_{aa}).

First of all, in lemma 4.5 and 4.7, we introduce some polynomials having the properties we need for Lemma 3.2, i.e., their expectation values have an extra factor $N^{-1/2}$ comparing with their stochastic domination. In the next subsection, we introduce some $\mathcal{F}$ sets, whose elements are the "basic" polynomials in our proof, i.e., the bricks of the polynomials in lemma 4.5 and 4.7.

4.1 Basic polynomials and their properties. We first introduce some notations.

Definition 4.1. $X^{(\mathbb{T},\mathbb{U})}$, $Y^{(\mathbb{T},\mathbb{U})}$, $G^{(\mathbb{T},\mathbb{U})}$ and $\mathcal{G}^{(\mathbb{T},\mathbb{U})}$

Let $\mathbb{T}, \mathbb{U}$ be some subsets of $\{1, 2, \cdots, N\}$. Then we define $Y^{(\mathbb{T},\mathbb{U})}$ as the $(N - |\mathbb{U}|) \times (N - |\mathbb{T}|)$ matrix obtained by removing all columns of Y indexed by $i \in \mathbb{T}$ and all rows of Y indexed by $i \in \mathbb{U}$. Notice that we keep the labels of indices of Y when defining $Y^{(\mathbb{T},\mathbb{U})}$. With the same method, we define $X^{(\mathbb{T},\mathbb{U})}$ with X.

14

Let $\mathbf{y}_i$ be the i-th column of Y and $\mathbf{y}_i^{(\mathbb{S})}$ be the vector obtained by removing $\mathbf{y}_i(j)$ for all $j \in \mathbb{S}$. Similarly we define y_i be the i-th row of Y. Define

$$G^{(\mathbb{T},\mathbb{U})} = \left[(Y^{(\mathbb{T},\mathbb{U})})^* Y^{(\mathbb{T},\mathbb{U})} - w \right]^{-1}, \quad m_G^{(\mathbb{T},\mathbb{U})} = \frac{1}{N} \operatorname{Tr} G^{(\mathbb{T},\mathbb{U})},$$

$$\mathcal{G}^{(\mathbb{T},\mathbb{U})} = \left[Y^{(\mathbb{T},\mathbb{U})} (Y^{(\mathbb{T},\mathbb{U})})^* - w \right]^{-1}, \quad m_{\mathcal{G}}^{(\mathbb{T},\mathbb{U})} = \frac{1}{N} \operatorname{Tr} \mathcal{G}^{(\mathbb{T},\mathbb{U})}.$$

By definition, $m^{(\emptyset,\emptyset)} = m$. Since the eigenvalues of Y^*Y and YY^* are the same except the zero eigenvalue, it is easy to check that

$$m_G^{(\mathbb{T},\mathbb{U})}(w) = m_{\mathcal{G}}^{(\mathbb{T},\mathbb{U})} + \frac{|\mathbb{U}| - |\mathbb{T}|}{Nw} \tag{4.1}$$

For $|\mathbb{U}| = |\mathbb{T}|$, we define

$$m^{(\mathbb{T},\mathbb{U})} := m_G^{(\mathbb{T},\mathbb{U})} = m_{\mathcal{G}}^{(\mathbb{T},\mathbb{U})} \tag{4.2}$$

There is a crude bound for $(m_G^{(\mathbb{T},\mathbb{U})} - m)$ proved in (6.6) of [9]:

$$\left| m_G^{(\mathbb{T},\mathbb{U})} - m \right| + \left| m_{\mathcal{G}}^{(\mathbb{T},\mathbb{U})} - m \right| \leqslant C \frac{|\mathbb{T}| + |\mathbb{U}|}{N\eta} \tag{4.3}$$

Definition 4.2. Notations for general sets.

As usual, if $x \in \mathbb{R}$ or $\mathbb{C}$, and $\mathcal{S}$ is a set of random variables then $x\mathcal{S}$ denotes the following set as

$$x\mathcal{S} := \{xs : s \in \mathcal{S}\}$$

For two sets $\mathcal{S}_1$ and $\mathcal{S}_2$ of random variables, we define the following set as

$$\mathcal{S}_1 \cdot \mathcal{S}_2 := \left\{ s_1 \cdot s_2 \mid s_1 \in \mathcal{S}_1, \ s_2 \in \mathcal{S}_2 \right\}$$

For simplicity, we call $s \in_n \mathcal{S}$ if and only if s can be written as the sum of $O(1)$ elements in $\mathcal{S}$, i.e.,

$$s \in_n \mathcal{S} \iff s \in \left\{ \sum_{i=1}^n s_i \mid s_i \in \mathcal{S}, \ n \in \mathbb{N}, \ n = O(1) \right\}$$

Definition 4.3. Definition of $\mathcal{F}_0$, $\mathcal{F}_1$, $\mathcal{F}_{1/2}$ and $\mathcal{F}$.

For fixed indeces a, b and ensemble X in lemma 3.2, we define $\mathcal{F}_0$ as the set of random variables (depending on X) which are stochastically dominated by 1 and independent of any X_{ak} and X_{ka} ($1 \leqslant k \leqslant N$), i.e.,

$$\mathcal{F}_0 = \{V \ : \ V \prec 1, \ V \text{ is independent of the } a-\text{th row and column of X}\}$$

Note: $\mathcal{F}_0$ depends on a, not b. One example element in $\mathcal{F}_0$ is $\operatorname{Tr} X - X_{aa}$.

For simplicity, we define

$$\sum_i^{(a)} := \sum_{i \neq a}, \qquad \sum_{ij}^{(a)} := \sum_{ij \neq a}$$

15

Next we define $\mathcal{F}_1$ as the union of the set $(N^{1/2}X_{aa}\mathcal{F}_0)$ and the sets of some quadratic forms as follows

$$\mathcal{F}_1 := \left(N^{1/2}X_{aa}\mathcal{F}_0\right)\bigcup\left\{\overset{(a)}{\sum_{kl}}X_{ka}V_{kl}X_{la} \text{ or } \overset{(a)}{\sum_{kl}}X_{ak}V_{kl}X_{al} \ \middle| \ \max_{kl}|V_{kl}| \prec 1, \ V_{kl} \in \mathcal{F}_0\right\}$$

$$\bigcup\left\{\overset{(a)}{\sum_{k\neq l}}X_{ak}V_{kl}X_{la} + N^{1/2}\overset{(a)}{\sum_{k}}X_{ak}V_{kk}X_{ka} \ \middle| \ \max_{kl}|V_{kl}| \prec 1, \ V_{kl} \in \mathcal{F}_0\right\}$$

(Note it is $X_{ka}V_{kl}X_{la}$ or $X_{ak}V_{kl}X_{al}$ in the first line and $X_{ak}V_{kl}X_{la}$ in the second line, and the diagonal terms in the second case is allowed to be larger than the others by a factor $N^{1/2}$.)

Furthermore, we define $\mathcal{F}$ as the set of following random variables

$$\mathcal{F} := \left\{V \ \middle| \ V \in_n \mathcal{F}_0\bigcup\left(\bigcup_{n=O(1)}(\mathcal{F}_1)^n\right)\right\}$$

where $(\mathcal{F}_1)^n$ represents the set of the products of n elements in $\mathcal{F}_1$. For simplicity, sometimes we write $\mathcal{F}$

$$\mathcal{F} = \mathcal{F}_\emptyset$$

i.e., with the subscription empty set $\emptyset$.
Similarly, we define

$$\mathcal{F}_{1/2} = \left\{\overset{(a)}{\sum_{k}}X_{ak}V_k \text{ or } \overset{(a)}{\sum_{k}}V_kX_{ka} \ \middle| \ \max_{k}|V_k| \prec 1, \ V_k \in \mathcal{F}_0\right\}$$

Note: For fixed $k \neq a$, the total number of X_{ak} and X_{ka} $(1 \leqslant k \leqslant N)$, in each monomial of the element in $\mathcal{F}$ is always even. On the other hand, this number in $\mathcal{F}_{1/2} \cdot \mathcal{F}$ is always odd. By the definition, it is easy to see that

$$\mathcal{F}_0, \ \mathcal{F}_1 \in \mathcal{F}$$

$$\mathcal{F}_0 \cdot \mathcal{F}_\alpha = \mathcal{F}_\alpha, \quad \alpha = 0, \ 1/2, \ 1, \ \emptyset$$

and

$$\mathcal{F}_{1/2} \cdot \mathcal{F}_{1/2} \subset \mathcal{F}_1, \quad \mathcal{F} \cdot \mathcal{F} \subset \mathcal{F} \tag{4.4}$$

Examples: by definition, $G_{kl}^{(a,a)} \leqslant \eta^{-1}$ for any $k, l \neq a$. Hence we have

$$\overset{(a)}{\sum_{kl}}X_{ka}G_{kl}^{(a,a)}X_{la} \in \eta^{-1}\mathcal{F}_1,$$

$$\left(\overset{(a)}{\sum_{kl}}X_{ka}G_{kl}^{(a,a)}X_{la}\right)\left(\overset{(a)}{\sum_{kl}}X_{ak}G_{kl}^{(a,a)}X_{al}\right) \in \eta^{-2}\mathcal{F},$$

16

and if $\eta = O(1)$

$$\left(\sum_{kl}^{(a)} X_{ka} G_{kl}^{(a,a)} X_{la} \right) \left(\sum_{kl}^{(a)} X_{ak} G_{kl}^{(a,a)} X_{al} \right) + (\operatorname{Tr} X - X_{aa}) \in_n \eta^{-2} \mathcal{F}$$

Definition 4.4. Uniformness *Let F_T, $T \in \mathcal{T}_N$ be a family of random variables, where $\mathcal{T}_N$ is parameter set which may depends on N. We say*

$$F_T \in_n \mathcal{F}, \quad T \in \mathcal{T}_N,$$

are **uniform** *for all $T \in \mathcal{T}_N$, if the following two uniform conditions hold.*

(i) *There exist uniform integers m and n independent of N such that for all $T \in \mathcal{T}_N$, we can write F_T as the sum of m elements in $(\mathcal{F}_0 \cup (\mathcal{F}_1)^n)$, i.e.,*

$$F_T = \sum_{i=1}^{m} F_{T,i}, \quad F_{T,i} \in \mathcal{F}_0 \cup (\mathcal{F}_1)^n.$$

(ii) *All of the stochastic domination relations, i.e., $\prec$, appearing in all F_T's $(T \in \mathcal{T}_N)$ hold uniformly.*

Similarly, for $\mathcal{F}_0$, $\mathcal{F}_{1/2}$ and $\mathcal{F}_1$, we call

$$F_T \in_n \mathcal{F}_\alpha, \quad T \in \mathcal{T}_N, \quad \alpha = 0, \ \frac{1}{2}, \ 1$$

uniformly for all $T \in \mathcal{T}_N$, if there exist uniform m independent of N such that

$$F_T = \sum_{i=1}^{m} F_{T,i}, \quad F_{T,i} \in \mathcal{F}_\alpha, \quad \alpha = 0, \ \frac{1}{2}, \ 1$$

and the above uniform condition (ii) holds.

More general, if $\mathcal{F}_\alpha$ is one of $\mathcal{F}_0$, $\mathcal{F}_{1/2}$, $\mathcal{F}_1$, $\mathcal{F}$, so as $\mathcal{F}_\beta$, i.e., $\alpha, \beta = 0, 1/2, 1$ or $\emptyset$, we say

$$F_T \in_n \mathcal{F}_\alpha \cdot \mathcal{F}_\beta,$$

uniformly for all $T \in \mathcal{T}_N$ if there exists uniform m independent of N such that F_T can be written as the sum of the m terms in $\mathcal{F}_\alpha \cdot \mathcal{F}_\beta$, i.e.,

$$F_T = \sum_{i=1}^{m} F_{T,\alpha,i} F_{T,\beta,i}$$

and

$$F_{T,\alpha,i} \in \mathcal{F}_\alpha, \quad F_{T,\beta,i} \in \mathcal{F}_\beta$$

hold uniformly for all $T \in \mathcal{T}_N$.

Furthermore, with fixed $D > 0$ and random (or deterministic) variable a_T, we say

$$F_T \in_n a_T \mathcal{F}_\alpha \cdot \mathcal{F}_\beta + O_\prec(N^{-D}), \qquad \mathcal{F}_\alpha, \ \mathcal{F}_\beta = \mathcal{F}_0, \ \mathcal{F}_{1/2}, \ \mathcal{F}_1, \ \mathcal{F}$$

uniformly for all $T \in \mathcal{T}_N$ if F_T can be written as

$$F_T = a_T F_{T,1} + F_{T,2}$$

where

$$F_{T,1} \in_n \mathcal{F}_\alpha \cdot \mathcal{F}_\beta, \quad \text{and} \quad F_{T,2} \prec N^{-D}$$

hold uniformly for all $T \in \mathcal{T}_N$.

Now we estimate the expectation values of the elements in $\mathcal{F} \cdot \mathcal{F}_{1/2}$. Let $F_{1/2} \in \mathcal{F}_{1/2}$, $F \in \mathcal{F}$. With large deviation theory, we can only obtain

$$F_{1/2} \prec 1, \quad F \prec 1, \quad F_{1/2} \cdot F \prec 1$$

But we will show that the elements in $\mathcal{F}_{1/2} \cdot \mathcal{F}$ may have much smaller expectation value.

Lemma 4.5. *For fixed indeces a, b and ensemble X in lemma 3.2, let F_0 and F be two random variables bounded by N^C for some C, i.e.,*

$$|F_0| + |F| \leqslant N^C$$

We assume that

$$F_0 \in N^C \mathcal{F}_0, \quad \text{and} \quad F \in_n \mathcal{F}_{1/2} \cdot \mathcal{F}$$

Then we have

$$|\mathbb{E}\, F_0 F| \prec N^{-1/2} \mathbb{E}\, |F_0| + N^{-D} \tag{4.5}$$

for any fixed $D > 0$.

Proof of Lemma 4.5: For simplicity, we assume $F \in \mathcal{F}_{1/2} \cdot \mathcal{F}$ (not $\in_n$). The general case can be proved with the same method. Furthermore, by definition, $\mathbb{E}F_0 F = 0$ if $F \in \mathcal{F}_{1/2} \cdot \mathcal{F}_0$. Hence one only needs to prove the following case: for some fixed m, $F \in \mathcal{F}_{1/2} \cdot (\mathcal{F}_1)^m$, i.e., F can be written as the product of one element of $\mathcal{F}_{1/2}$ and m elements of $\mathcal{F}_1$, i.e.,

$$F = F_{1/2} F_1 F_2 F_3 \cdots F_m, \quad F_{1/2} \in \mathcal{F}_{1/2}, \ F_i \in \mathcal{F}_1, \quad 1 \leqslant i \leqslant m$$

By definition, $F_{1/2} F_1 F_2 F_3 \cdots F_m$ can be consider as a polynomials of X_{ak}'s and X_{ka}'s ($1 \leqslant k \leqslant N$), whose coefficients are independents of the a-th row and column of X. Then, we can decompose F as

$$F = F_{1/2} F_1 F_2 F_3 \cdots F_m \tag{4.6}$$

$$= \sum_{n \leqslant 2m+1} \sum_{k_1, k_2, \ldots, k_n} \sum_{s_1, \ldots, s_n} \sum_{t_1, \ldots, t_n} \mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big) \left(\prod_{i=1}^n (X_{ak_i})^{s_i} (X_{k_i a})^{t_i}\right)$$

where k_i's are all different in the summation, and $\mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big)$ is the coefficient of $\prod_{i=1}^n (X_{ak_i})^{s_i} (X_{k_i a})^{t_i}$ and it is independent of the a-th row and column of X. We separate the parameter region into two cases.

First case: $k_i \neq a$ for all $1 \leqslant i \leqslant n$. By definition of $\mathcal{F}_1$, we have

$$\mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big) \prec \mathbf{1}\left(\sum s_i + \sum t_i = 2m+1\right) \prod_{i=1}^n (N^{1/2})^{\min\{s_i, t_i\}} \tag{4.7}$$

18

where the last factor come from the $N^{1/2}$ factor in the definition of $\mathcal{F}_1$ (see the $N^{1/2} \sum_k^{(a)} X_{ak} V_{kk} X_{ka}$ term in the definition of $\mathcal{F}_1$.

Second case: $k_j = a$ for some $1 \leqslant j \leqslant n$. Since the k_i's are all different, hence the other k_i's are not equal to a. Let $s_j = s$, $t_j = 0$, we have

$$\mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big) \prec \mathbf{1}\Bigg(\sum_{i:\, i \neq j}(s_i + t_i) \in 2\mathbb{N} + 1\Bigg) \prod_{i:\, i \neq j}(N^{1/2})^{\min\{s_i, t_i\}} N^{s/2} \tag{4.8}$$

By definition of $\mathcal{F}_1$ and $\mathcal{F}$, we know that for any $\delta > 0$ and $D > 0$, there exists probability set Ω, which is independent of the a-th row and column of X, such that $\mathbb{P}(\Omega) \geqslant 1 - N^{-D}$, and the $\prec$'s in (4.7) and (4.8) can be replaced with $\leqslant$. More precisely,

$$\mathbf{1}_\Omega |\mathcal{A}_{\text{first case}}| \leqslant N^\delta \cdot \text{r.h.s of } (4.7), \quad \mathbf{1}_\Omega |\mathcal{A}_{\text{second case}}| \leqslant N^\delta \cdot \text{r.h.s of } (4.8) \tag{4.9}$$

With this Ω and $|F_0| + |F| \leqslant N^C$, we have

$$\mathbb{E} F_0 F = \mathbb{E} \mathbf{1}_\Omega F_0 F + \mathbb{E} \mathbf{1}_{\Omega^c} F_0 F = \mathbb{E} \mathbf{1}_\Omega F_0 F + O(N^{3C-D}) \tag{4.10}$$

Hence to prove (4.5), we only need to bound $\mathbb{E} \mathbf{1}_\Omega F_0 F$. For the first case, i.e., $k_i \neq a$ $(1 \leqslant i \leqslant n)$, using (4.9), and the fact that F_0 and Ω are independent of the a-th row and column of X, we have

$$\mathbb{E} \sum_n \sum_{k_1, k_2, \dots, k_n}^{(a)} \sum_{s_1, \dots, s_n} \sum_{t_1, \dots, t_n} \mathbf{1}_\Omega F_0 \mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big) \left(\prod_{i=1}^n (X_{ak_i})^{s_i}(X_{k_i a})^{t_i}\right)$$

$$= \sum_n \sum_{k_1, k_2, \dots, k_n}^{(a)} \sum_{s_1, \dots, s_n} \sum_{t_1, \dots, t_n} \mathbb{E} \mathbf{1}_\Omega F_0\, \mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big) \mathbb{E}\left(\prod_{i=1}^n (X_{ak_i})^{s_i}(X_{k_i a})^{t_i}\right)$$

$$\leqslant \sum_n \sum_{\{s_i\}} \sum_{\{t_i\}} \mathbf{1}\left(\sum s_i + \sum t_i = 2m + 1\right)\left(\prod_{i=1}^n \mathbf{1}(s_i \neq 1)\,\mathbf{1}(t_i \neq 1)\,\mathbf{1}(s_i + t_i \neq 0)(N^{-1/2})^{\max\{s_i, t_i\} - 2}\right)(\mathbb{E}|F_0|)\, N^\delta$$

for any $\delta > 0$, where the factor $(N^{-1/2})^{-2} = N^1$ comes from summation of $k_i : 1 \leqslant k_i \leqslant N$. It is easy to check:

$$\prod_i \mathbf{1}(s_i \neq 1)\,\mathbf{1}(t_i \neq 1)\,\mathbf{1}(s_i + t_i \neq 0)(N^{-1/2})^{\max\{s_i, t_i\} - 2} \leqslant (N^{-1/2})^{\mathbf{1}(s_i + t_i \in 2\mathbb{N} - 1)}$$

Therefore, for any $\delta > 0$,

$$\mathbb{E} \sum_n \sum_{k_1, k_2, \dots, k_n}^{(a)} \sum_{s_1, \dots, s_n} \sum_{t_1, \dots, t_n} \mathbf{1}_\Omega F_0 \mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big)\left(\prod_{i=1}^n (X_{ak_i})^{s_i}(X_{k_i a})^{t_i}\right) \leqslant (\mathbb{E}|F_0|)\, N^{-1/2+\delta} \tag{4.11}$$

Similarly for the second case: without loss of generality, we assume $k_1 = a$. Then as above, using (4.9), and

19

the fact Ω independent of the a-th row and column of X, we have

$$
\mathbb{E} \sum_n \sum_{k_2,\ldots,k_n \neq a} \sum_s \sum_{s_2,\ldots,s_n} \sum_{t_2,\ldots,t_n} \mathbf{1}_\Omega F_0 \mathcal{A}\Big(\{k_i\}_{i=1}^n, \{s_i\}_{i=1}^n, \{t_i\}_{i=1}^n\Big) \left(\prod_{i \neq 1}^n (X_{ak_i})^{s_i}(X_{k_i a})^{t_i}\right)(X_{aa})^s
$$

$$
\leqslant \sum_n \sum_{\{s_i\}} \sum_{\{t_i\}} \mathbf{1}\left(\sum_{i \geqslant 2}(s_i + t_i) \in 2\mathbb{N}+1\right)\left(\prod_{i \geqslant 2} \mathbf{1}\,(s_i \neq 1)\,\mathbf{1}\,(t_i \neq 1)\,\mathbf{1}\,(s_i + t_i \neq 0)\,(N^{-1/2})^{\max\{s_i,t_i\}+2}\right)(\mathbb{E}|F_0|)\,N^\delta
$$

$$
\leqslant (\mathbb{E}|F_0|)\,N^{-1/2+\delta} \tag{4.12}
$$

Combining (4.11) and (4.12), we obtain

$$
\mathbb{E}\mathbf{1}_\Omega F_0 F \prec (\mathbb{E}|F_0|)\,N^{-1/2}, \tag{4.13}
$$

Then together with (4.10), we obtain (4.5) and complete the proof of Lemma 4.5. $\qquad\square$

Now we slightly extend the above lemma. Instead of assuming $F \in_n \mathcal{F}_{1/2} \cdot \mathcal{F}$, we assume that $F = F \in_n \mathcal{F}_{1/2} \cdot \mathcal{F} + O_\prec(N^{-D})$ for some fixed $D > 0$.

Corollary 4.6. *For fixed indeces a, b and ensemble X in lemma 3.2, let F_0 and F be two random variables bounded by N^C for some C, i.e.,*

$$
|F_0| + |F| \leqslant N^C
$$

We assume that

$$
F_0 \in N^C \mathcal{F}_0
$$

and for some fixed $D > 0$,

$$
F =\in_n \mathcal{F}_{1/2} \cdot \mathcal{F} + O_\prec(N^{-D})
$$

Then we have

$$
|\mathbb{E}\,F_0 F| \prec N^{-1/2}\mathbb{E}\,|F_0| + N^{-D+2C+1} \tag{4.14}
$$

Proof of Corollary 4.6: Write

$$
F = F^M + F^e, \quad F^M \in_n \mathcal{F}_{1/2} \cdot \mathcal{F}, \quad F^e = O_\prec(N^{-D})
$$

Here superscription M and e are for *main* and *error*. (Note F^M and F^e are not assumed to be bounded by N^C, otherwise the proof is much simpler.) For simplicity, we assume $F^M \in \mathcal{F}_{1/2} \cdot \mathcal{F}$ (not $\in_n$) and for some $m \geqslant 0$, $F^M \in F_{1/2}(F_1)^m$. Then we repeat the same argument as above, i.e., from (4.6) to (4.9). Then for any (small) $\delta > 0$ and (large) $\widetilde{D} > 0$, there exists probability set Ω, which is independent of the a-th row and column of X, such that $\mathbb{P}(\Omega) \geqslant 1 - N^{-\widetilde{D}}$, and (4.9) holds. Next we write

$$
|\mathbb{E}\,F_0 F| = |\mathbb{E}\,\mathbf{1}_{\Omega^c} F_0 F| + \left|\mathbb{E}\,\mathbf{1}_\Omega F_0 F^M\right| + \left|\mathbb{E}\,\mathbf{1}_\Omega F_0 F^e\right|
$$

$$
= N^{-\widetilde{D}+2C} + N^{-1/2+\delta}\mathbb{E}\,|F_0| + \left|\mathbb{E}\,\mathbf{1}_\Omega F_0 F^e\right|
$$

where we used $|F_0| + |F| \leqslant N^C$ and (4.13).

Now we bound $|\mathbb{E}\,\mathbf{1}_\Omega F_0 F^e|$. By the definition of $\prec$ again, there exists $\widetilde{\Omega}$ such that $\mathbb{P}(\widetilde{\Omega}) \geqslant 1 - N^{-\widetilde{D}}$ and

$$
F^e \leqslant N^{-D+\delta}
$$

With this $\widetilde{\Omega}$, and $|F_0| + |F| \leqslant N^C$ we write

$$
\begin{aligned}
\left|\mathbb{E}\,1_\Omega F_0 F^e\right| &\leqslant \left|\mathbb{E}\,1_{\Omega\cap\widetilde{\Omega}} F_0 F^e\right| + \left|\mathbb{E}\,1_{\Omega\cap\widetilde{\Omega}^c} F_0 F^e\right| && (4.15) \\
&= \left|\mathbb{E}\,1_{\Omega\cap\widetilde{\Omega}} F_0 F^e\right| + \left|\mathbb{E}\,1_{\Omega\cap\widetilde{\Omega}^c} F_0 F\right| + \left|\mathbb{E}\,1_{\Omega\cap\widetilde{\Omega}^c} F_0 F^M\right| \\
&\leqslant N^{-D+C+\delta} + N^{-\widetilde{D}+2C} + \left|\mathbb{E}\,1_{\Omega\cap\widetilde{\Omega}^c} F_0 F^M\right|
\end{aligned}
$$

For the last term, we note that by the definition of Ω we can simply bound the term in $\mathcal{F}$ which are independent of the a-th row and column of X by $N^{0.1}$. Then using the assumption $F^M \in \mathcal{F}_{1/2}(\mathcal{F}_1)^m$, we have

$$
|1_\Omega F^M| \leqslant N^{4m+1} \sum_{n=1}^{2m+1} \left(\sum_{k_1,k_2,\cdots,k_n} \prod_{j=1}^{2m+1} \left(|X_{ak_j}| + |X_{ak_j}| \right) \right)
$$

Together with Cauchy-Schwarz inequality, and subexponential decay property (1.7), we obtain that

$$
\left|\mathbb{E}\,1_{\Omega\cap\widetilde{\Omega}^c} F_0 F^e\right| \leqslant \mathbb{E}\,1_\Omega |F_0 F^e|^2\, \mathbb{P}(\widetilde{\Omega}^c) \leqslant N^{-\widetilde{D}+C_m}
$$

Inserting it into (4.15), choosing large enough $\widetilde{D}$, we obtain (4.14) and complete the proof. $\qquad\square$

More general, if $F_T \in_n \mathcal{F}_{1/2} \cdot \mathcal{F}$ hold uniformly for $T \in \mathcal{T}$, corollary 4.6 can be extended to the following integration version.

Lemma 4.7. *For fixed indeces a, b and ensemble X in lemma 3.2, let F_T be a family of random variables such that for some deterministic x_T and uniform $D > 0$*

$$
F_T \in_n x_T \mathcal{F}_{1/2} \cdot \mathcal{F} + O_\prec(N^{-D})
$$

hold uniformly for $T \in \mathcal{T} = \mathcal{T}_N$, i.e., $F_T = F_T^M + F_T^e$ and

$$
F_T^M \in_n x_T \mathcal{F}_{1/2} \cdot \mathcal{F}, \quad F_T^e = O_\prec(N^{-D})
$$

hold uniformly for $T \in \mathcal{T} = \mathcal{T}_N$. Here we assume that $\cup_N \mathcal{T}_N$ can be covered by a compact set in $\mathbb{R}^p$ for some $p \in \mathbb{N}$, this compact set and p are independent of N.

We also assume that $|x_T| + |F_T| \leqslant N^C$ for some uniform $C > 0$. Let F_0 be a random variable satisfying $F_0 \prec N^C \mathcal{F}$ and $|F_0| \leqslant N^C$. Then

$$
\left| \mathbb{E}\, F_0 \int_{T\in\mathcal{T}} F_T \mathrm{d}T \right| \prec N^{-1/2}\left(\mathbb{E}\,|F_0|\right) \int_{T\in\mathcal{T}} |x_T| \mathrm{d}T + N^{-D+2C+1} \qquad (4.16)
$$

Proof of Lemma 4.7: Since F_0 and F_T are bounded by N^C, one can exchange the order of integration and expectation, i.e.,

$$
\mathbb{E}\, F_0 \int_{T\in\mathcal{T}} F_T \mathrm{d}T = \int_{T\in\mathcal{T}} \left(\mathbb{E}\, F_0 \cdot F_T\right) \mathrm{d}T
$$

Then with the uniformness, one can easily extend the proof of Lemma 4.5 and corollary 4.6, and prove this lemma.

$\qquad\square$

4.2 Proof of Lemma 3.2. The Lem. 4.5 and 4.7 are the key observations for the proof of Lemma 3.2. Now to prove Lemma 3.2, we claim that the following lemma 4.9, which shows that the terms in Lemma 3.2 can be represented by $\mathcal{F}$ and $\mathcal{F}_{1/2} \cdot \mathcal{F}$(with negligible error term). We first introduce a cutoff function on $\operatorname{Re} m^{(a,a)}$. (Recall the definition in Def. 4.1)

Definition 4.8. *Define* χ_a *as*

$$\chi_a := \chi_a(\varepsilon, w, z) = \mathbf{1}\left(|\operatorname{Re} m^{(a,a)}| \geqslant \frac{1}{2} N^\varepsilon (N\eta)^{-1}\right) \tag{4.17}$$

Note: By definition and (4.3), $h(t_X) > 0$ implies $\chi_a = 1$, and for any $|\mathbb{U}| + |\mathbb{T}| = O(1)$, we have

$$h(t_X) > 0 \quad \Longrightarrow \quad \chi_a = 1 \quad \Longrightarrow \quad |\operatorname{Re} m^{(\mathbb{U}, \mathbb{T})}| \geqslant \frac{1}{4} N^\varepsilon (N\eta)^{-1} \tag{4.18}$$

Lemma 4.9. *Recall* $X^{(a,a)}$ *and* $m^{(a,a)}$ *defined in Definition* (4.1). *Under the assumption of Lemma 3.2, for any fixed large* $D > 0$, *we have*

$$h(t_X) \operatorname{Re} m - h(t_{X^{(a,a)}}) \operatorname{Re} m^{(a,a)} \in_n \frac{1}{N\eta} \mathcal{F} + O_\prec(N^{-D}) \tag{4.19}$$

$$B_m(X) \in_n (N\eta)^{m-1} \mathcal{F} + O_\prec(N^{-D}), \quad m = 1, 2, 3 \tag{4.20}$$

$$\chi_a P_m(X) \in_n \frac{1}{N\eta} \mathcal{F}_{1/2} \cdot \mathcal{F} + O_\prec(N^{-D}), \quad m = 1, 3 \tag{4.21}$$

$$\chi_a P_2(X) \in_n \frac{1}{N\eta} \mathcal{F} + O_\prec(N^{-D})$$

uniformly hold for

$$a, b: \ 1 \leqslant a \neq b \leqslant N, \quad z: ||z| - 1| \leqslant 2\varepsilon, \quad \text{and} \quad w \in I_\varepsilon.$$

We postpone the proof of this lemma to the next section. In the remainder of this section, we will prove Lemma 3.2 with Lemma 4.9. First we introduce a simple lemma for the calculation of $\mathcal{F}$ sets.

Lemma 4.10. *Let* A *and* B *be two variables stochastically dominated by* N^C *for some* $C > 0$, *i.e.,* $|A| + |B| \prec N^C$. *If for random variable* A_0 *and* B_0, *we have*

$$A = A_0 + O_\prec(N^{-D}), \quad B = B_0 + O_\prec(N^{-D}),$$

for some $D > 0$. *Then*

$$AB = A_0 B_0 + O_\prec(N^{C-D}) \tag{4.22}$$

Proof: By assumption,

$$\left(A - O_\prec(N^{-D})\right)\left(B - O_\prec(N^{-D})\right) = A_0 B_0$$

With $|A| + |B| \prec N^C$, we obtain (4.22).

$\square$

Now we return to finish the proof of Lemma 3.2.

Proof of Lemma 3.2: For simplicity, we introduce the notation $\widetilde{A}(w, z)$ as

$$\widetilde{A}(w, z) := h(t_X) \operatorname{Re} m(w, z) - h(t_{X^{(a,a)}}) \operatorname{Re} m^{(a,a)}(w, z)$$

First as in (3.30), (3.18) and (3.23), one can see that there exists uniform $C > 0$, such that

$$|A_X^{(f)}| + |A_{X^{(a,a)}}^{(f)}| + |\widetilde{A}(w,z)| + \sum_{n=1,2,3} |\mathscr{P}_n(X)| + \sum_{n=1,2,3} |P_n(w,z)| + \sum_{n=1,2,3} |B_n(w,z)| \leqslant N^C \qquad (4.23)$$

With $A_X^{(f)} = A_{X^{(a,a)}}^{(f)} + (A_X^{(f)} - A_{X^{(a,a)}}^{(f)})$, we write

$$(A_X^{(f)})^{p-3} \mathscr{P}_1^3(X) = \sum_l C_l (A_{X^{(a,a)}}^{(f)})^{p-3-l} \left(A_X^{(f)} - A_{X^{(a,a)}}^{(f)} \right)^l \mathscr{P}_1^3(X)$$

Recall the definitions in (2.7), (3.28) and (3.17), for fixed l, with the notation $\widetilde{A}(w,z)$ and (4.18), we can write:

$$(A_{X^{(a,a)}}^{(f)})^{p-3-l} \left(A_X^{(f)} - A_{X^{(a,a)}}^{(f)} \right)^l \mathscr{P}_1^3(X) \qquad (4.24)$$

$$= (A_{X^{(a,a)}}^{(f)})^{p-3-l} N^{l+3} \int_{\mathcal{T}} \prod_{i=1}^l \widetilde{A}(w_i, z_i) \prod_{i=l+1}^{l+3} (\chi_a P_1 B_1)(w_i, z_i) \prod_{i=1}^{l+3} \Delta f(\xi_i) \chi(\eta_i) \phi'(E_i) \mathrm{d}T$$

where $\mathrm{d}T = \prod_i \mathrm{d}E_i \mathrm{d}\eta_i \mathrm{d}A(\xi_i)$ and $\mathcal{T} = (I_\varepsilon \times \mathrm{supp}\, f)^{l+3}$. Using (4.23), Lemma 4.9 and Lemma 4.10, for any fixed $D > 0$, we have

$$\prod_{i=1}^l \widetilde{A}(w_i, z_i) \prod_{i=l+1}^{l+3} (\chi_a P_1 B_1)(w_i, z_i) \in_n \left(\prod_{i=1}^{l+3} \frac{1}{N\eta_i} \right) \mathcal{F}_{1/2} \cdot \mathcal{F} + O_\prec(N^{-D}), \quad \eta_i = \mathrm{Im}\, w_i \qquad (4.25)$$

uniformly hold for $T \in \mathcal{T}$. Applying Lemma 4.7 by choosing

$$F_0 = (A_{X^{(a,a)}}^{(f)})^{p-3-l}, \quad F_T = \prod_{i=1}^l \widetilde{A}(w_i, z_i) \prod_{i=l+1}^{l+3} (\chi_a P_1 B_1)(w_i, z_i), \quad x_T = \prod_{i=1}^{l+3} \eta_i^{-1} \Delta f(\xi_i) \chi(\eta_i) \phi'(E_i)$$

and $\mathcal{T} = (I_\varepsilon \times \mathrm{supp}\, f)^{l+3}$, with (4.24) and (4.24), we obtain

$$\mathbb{E}(A_{X^{(a,a)}}^{(f)})^{p-3-l} \left(A_X^{(f)} - A_{X^{(a,a)}}^{(f)} \right)^l \mathscr{P}_1^3(X) \prec N^{-1/2} \left(\mathbb{E}\left(|A_{X^{(a,a)}}^{(f)}|^{p-l-3} \right) \right) + N^{-D}$$

for any fixed $D > 0$. Then use Holder inequality, we have

$$\left| \mathbb{E}(A_{X^{(a,a)}}^{(f)})^{p-3-l} \left(A_X^{(f)} - A_{X^{(a,a)}}^{(f)} \right)^l \mathscr{P}_1^3(X) \right| \prec N^{-1/2} \left(O_\prec(1) + \mathbb{E}\left(|A_{X^{(a,a)}}^{(f)}|^p \right) \right)$$

Similarly, one can prove

$$\left| \mathbb{E}(A_X^{(f)})^{p-3} \mathscr{P}_1^3(X) \right| + \left| \mathbb{E}(A_X^{(f)})^{p-2} \mathscr{P}_1(X) \mathscr{P}_2(X) \right| + \left| \mathbb{E}(A_X^{(f)})^{p-1} \mathscr{P}_3(X) \right| \prec N^{-1/2} \left(O_\prec(1) + \mathbb{E}\left((A_{X^{(a,a)}}^{(f)})^p \right) \right) \qquad (4.26)$$

It follows from (4.3) that $m - m^{(a,a)} = O(N\eta)^{-1}$. Then it is easy to check that $|A_{X^{(a,a)}}^{(f)} - A_X^{(f)}| \leqslant C$. Inserting it into (4.26), we complete the proof of Lemma 3.2.

$\square$

5   Polynomialization of Green's functions

As showed in the previous sections, to complete the proof of Theorem 1.2, it only remains to prove Lemma 4.9. In this section, we will prove Lemma 4.9, i.e., write the terms in (4.19) as polynomials in $\mathcal{F}$ or $\mathcal{F}_{1/2} \cdot \mathcal{F}$ (up to negligible error). Since the uniformness can be easily checked, we will only focus on the fixed a, b, z, w:

$$a, b:\ 1 \leqslant a \neq b \leqslant N, \quad z: ||z| - 1| \leqslant 2\varepsilon, \quad \text{and} \quad w \in I_\varepsilon.$$

First we need to write the single matrix elements of G's and $\mathcal{G}'s$ as this type of polynomials. To do so, we start with deriving some bounds on G's under the condition:

$$|\operatorname{Re} m| \geqslant \frac{1}{4} N^\varepsilon (N\eta)^{-1} \tag{5.1}$$

Note: this condition is guaranteed by $\chi_a > 0$, $h(t_X) > 0$ or $h(t_{X^{(a,a)}}) > 0$.

5.1   Preliminary lemmas. This subsection summarizes some elementary results from [9] and [10]. Note that all the inequalities in this subsection hold uniformly for bounded z and w. Furthermore, they hold without the condition (5.1).

Recall the definitions of $Y^{(U,T)}$, $G^{(U,T)}$, $\mathcal{G}^{(U,T)}$, $\mathbf{y}_i$ and y_i in the definition 4.1.

Lemma 5.1 (Relation between G, $G^{(\mathbb{T},\emptyset)}$ and $G^{(\emptyset,\mathbb{T})}$). *For $i, j \neq k$ ($i = j$ is allowed) we have*

$$G_{ij}^{(k,\emptyset)} = G_{ij} - \frac{G_{ik}G_{kj}}{G_{kk}}, \quad \mathcal{G}_{ij}^{(\emptyset,k)} = \mathcal{G}_{ij} - \frac{\mathcal{G}_{ik}\mathcal{G}_{kj}}{\mathcal{G}_{kk}}, \tag{5.2}$$

$$G^{(\emptyset,i)} = G + \frac{(G\mathrm{y}_i^*)(\mathrm{y}_iG)}{1 - \mathrm{y}_iG\mathrm{y}_i^*}, \quad G = G^{(\emptyset,i)} - \frac{(G^{(\emptyset,i)}\mathrm{y}_i^*)(\mathrm{y}_iG^{(\emptyset,i)})}{1 + \mathrm{y}_iG^{(\emptyset,i)}\mathrm{y}_i^*}, \tag{5.3}$$

and

$$\mathcal{G}^{(i,\emptyset)} = \mathcal{G} + \frac{(\mathcal{G}\mathbf{y}_i)(\mathbf{y}_i^*\mathcal{G})}{1 - \mathbf{y}_i^*\mathcal{G}\mathbf{y}_i}, \quad \mathcal{G} = \mathcal{G}^{(i,\emptyset)} - \frac{(\mathcal{G}^{(i,\emptyset)}\mathbf{y}_i)(\mathbf{y}_i^*\mathcal{G}^{(i,\emptyset)})}{1 + \mathbf{y}_i^*\mathcal{G}^{(i,\emptyset)}\mathbf{y}_i}.$$

Definition 5.2. *In the following, $\mathbb{E}_X$ means the integration with respect to the random variable X. For any $\mathbb{T} \subset [\![1, N]\!]$, we introduce the notations*

$$Z_i^{(\mathbb{T})} := (1 - \mathbb{E}_{\mathrm{y}_i})\mathrm{y}_i^{(\mathbb{T})}G^{(\mathbb{T},i)}\mathrm{y}_i^{(\mathbb{T})*}$$

and

$$\mathcal{Z}_i^{(\mathbb{T})} := (1 - \mathbb{E}_{\mathbf{y}_i})\mathbf{y}_i^{(\mathbb{T})*}\mathcal{G}^{(i,\mathbb{T})}\mathbf{y}_i^{(\mathbb{T})}.$$

Recall by our convention that $\mathbf{y}_i$ is a $N \times 1$ column vector and y_i is a $1 \times N$ row vector. For simplicity we will write

$$Z_i = Z_i^{(\emptyset)}, \quad \mathcal{Z}_i = \mathcal{Z}_i^{(\emptyset)}.$$

Lemma 5.3 (Identities for G, $\mathcal{G}$, Z and $\mathcal{Z}$). *For any $\mathbb{T} \subset [\![1, N]\!]$, we have*

$$G_{ii}^{(\emptyset,\mathbb{T})} = -w^{-1}\left[1 + m_{\mathcal{G}}^{(i,\mathbb{T})} + |z|^2\mathcal{G}_{ii}^{(i,\mathbb{T})} + \mathcal{Z}_i^{(\mathbb{T})}\right]^{-1}, \tag{5.4}$$

$$G_{ij}^{(\emptyset,\mathbb{T})} = -wG_{ii}^{(\emptyset,\mathbb{T})}G_{jj}^{(i,\mathbb{T})}\left(\mathbf{y}_i^{(\mathbb{T})*}\mathcal{G}^{(ij,\mathbb{T})}\mathbf{y}_j^{(\mathbb{T})}\right), \quad i \neq j, \tag{5.5}$$

24

where, by definition, $\mathcal{G}_{ii}^{(i,\mathbb{T})} = 0$ if $i \in \mathbb{T}$. Similar results hold for $\mathcal{G}$:

$$\left[\mathcal{G}_{ii}^{(\mathbb{T},\emptyset)}\right]^{-1} = -w\left[1 + m_G^{(\mathbb{T},i)} + |z|^2 G_{ii}^{(\mathbb{T},i)} + Z_i^{(\mathbb{T})}\right] \tag{5.6}$$

$$\mathcal{G}_{ij}^{(\mathbb{T},\emptyset)} = -w\mathcal{G}_{ii}^{(\mathbb{T},\emptyset)}\mathcal{G}_{jj}^{(\mathbb{T},i)}\left(y_i^{(\mathbb{T})} G^{(\mathbb{T},ij)} y_j^{(\mathbb{T})*}\right), \quad i \neq j. \tag{5.7}$$

Definition 5.4 (ζ-High probability events)**.** *Define*

$$\varphi := (\log N)^{\log\log N}. \tag{5.8}$$

Let $\zeta > 0$. We say that an N-dependent event Ω holds with ζ-high probability if there is some constant C such that

$$\mathbb{P}(\Omega^c) \leqslant N^C \exp(-\varphi^\zeta)$$

for large enough N. Furthermore, we say that $\Omega(u)$ holds with ζ-high probability uniformly for $u \in U_N$, if there is some uniform constant C such that

$$\max_{u \in U_N} \mathbb{P}(\Omega^c(u)) \leqslant N^C \exp(-\varphi^\zeta) \tag{5.9}$$

for uniformly large enough N.

Note: Usually we choose ζ to be 1. By the definition, if some event Ω holds with ζ-high probability for some $\zeta > 0$, then Ω holds with probability larger then $1 - N^{-D}$ for any $D > 0$.

Lemma 5.5 (Large deviation estimate)**.** *Let X be defined as in Theorem 1.2. For any $\zeta > 0$, there exists $Q_\zeta > 0$ such that for $\mathbb{T} \subset [\![1,N]\!]$, $|\mathbb{T}| \leqslant N/2$ the following estimates hold with ζ-high probability uniformly for $1 \leqslant i, j \leqslant N$, $|w| + |z| \leqslant C$:*

$$|Z_i^{(\mathbb{T})}| = \left|(1 - \mathbb{E}_{y_i})\left(y_i^{(\mathbb{T})} G^{(\mathbb{T},i)} y_i^{(\mathbb{T})*}\right)\right| \leqslant \varphi^{Q_\zeta/2}\sqrt{\frac{\operatorname{Im} m_G^{(\mathbb{T},i)} + |z|^2 \operatorname{Im} G_{ii}^{(\mathbb{T},i)}}{N\eta}}, \tag{5.10}$$

$$|\mathcal{Z}_i^{(\mathbb{T})}| = \left|(1 - \mathbb{E}_{\mathbf{y}_i})\left(\mathbf{y}_i^{(\mathbb{T})*} \mathcal{G}^{(i,\mathbb{T})} \mathbf{y}_i^{(\mathbb{T})}\right)\right| \leqslant \varphi^{Q_\zeta/2}\sqrt{\frac{\operatorname{Im} m_{\mathcal{G}}^{(i,\mathbb{T})} + |z|^2 \operatorname{Im} \mathcal{G}_{ii}^{(i,\mathbb{T})}}{N\eta}}.$$

Furthermore, for $i \neq j$, we have

$$\left|(1 - \mathbb{E}_{y_i y_j})\left(y_i^{(\mathbb{T})} G^{(\mathbb{T},ij)} y_j^{(\mathbb{T})*}\right)\right| \leqslant \varphi^{Q_\zeta/2}\sqrt{\frac{\operatorname{Im} m_G^{(\mathbb{T},ij)} + |z|^2 \operatorname{Im} G_{ii}^{(\mathbb{T},ij)} + |z|^2 \operatorname{Im} G_{jj}^{(\mathbb{T},ij)}}{N\eta}}, \tag{5.11}$$

$$\left|(1 - \mathbb{E}_{\mathbf{y}_i\mathbf{y}_j})\left(\mathbf{y}_i^{(\mathbb{T})*} \mathcal{G}^{(ij,\mathbb{T})} \mathbf{y}_j^{(\mathbb{T})}\right)\right| \leqslant \varphi^{Q_\zeta/2}\sqrt{\frac{\operatorname{Im} m_{\mathcal{G}}^{(ij,\mathbb{T})} + |z|^2 \operatorname{Im} \mathcal{G}_{ii}^{(ij,\mathbb{T})} + |z|^2 \operatorname{Im} \mathcal{G}_{jj}^{(ij,\mathbb{T})}}{N\eta}}, \tag{5.12}$$

where

$$\mathbb{E}_{y_i y_j}\left(y_i^{(\mathbb{T})} G^{(\mathbb{T},ij)} y_j^{(\mathbb{T})*}\right) = |z|^2 G_{ij}^{(\mathbb{T},ij)} + \delta_{ij} m_G^{(\mathbb{T},ij)}, \quad \mathbb{E}_{\mathbf{y}_i\mathbf{y}_j}\left(\mathbf{y}_i^{(\mathbb{T})*} \mathcal{G}^{(ij,\mathbb{T})} \mathbf{y}_j^{(\mathbb{T})}\right) = |z|^2 \mathcal{G}_{ij}^{(ij,\mathbb{T})} + \delta_{ij} m_{\mathcal{G}}^{(ij,\mathbb{T})}. \tag{5.13}$$

Lemma 5.6. *Let X be defined as in Theorem 1.2. Suppose $|w| + |z| \leqslant C$. For any $\zeta > 0$, there exists C_ζ such that if the assumption*

$$\eta \geqslant \varphi^{C_\zeta} N^{-1} |w|^{1/2} \tag{5.14}$$

holds then the following estimates hold

$$\max_i |G_{ii}| \leqslant 2(\log N)|w|^{-1/2}, \tag{5.15}$$

$$\max_i |w||G_{ii}||\mathcal{G}_{ii}^{(i,\emptyset)}| \leqslant (\log N)^4, \tag{5.16}$$

$$\max_{ij} |G_{ij}| \leqslant C(\log N)^2 |w|^{-1/2}, \tag{5.17}$$

$$|m| \leqslant 2(\log N)|w^{-1/2}| \tag{5.18}$$

with ζ-high probability uniformly for $|w| + |z| \leqslant C$.

5.2 Improved bounds on G's.

The next lemma gives the bounds on G, $\mathcal{G}$ and m under the condition (5.1). Note: with (4.3), it implies that for any U, T: $|U| + |T| = O(1)$,

$$|\operatorname{Re} m^{(U,T)}| \gg (N\eta)^{-1}. \tag{5.19}$$

Before we give the rigorous proof for the bounds on G, $\mathcal{G}$, we provide a rough picture on the sizes of these terms under the condition (5.1), $w \in I_\varepsilon$ and $||z| - 1| \leqslant 2\varepsilon$. We note that the typical size of the $G_{kl}^{(U,T)}$ heavily relies on whether $k = l$ and whether k, l are in U, T.

(i) If $k = l \notin U \cup T$, the typical size of $G_{kk}^{(U,T)}(w,z)$ is $m(w,z) = \frac{1}{N} \operatorname{Tr} G(w,z)$.

(ii) If $k \neq l$, and $k, l \notin U \cup T$, the typical size of $G_{kl}^{(U,T)}(w,z)$ is $\sqrt{|m|/(N\eta)}$.

(iii) If $\{k,l\} \cap U \neq \emptyset$, then $G_{kl}^{(U,T)} = 0$. This result follows from the definition, and it worth to emphasize:

$$\{k,l\} \cap U \neq \emptyset \implies G_{kl}^{(U,T)} = \mathcal{G}_{kl}^{(T,U)} = 0 \tag{5.20}$$

(iv) If $k = l \in T$, then the typical size of $G_{kk}^{(U,T)}$ is $|wm|^{-1}$

(v) If $k \neq l$, and $k \in T$ and $l \notin T$, then the typical size of $G_{kl}^{(U,T)}$ is $(|w^{1/2}m|)^{-1}\sqrt{|m|/(N\eta)}$

(vi) If $k \neq l$, and $k, l \in T$ then the typical size of $G_{kl}^{(U,T)}$ is $|wm^2|^{-1}\sqrt{|m|/(N\eta)}$

(vii) With the definition of $G^{(U,T)}$ and $\mathcal{G}^{(T,U)}$ in Def. 4.1, one can easily see that $\mathcal{G}_{kl}^{(T,U)}$ has the same typical size as $G_{kl}^{(U,T)}$ (Here the superscript of $\mathcal{G}$ is (T,U) not (U,T)).

We note: The m is bounded by $(\log N)^C |w|^{-1/2}$ in (5.18) (no better bound is obtained in this paper), but we believe that it could be much smaller.

Lemma 5.7. *Let X be defined as in Theorem 1.2. Let ε be small enough positive number, $||z^2| - 1| \leqslant 2\varepsilon$ and $w \in I_\varepsilon$ (see definition in (2.5)). If (5.1) holds , i.e., $|\operatorname{Re} m(w, z)| \geqslant \frac{1}{4} N^\varepsilon (N\eta)^{-1}$ in $\Omega = \Omega(\varepsilon, w, z)$. Then there exists $\widetilde{\Omega} \subset \Omega$, and $C > 0$ such that $\widetilde{\Omega}$ holds in Ω with 1-high probability uniformly for z, w: $||z^2| - 1| \leqslant 2\varepsilon$ and $w \in I_\varepsilon$, (see definition in (5.9)) and the following bounds hold in $\widetilde{\Omega}$ for any $1 \leqslant i \neq j \leqslant N$, (Here $A \sim B$ denotes there exists $C > 0$ such that $C^{-1}|B| \leqslant |A| \leqslant C|B|$)*

$$|1 + m| \geqslant N^{\frac{3}{4}\varepsilon}(N\eta)^{-1} \tag{5.21}$$

$$|1 + m^{(i,i)}| \geqslant N^{\frac{1}{4}\varepsilon}\left|\mathcal{Z}_i^{(i)}\right| \tag{5.22}$$

$$G_{ii}^{(\emptyset,i)} = (1 + O(N^{-\frac{1}{4}\varepsilon}))\frac{-1}{w}\frac{1}{1 + m^{(i,i)}} \tag{5.23}$$

$$|1 + m| \sim |m| \tag{5.24}$$

$$|G_{ii}| \leqslant (\log N)^C |m| \tag{5.25}$$

$$|G_{ij}^{(\emptyset,i)}| \leqslant \frac{\varphi^C}{|w^{1/2}m|}\sqrt{\frac{|m|}{N\eta}} \tag{5.26}$$

$$|G_{ii}^{(\emptyset,j)}| \leqslant (\log N)^C |m| \tag{5.27}$$

$$|G_{ii}^{(\emptyset,ij)}| \leqslant \frac{C}{|wm|} \tag{5.28}$$

$$|G_{ij}| \leqslant \varphi^C \sqrt{\frac{|m|}{N\eta}} \tag{5.29}$$

$$|wG_{ii}|^{-1} \geqslant N^{\frac{1}{2}\varepsilon}|\mathcal{Z}_i| \tag{5.30}$$

$$|m^{(i,i)}| \geqslant (\log N)^{-1} \tag{5.31}$$

Furthermore, with the symmetry and the definition of $G^{(\mathbb{U},\mathbb{T})}$ and $\mathcal{G}^{(\mathbb{T},\mathbb{U})}$, these bounds also hold under the following exchange

$$G^{(\mathbb{U},\mathbb{T})} \leftrightarrow \mathcal{G}^{(\mathbb{T},\mathbb{U})}, \quad \mathcal{Z} \leftrightarrow Z. \tag{5.32}$$

Proof of Lemma 5.7: In the following proof, we only focus on the fixed z, w, i and j, since the uniformness can be easily checked.

We choose $\zeta = 1$. Because $\varphi \ll N^\varepsilon$ for any fixed $\varepsilon > 0$ (see (5.8)) and in this lemma $w \in I_\varepsilon$, one can easily check that the assumption in this lemma implies the conditions of lemma 5.6 i.e.,

$$w \in I_\varepsilon \implies (5.14) \text{ holds for } \forall C_\zeta \tag{5.33}$$

Therefore we can use all of the results (with $\zeta = 1$) of lemma 5.6 in the following proof.

1. We first prove (5.21). The condition (5.1) implies that $|\frac{1}{N}\sum_i \operatorname{Re} G_{ii}| \geqslant \frac{1}{4}N^\varepsilon(N\eta)^{-1}$, then there exists $i : 1 \leqslant i \leqslant N$ such that $|G_{ii}| \geqslant \frac{1}{4}N^\varepsilon(N\eta)^{-1}$. Together with (5.16), it implies that $|\mathcal{G}_{ii}^{(i,\emptyset)}| \leqslant |w|^{-1}N^{-\frac{4}{5}\varepsilon}N\eta$ with 1 - high probability in Ω. Inserting it into (5.6) with $\mathbb{T} = i$, using $G_{ii}^{(i,i)} = 0$ from (5.20), we have

$$|1 + m^{(i,i)} + Z_i^{(i)}| \geqslant N^{\frac{4}{5}\varepsilon}(N\eta)^{-1} \tag{5.34}$$

Applying (5.10) to bound $Z_i^{(i)}$ with $\mathbb{T} = i$, using Schwarz's inequality and the fact $G_{ii}^{(i,i)} = 0$ again, we obtain

$$|Z_i^{(i)}| \leqslant N^{-\varepsilon/20}\operatorname{Im} m^{(i,i)} + N^{\varepsilon/10}(N\eta)^{-1} \tag{5.35}$$

holds with 1-high probability in Ω. Together with (5.34), it implies that with 1-high probability in Ω,

$$|1 + m^{(i,i)}| \geqslant 2N^{\frac{3}{4}\varepsilon}(N\eta)^{-1}$$

Then replacing $m^{(i,i)}$ with m by (4.3), we obtain (5.21).

2. For (5.22), first using (4.3) and (5.21), we have that for *any* $i : 1 \leqslant i \leqslant N$

$$|1 + m^{(i,i)}| \geqslant N^{\frac{2}{3}\varepsilon}(N\eta)^{-1} \tag{5.36}$$

holds with 1-high probability in Ω. Together with the $\mathcal{Z}$ version of (5.35):

$$|\mathcal{Z}_i^{(i)}| \leqslant N^{-\varepsilon/4} \operatorname{Im} m^{(i,i)} + N^{\varepsilon/3}(N\eta)^{-1}$$

we obtain (5.22).

3. For (5.23), it follows from (5.4) with $\mathbb{T} = i$, (5.20) and (5.22).

4. Now we prove (5.24). Suppose (5.21), (5.23) and (5.10) holds in $\Omega_0 \subset \Omega$. From our previous results, Ω_0 holds with 1-high probability in Ω. Now we prove that (5.24) holds in Ω_0. First we assume that $|1+m| \leqslant 3$, clearly otherwise (5.24) holds. Together with (5.21), it implies that $(N\eta)^{-1} \leqslant 3N^{-\frac{1}{2}\varepsilon}$. Using (4.3) and $|1 + m| \leqslant 3$, we obtain $|1 + m^{(i,i)}| \leqslant 4$ and $|m_G^{(\emptyset,i)}| \leqslant 5$. With (5.23), the bound $|1 + m^{(i,i)}| \leqslant 4$ implies $|G_{ii}^{(\emptyset,i)}| \geqslant |5w|^{-1}$. The assumption $w \in I_\varepsilon$ implies $|w| \leqslant \varepsilon$ (see definition of I_ε in (2.5)). Then applying (5.10) on Z_i, and using $||z| - 1| \leqslant 2\varepsilon$ and the bounds we just proved on $(N\eta)^{-1}$, $m_G^{(\emptyset,i)}$ and $G_{ii}^{(\emptyset,i)}$, we obtain that in Ω_0,

$$|Z_i| \leqslant N^{\frac{-1}{3}\varepsilon}|G_{ii}^{(\emptyset,i)}| \tag{5.37}$$

Together with $|G_{ii}^{(\emptyset,i)}| \geqslant |5w|^{-1}$ and the assumption $||z| - 1| \leqslant 2\varepsilon$ and $|w| \leqslant \varepsilon$, we have

$$\left| |z|^2 |G_{ii}^{(\emptyset,i)}| + Z_i \right| \geqslant |10w|^{-1} \tag{5.38}$$

Now inserting (5.38) into the identity (5.6) with $\mathbb{T} = \emptyset$, using $|m_G^{(\emptyset,i)}| \leqslant 5$, and $|w| \leqslant \varepsilon$ again, we obtain that

$$\mathcal{G}_{ii} = \frac{1}{-w\left(|z|^2 G_{ii}^{(\emptyset,i)} + Z_i\right)} + \varepsilon_i, \quad |\varepsilon_i| \leqslant |60w|\frac{1}{|w|\left(|z|^2 G_{ii}^{(\emptyset,i)} + Z_i\right)} \tag{5.39}$$

Then together with (5.37) and (5.23), in Ω_0, we have

$$\left| \mathcal{G}_{ii} - |z|^{-2}(1 + m^{(i,i)}) \right| \leqslant (O(|w|) + o(1))\left|(1 + m^{(i,i)})\right| \tag{5.40}$$

Combining (5.21) and (4.3), we have

$$(1 + m^{(i,i)}) = (1 + o(1))(1 + m)$$

Inserting it into (5.40), we have

$$\left| \mathcal{G}_{ii} - |z|^{-2}(1 + m) \right| \leqslant (O(|w|) + o(1))\left|(1 + m)\right|, \quad \text{in } \Omega_0 \tag{5.41}$$

It is easy to extend this result to the following one:

$$\max_i \left| \mathcal{G}_{ii} - |z|^{-2}(1+m) \right| \leqslant \left(O(|w|) + o(1) \right)|(1+m)|, \quad \text{in } \widetilde{\Omega} \tag{5.42}$$

holds in a probability set $\widetilde{\Omega} \subset \Omega$ such that $\widetilde{\Omega}$ holds with 1-high probability in Ω. Since $m = \frac{1}{N}\sum_i \mathcal{G}_{ii}$, for small enough ε, with $|w| \leqslant \varepsilon$ and $||z^2| - 1| \leqslant 2\varepsilon$, (5.42) implies that

$$\frac{9}{10}|1+m| \leqslant |m| \leqslant \frac{11}{10}|1+m|, \quad \text{in } \widetilde{\Omega}$$

It completed the proof of (5.24).

We note: combining (4.3), (5.1), (5.21) and (5.24), we have for any $|U|$, $|T| = O(1)$,

$$m^{(U,T)} \sim m \sim 1 + m \sim 1 + m^{(U,T)}, \quad |U|, \ |T| = O(1) \tag{5.43}$$

5. For (5.25), it follows from (5.23)(with $\mathcal{G}_{ii}^{(i,\emptyset)}$ in the l.h.s.), (5.43) and (5.16).

6. For (5.26), first using (5.5), (5.12), (5.13) and (5.20), we obtain that

$$|G_{ij}^{(\emptyset,i)}| \leqslant \varphi^C |w| |G_{ii}^{(\emptyset,i)}| |G_{jj}^{(i,i)}| \sqrt{\frac{\operatorname{Im} m_{\mathcal{G}}^{(ij,i)} + |z|^2 \operatorname{Im} \mathcal{G}_{jj}^{(ij,i)}}{N\eta}} \tag{5.44}$$

holds with 1-high probability in Ω. Applying (5.16) on $X^{(i,i)}$ instead of X, we obtain that

$$|w| |G_{jj}^{(i,i)}| |\mathcal{G}_{jj}^{(ij,i)}| \leqslant (\log N)^4 \tag{5.45}$$

Recall (5.1) implies (5.19). Applying (5.25) on $G_{jj}^{(i,i)}$, we have that

$$|G_{jj}^{(i,i)}| \leqslant (\log N)^C |m^{(i,i)}| \tag{5.46}$$

holds with 1-high probability in Ω. Then inserting (5.45), (5.46), (5.23) and (5.43) into (5.44), with (5.18) we obtain (5.26).

7. For (5.27), from (5.3), we have

$$G_{ii} = G_{ii}^{(\emptyset,j)} - \frac{(G^{(\emptyset,j)}\mathbf{y}_j^*)_i \, (\mathbf{y}_j G^{(\emptyset,j)})_i}{1 + \mathbf{y}_j G^{(\emptyset,j)} \mathbf{y}_j^*},$$

On the other hand, (5.6) and (5.13) show that (similar result can be seen in (6.18) of [9])

$$\mathcal{G}_{jj} = -w^{-1}(1 + \mathbf{y}_j G^{(\emptyset,j)}\mathbf{y}_j^*)^{-1}$$

Then

$$G_{ii} = G_{ii}^{(\emptyset,j)} + w\mathcal{G}_{jj}\left((G^{(\emptyset,j)}X^T)_{ij} - G_{ij}^{(\emptyset,j)}z^* \right)\left((XG^{(\emptyset,j)})_{ji} - G_{ji}^{(\emptyset,j)}z \right) \tag{5.47}$$

Since X_{jk}'s $(1 \leqslant k \leqslant N)$ are independent of $G^{(\emptyset,j)}$, using large deviation lemma (e.g. see Lemma 6.7 [9]), as in (3.44) of [10], we have that with 1-high probability,

$$|(XG^{(\emptyset,j)})_{ji}| + |(G^{(\emptyset,j)}X^T)_{ij}| \leqslant \varphi^C \sqrt{\frac{\operatorname{Im} G_{ii}^{(\emptyset,j)}}{N\eta}} \tag{5.48}$$

Inserting this bound, (5.25), (5.26) and (5.43) into (5.47), we have

$$|G_{ii} - G_{ii}^{(\emptyset,j)}| \leqslant \varphi^C w|m| \left(\frac{\operatorname{Im} G_{ii}^{(\emptyset,j)}}{N\eta} + \frac{1}{w|m|N\eta} \right)$$

i.e.,

$$G_{ii} = \left(1 + O(\frac{|w|m}{N\eta}) \right) G_{ii}^{(\emptyset,j)} + O(\frac{\varphi^C}{N\eta})$$

It implies that

$$G_{ii}^{(\emptyset,j)} = \left(1 + O(\frac{|w|m}{N\eta}) \right) G_{ii} + O(\frac{\varphi^C}{N\eta})$$

Then with (5.15) and (5.18), it implies

$$|G_{ii} - G_{ii}^{(\emptyset,j)}| \leqslant \varphi^C (N\eta)^{-1}$$

and we obtain (5.27).

8. For (5.28), using (5.4) and (5.20), we have

$$G_{ii}^{(\emptyset,ij)} = -w^{-1}[1 + m_{\mathcal{G}}^{(i,ij)} + \mathcal{Z}_i^{(ij)}]^{-1}$$

Using (5.10) and (5.20) again, we can bound $\mathcal{Z}_i^{(ij)}$ as

$$|\mathcal{Z}_i^{(ij)}| \leqslant \varphi^C \sqrt{\frac{\operatorname{Im} m_{\mathcal{G}}^{(i,ij)}}{N\eta}}$$

Together with (5.43) and (5.21), we obtain (5.28).

9. For (5.29), using (5.5), (5.12) and (5.13), we obtain that

$$|G_{ij}| \leqslant \varphi^C |w||G_{ii}||G_{jj}^{(i,\emptyset)}| \sqrt{\frac{\operatorname{Im} m_{\mathcal{G}}^{(ij,\emptyset)} + |z|^2 \operatorname{Im} \mathcal{G}_{jj}^{(ij,\emptyset)} + |z|^2 \operatorname{Im} \mathcal{G}_{ii}^{(ij,\emptyset)}}{N\eta}} + \varphi^C |wz^2||G_{ii}||G_{jj}^{(i,\emptyset)}||\mathcal{G}_{ij}^{(ij,\emptyset)}| \quad (5.49)$$

Furthermore, with (5.7), (5.11), (5.20) and (5.43), we have

$$|\mathcal{G}_{ij}^{(ij,\emptyset)}| \leqslant \varphi^C |w||\mathcal{G}_{ii}^{(ij,\emptyset)}||\mathcal{G}_{jj}^{(ij,i)}| \sqrt{\frac{\operatorname{Im} m^{(ij,ij)}}{N\eta}} \leqslant \varphi^C |w||\mathcal{G}_{ii}^{(ij,\emptyset)}||\mathcal{G}_{jj}^{(ij,i)}| \sqrt{\frac{|m|}{N\eta}} \quad (5.50)$$

Here these two bounds holds with 1-high probability. As in (5.46), applying (5.23) on $\mathcal{G}_{jj}^{(ij,i)}$, with (5.43) we have

$$|\mathcal{G}_{jj}^{(ij,i)}| \leqslant C|w|^{-1}|m^{(ij,ij)}|^{-1} \leqslant C|w|^{-1}|m|^{-1}$$

with 1-high probability in Ω. With (5.25), (5.28), (4.3) and (5.21), we also have

$$|G_{ii}| \leqslant (\log N)^C |m|, \quad |\mathcal{G}_{ii}^{(ij,\emptyset)}| + |\mathcal{G}_{jj}^{(ij,\emptyset)}| \leqslant C|w|^{-1}|m|^{-1}, \quad |m_{\mathcal{G}}^{(ij,\emptyset)}| \leqslant C|m|,$$

For the $G_{jj}^{(i,\emptyset)}$ in (5.49), as in (5.47) and (5.48), with (5.20), we have

$$G_{jj}^{(i,\emptyset)} - G_{jj}^{(i,i)} = w\mathcal{G}_{ii}^{(i,\emptyset)}(G^{(i,i)}X^T)_{ji}(XG^{(i,i)})_{ij} \tag{5.51}$$
$$= O\left(\varphi^C |w\mathcal{G}_{ii}^{(i,\emptyset)}|\,\mathrm{Im}\,G_{jj}^{(i,i)}(N\eta)^{-1}\right)$$

Then applying (5.25) on $G_{jj}^{(i,i)}$, and applying (5.23) on $\mathcal{G}_{ii}^{(i,\emptyset)}$, with (5.43) we obtain that

$$|G_{jj}^{(i,\emptyset)}| \leqslant (\log N)^C |m|$$

Inserting these bounds into (5.49) and (5.50), we obtain (5.29).

10. For (5.30), using (5.10) (with $\mathbb{T} = \emptyset$) and (5.23), (5.43) , we have

$$|\mathcal{Z}_i| \leqslant \varphi^C \sqrt{\frac{|m| + (|wm|)^{-1}}{N\eta}} \tag{5.52}$$

holds with 1-high probability in Ω. Together with (5.18), we obtain

$$|\mathcal{Z}_i| \leqslant \varphi^C \sqrt{\frac{(|wm|)^{-1}}{N\eta}} \tag{5.53}$$

Together with (5.25) and (5.18), we have

$$|\mathcal{Z}_i||wG_{ii}| \leqslant \varphi^C \sqrt{\frac{|w|^{1/2}}{N\eta}}.$$

Then with (2.6), we obtain (5.30).

11. For (5.31), we note that (5.24) implies $|m| \geqslant (\log N)^{-1}$. Then with (5.43), we obtain (5.31). $\qquad\square$

5.3 Polynomialization of Green's functions: In this subsection, using the bounds we proved in the last subsection, we write the G's and $\mathcal{G}$'s as the polynomials in $\mathcal{F}$ and $\mathcal{F}_{1/2} \cdot \mathcal{F}$ (with negligible error).

We note: In the Lemma 3.2 and 4.9 we assumed $X_{ab} = 0$, but the bounds we proved in Lemma 5.6 and Lemma 5.7 still hold for this type of X, the similar detailed argument was given in Remark 3.8 of [2].

Lemma 5.8. *Lemma 5.6 and Lemma 5.7 still hold if one enforces $X_{st} = 0$ for some fixed $1 \leqslant s, t \leqslant N$.*

Note: Here s, t are allowed to be the same as the i, j in Lemma 5.6 and Lemma 5.7. For example, from (5.29), we have $|G_{st}| \leqslant \varphi^C m^{1/2}(N\eta)^{-1/2}$, even if $X_{st} = 0$.

By the definitions of $A_X^{(f)}$, $\mathscr{P}_{1,2,3}(X)$, $B_{1,2,3}(X)$ and $P_{1,2,3}(X)$, one can see that the values of $A_X^{(f)}$, $\mathscr{P}_{1,2,3}(X)$ would not change if one replaced the G's inside with $\chi_a G$'s. Therefore, instead of G's, we will write $\chi_a G$ as the polynomials in $\mathcal{F}$ and $\mathcal{F}_{1/2} \cdot \mathcal{F}$ (with negligible error).

31

Definition 5.9. *For simplicity, we define the notations:*

$$\alpha := \chi_a |m^{(a,a)}|, \quad \beta := \frac{\chi_a}{|wm^{(a,a)}|}, \quad \gamma = \chi_a |w|^{1/2} \sqrt{\frac{|m^{(a,a)}|}{N\eta}}$$

We collect some basic properties of these quantities in the the following lemma.

Lemma 5.10. *Under the assumption of Lemma 3.2, for z, w:* $||z^2| - 1| \leqslant 2\varepsilon$ *and* $w \in I_\varepsilon$

$$\chi_a (\log N)^{-1} \leqslant \alpha \leqslant (\log N)^C \beta \leqslant (\log N)^C \eta^{-1} \tag{5.54}$$

$$\chi_a (\log N)^{-1} N^{-1/2} \leqslant \gamma \leqslant N^{-\varepsilon/2} \tag{5.55}$$

$$\beta\gamma^2 = \chi_a (N\eta)^{-1} \tag{5.56}$$

$$\frac{\chi_a (\log N)^C}{N\eta} \leqslant \alpha \leqslant \chi_a (\log N)^C |w^{-1/2}| \tag{5.57}$$

hold with 1-high probability.

Proof of Lemma 5.10: We note $\chi_a = 1$ implies the condition (5.1). Hence the results in Lemma 5.7 hold with 1- high probability. First from (5.31) and $|w| \geqslant \eta$, we have the first and the third inequalities of (5.54), and the first inequality of (5.55). The second inequality in (5.54) follows from (5.18) and (5.43). It also implies the second inequality of (5.57). Combining the second inequality of (5.54) with (2.6), we obtain the second inequality in (5.55). For (5.56), one can easily check this identity by the definition of β and γ. For the first inequality of (5.57), it follows from (5.21) and (5.43). □

Definition 5.11. *Under the assumption of Lemma 3.2, for* $w \in I_\varepsilon$, $||z| - 1| \leqslant 2\varepsilon$ *and* $s, k \neq a$, *we define* S_{ks} *and* $\widetilde{S}_{sk}$ *as random variables which are independent of the a-th row and columns of X and*

$$\frac{G_{ka}^{(\emptyset,a)}}{G_{aa}^{(\emptyset,a)}} = \sum_s^{(a)} S_{ks} X_{sa} \quad \text{and} \quad \frac{G_{ak}^{(\emptyset,a)}}{G_{aa}^{(\emptyset,a)}} = \sum_s^{(a)} X_{sa} \widetilde{S}_{sk}$$

With (5.5), one can obtain their explicit expressions, e.g.,

$$S_{ks} := z^* w G_{kk}^{(a,a)} \mathcal{G}_{ks}^{(ak,a)} - w G_{kk}^{(a,a)} \sum_t^{(a)} \mathcal{G}_{st}^{(ak,a)} X_{tk}$$

Similarly, we define $\mathcal{S}_{ks}$ *and* $\widetilde{\mathcal{S}}_{sk}$ *as random variables which are independent of the a-th row and columns of X and*

$$\frac{\mathcal{G}_{ka}^{(\emptyset,a)}}{\mathcal{G}_{aa}^{(a,\emptyset)}} = \sum_s \mathcal{S}_{ks} X_{as} \quad \text{and} \quad \frac{\mathcal{G}_{ak}^{(\emptyset,a)}}{\mathcal{G}_{aa}^{(a,\emptyset)}} = \sum_s X_{as} \widetilde{\mathcal{S}}_{sk}$$

As one can see that S, $\widetilde{S}$, $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ have the same behaviors. Here we collect some basic properties of these quantities in the the following lemma.

Lemma 5.12. *We assume that* $||z| - 1| \leqslant 2\varepsilon$, $w \in I_\varepsilon$, $k \neq a$ *and X satisfies the assumption of Lemma 3.2. For some* $C > 0$, *with 1- high probability, we have*

$$|\chi_a S_{ks}| \leqslant \chi_a \varphi^C (\delta_{sk} + \gamma) \tag{5.58}$$

32

so as $\widetilde{S}$, $\mathcal{S}$ and $\widetilde{\mathcal{S}}$. Recall the definition $\mathcal{F}$'s in Def. 4.3, for some $C > 0$, we have

$$\chi_a X_{aa} \in_n \gamma\mathcal{F}, \quad \chi_a(XSX)_{aa} \in_n \gamma\mathcal{F}, \tag{5.59}$$

and

$$\chi_a(X^T \widetilde{S} SX)_{aa} \in_n N\gamma^2\mathcal{F} \tag{5.60}$$

Furthermore, (5.58), (5.59) and (5.60) hold uniformly for $||z| - 1| \leqslant 2\varepsilon$, $w \in I_\varepsilon$ and $k, s : k, s \neq a$, $1 \leqslant k, s \leqslant N$.

Note: With (5.59), we also have

$$\chi_a \left(G^{(\emptyset,a)}_{aa}\right)^{-1} (XG^{(\emptyset,a)})_{aa} = \chi_a \left(G^{(\emptyset,a)}_{aa}\right)^{-1} \sum_k X_{ak} G^{(\emptyset,a)}_{ka} = \chi_a \left((XSX)_{aa} + X_{aa}\right) \in_n \gamma\mathcal{F} \tag{5.61}$$

Proof of Lemma 5.12: Since the uniformness are easy to be checked, we will only focus on the fixed z, w, s and k.

1. For (5.58), the condition $\chi_a = 1$ implies that we can apply Lemma 5.7 on the $X^{(a,a)}$. Recall: these bounds also hold under the exchange (5.32). Then the bounds (5.25) and (5.26) imply that for $s \neq k$,

$$\chi_a |G^{(a,a)}_{kk}| \leqslant (\log N)^C |m^{(a,a)}|, \quad \chi_a |\mathcal{G}^{(ak,a)}_{ks}| \leqslant \frac{\varphi^C}{|w^{1/2} m^{(a,a)}|} \sqrt{\frac{|m^{(a,a)}|}{N\eta}}, \tag{5.62}$$

holds with 1-high probability. Similarly (5.23) and (5.43) implies that for $s = k$

$$\chi_a |\mathcal{G}^{(ak,a)}_{kk}| \leqslant C |wm^{(a,a)}|^{-1}$$

holds with 1-high probability. Then with the explicit expression of S_{ks} in Def. 5.11, we have

$$\chi_a S_{ks} = O(\delta_{ks} + \varphi^C \gamma) - wG^{(a,a)}_{kk} \sum_t^{(a)} \mathcal{G}^{(ak,a)}_{st} X_{tk} \tag{5.63}$$

holds with 1-high probability. Since X_{tk}'s are independent of $\mathcal{G}^{(ak,a)}_{st}$'s $(1 \leqslant t \leqslant N)$, using large deviation lemma (e.g. see Lemma 6.7 [9]), as in (3.44) of [10], we have for

$$|\sum_t^{(a)} \mathcal{G}^{(ak,a)}_{st} X_{tk}| \leqslant \varphi^C \sqrt{\frac{\operatorname{Im} \mathcal{G}^{(ak,a)}_{ss}}{N\eta}} \tag{5.64}$$

holds with 1-high probability. Applying Lemma 5.7 on the $X^{(a,a)}$ again, from (5.27), we have

$$|\mathcal{G}^{(ak,a)}_{ss}| \leqslant (\log N)^C |m^{(a,a)}| + C\delta_{ks} \frac{1}{|wm^{(a,a)}|}$$

with 1-high probability. Together with the first part of (5.62), (5.63) and (5.64), we obtain

$$|\chi_a S_{ks}| \leqslant C\delta_{sk} + \varphi^C \gamma + \varphi^C |w^{1/2} m^{(a,a)}| \gamma \tag{5.65}$$

with 1-high probability. At last, with (5.18) and (5.43), we obtain (5.58).

2. For (5.59), we recall the definition of $\mathcal{F}$ in Def. 4.3, especially the two $N^{1/2}$ factors in $\mathcal{F}$. It is easy to see that (5.59) follows from the first inequality of (5.55) and the bounds on S in (5.58).

3. For (5.60), since the (5.58) also holds for $\widetilde{S}$, then with the first inequality of (5.55), we have

$$|\chi_a(\widetilde{S}S)_{kl}| \leqslant \varphi^C \left(\delta_{kl} + \gamma + N\gamma^2 \right) \leqslant \varphi^C N\gamma^2$$

with 1-high probability. Together with definition of $\mathcal{F}$, we obtain (5.60).

$\square$

Now we introduce a method to track and show the dependence of the random variables on the indices. First we give a simple example to show the basic idea. Let A_{kl}, $1 \leqslant k, l \leqslant N$ be a family of random variables:

$$A_{kl} = \frac{G_{kk}^{(a,a)}}{|G_{kk}^{(a,a)}|} \frac{G_{ll}^{(a,a)}}{|G_{ll}^{(a,a)}|} (XG^{(a,a)}X^T)_{aa}, \quad 1 \leqslant k, l \leqslant N \tag{5.66}$$

where X^T is the transpose of X. By definition of $\mathcal{F}$ and $\mathcal{F}_0$, we can say,

$$A_{kl} \in \mathcal{F}_0 \cdot \mathcal{F}_0 \cdot \mathcal{F}_1 \in \mathcal{F}$$

But the first part of the r.h.s. of (5.66), i.e., $\frac{G_{kk}^{(a,a)}}{|G_{kk}^{(a,a)}|}$ only depends on the first index k, the second part $\frac{G_{ll}^{(a,a)}}{|G_{ll}^{(a,a)}|}$ only depends on the second index l and the third part is independent of the indices. Therefore, we prefer to write it as

$$A_{kl} \in \mathcal{F}_0^{[k]} \cdot \mathcal{F}_0^{[l]} \cdot \mathcal{F}_1^{[\emptyset]}.$$

More precisely, $A_{kl} \in \mathcal{F}_0^{[k]} \cdot \mathcal{F}_0^{[l]} \cdot \mathcal{F}_1^{[\emptyset]}$ means that $A_{kl} = f_1(k)f_2(l)f_3$ and $f_1(k) \in \mathcal{F}_0$, $f_2(l) \in \mathcal{F}_0$, $f_3 \in \mathcal{F}_1$, and $f_1(k)$ only depends on index k, $f_2(l)$ only depends on index l, and f_3 does not depends on index.

For general case, to show how the variable depends on the indices, we define the following notations.

Definition 5.13. *Let A_I be a family of random variables where I is indices (vector), not including index a. we write*

$$A_I \in \prod_i \mathcal{F}_{\alpha_i}^{[I_i]}, \quad \mathcal{F}_{\alpha_i} \in \{\mathcal{F}_0, \mathcal{F}_{1/2}, \mathcal{F}_1, \mathcal{F}\}$$

where I_i is a part of I, if and only if there exists $f_i(I_i) \in \mathcal{F}_{\alpha_i}$ such that $A_I = \prod_i f_i(I_i)$ and $f_i(I_i)$ only depends on the indices in I_i.

For the example in (5.66), we write $A_{kl} \in \mathcal{F}_0^{[k]} \cdot \mathcal{F}_0^{[l]} \cdot \mathcal{F}^{[\emptyset]}$, where $I = (k, l)$, $I_1 = (k)$, $I_2 = (l)$ and $I_3 = (\emptyset)$, $\alpha_1 = \alpha_2 = 0$ and $\alpha_3 = 1$

The following lemma shows the G's can be written as the polynomials in $\mathcal{F}$'s.

Lemma 5.14. *For simplicity, we introduce the notaion:*

$$\mathcal{F}_{0,X}^{[k]} := X_{ak}\mathcal{F}_0^{[k]} + X_{ka}\mathcal{F}_0^{[k]} \tag{5.67}$$

i.e.,

$$f_k \in \mathcal{F}_{0,X}^{[k]} \iff \exists g_k, h_k \in \mathcal{F}_0^{[k]} : f_k = X_{ak}g_k + X_{ka}h_k$$

Let $w \in I_\varepsilon$ and $||z| - 1| \leqslant 2\varepsilon$. Under the assumption of Lemma 3.2, for any $D > 0$, we have

34

$$\chi_a G_{aa}^{(\emptyset,a)}, \in_n \beta\mathcal{F} + O_\prec(N^{-D}) \tag{5.68}$$

and

$$\chi_a G_{aa} \in_n \alpha\mathcal{F} + O_\prec(N^{-D}) \tag{5.69}$$

For any $k \neq a$,

$$\chi_a (G_{aa}^{(\emptyset,a)})^{-1} G_{ka}^{(\emptyset,a)} \in_n \gamma\mathcal{F}_{1/2}^{[k]} + \mathcal{F}_{0,X}^{[k]} + O_\prec(N^{-D}) \tag{5.70}$$

and,

$$\chi_a G_{ak} \in_n \sqrt{\frac{\alpha}{N\eta}}\mathcal{F}_{1/2}^{[k]} \cdot \mathcal{F}^{[\emptyset]} + (\alpha + \beta\gamma)\,\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}_{0,X}^{[k]} + O_\prec(N^{-D}) \tag{5.71}$$

For any $k,l \neq a$,

$$\chi_a \left(G_{kl} - G_{kl}^{(a,a)}\right) \in_n \left(\frac{\chi_a}{N\eta}\mathcal{F}_{1/2}^{[k]}\mathcal{F}_{1/2}^{[l]} + \beta\gamma\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{1/2}^{[l]} + \beta\gamma\mathcal{F}_{0,X}^{[l]}\mathcal{F}_{1/2}^{[k]} + \beta\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{0,X}^{[l]}\right)\mathcal{F}^{[\emptyset]} + O_\prec(N^{-D}) \tag{5.72}$$

Furthermore, (5.68)-(5.72) hold uniformly for $||z| - 1| \leqslant 2\varepsilon$, $w \in I_\varepsilon$ and $1 \leqslant k,l \neq a \leqslant N$

Proof of Lemma 5.14: Because one can easily check the uniformness, in the following proof we will only focus on the fixed w, z, k and l. Recall (4.18) and (5.33), with the assumption $w \in I_\varepsilon$ and $||z| - 1| \leqslant 2\varepsilon$, we know the results in Lemma 5.6 and 5.7 hold under the assumption of this lemma. Furthermore, these results also hold for $X^{(a,a)}$(instead of X).

1. We first prove (5.68). Applying Lemma 5.7 on $X^{(a,a)}$, with (5.25), (5.29) and the first inequality of (5.57), we have

$$\chi_a G_{kl}^{(a,a)} \in \left(\delta_{kl}\,\alpha + |w^{-\frac{1}{2}}|\gamma\right)\mathcal{F}_0, \quad \text{and} \quad |w^{-\frac{1}{2}}|\gamma \leqslant \alpha \tag{5.73}$$

Then with

$$Z_a^{(a)} = \left((XG^{(a,a)}X^T)_{aa} - m^{(a,a)}\right) \tag{5.74}$$

and $\alpha := \chi_a m^{(a,a)}$, we have

$$\chi_a(XG^{(a,a)}X^T)_{aa} \in_n \alpha\,\mathcal{F} \quad \text{and} \quad \chi_a Z_a^{(a)} \in_n \alpha\,\mathcal{F}. \tag{5.75}$$

From (5.6) and (5.20) with $i = a$, $\mathbb{T} = a$, we have

$$\chi_a \mathcal{G}_{aa}^{(a,\emptyset)} = \chi_a \frac{1}{-w}\frac{1}{1 + m^{(a,a)} + Z_a^{(a)}}$$

Then with (5.22), for any $\varepsilon, D > 0$, there exists $C_{\varepsilon,D}$ depending ε and D, such that

$$\chi_a \mathcal{G}_{aa}^{(a,\emptyset)} = \chi_a \frac{1}{-w}\sum_{k=1}^{C_{\varepsilon,D}}\left(\frac{1}{(1 + m^{(a,a)})^k}(Z_a^{(a)})^{k-1}\right) + O_\prec(N^{-D})$$

holds with 1-high probability. Hence with (5.43) and $\chi_a Z_a^{(a)} \in_n m^{(a,a)}\mathcal{F}$ in (5.75), we obtain that

$$\chi_a \mathcal{G}_{aa}^{(a,\emptyset)} \in_n \frac{1}{wm^{(a,a)}}\mathcal{F} + O_\prec(N^{-D}) = \beta\mathcal{F} + O_\prec(N^{-D}) \tag{5.76}$$

which implies (5.68) with the fact: $\mathcal{G}_{aa}^{(a,\emptyset)}$ and $G_{aa}^{(\emptyset,a)}$ have the same behavior.

2. Now we prove (5.69). From (5.30) and (5.4), with $i = a$ and $T = \emptyset$, for any $\varepsilon, D > 0$, there exists $C_{\varepsilon,D}$ depending ε and D such that with 1-high probability,

$$\chi_a G_{aa} = \sum_{k=1}^{C_{\varepsilon,D}} \frac{-w^{-1}\chi_a}{(1 + m_{\mathcal{G}}^{(a,\emptyset)} + |z|^2 \mathcal{G}_{aa}^{(a,\emptyset)})^k}(\mathcal{Z}_a)^{k-1} + O_{\prec}(N^{-D}) \tag{5.77}$$

Note: $1 + m_{\mathcal{G}}^{(a,\emptyset)} + |z|^2 \mathcal{G}_{aa}^{(a,\emptyset)}$ is independent of the a-th column of X, but depends on the a-th row of X. From (5.10) and (5.13), we have

$$\mathcal{Z}_a = z\sum_k (X^T)_{ak}\mathcal{G}_{ka}^{(a,\emptyset)} + z^*\sum_k \mathcal{G}_{ak}^{(a,\emptyset)} X_{ka} + \sum_{kl}(X^T)_{ak}\mathcal{G}_{kl}^{(a,\emptyset)} X_{la} - m_{\mathcal{G}}^{(a,\emptyset)} - |z|^2 \mathcal{G}_{aa}^{(a,\emptyset)} \tag{5.78}$$

Now we claim that for any D,

$$\chi_a \mathcal{Z}_a \in_n \beta \mathcal{F} + O_{\prec}(N^{-D}) \tag{5.79}$$

and

$$\chi_a \left(1 + m_{\mathcal{G}}^{(a,\emptyset)} + |z|^2 \mathcal{G}_{aa}^{(a,\emptyset)}\right)^{-1} \in_n w\alpha \mathcal{F} + O_{\prec}(N^{-D}) \tag{5.80}$$

Combining (5.79), (5.80) and (5.77), we obtain (5.69).

2.a We prove (5.79) first. Using the $\mathcal{G}$ version of (5.61) and (5.76), we can write the first two terms of the r.h.s. of (5.78) as we can write

$$\chi_a z\sum_k (X^T)_{ak}\mathcal{G}_{ka}^{(a,\emptyset)} + \chi_a z^*\sum_k \mathcal{G}_{ak}^{(a,\emptyset)} X_{ka} = 2\chi_a \operatorname{Re} z \left(\mathcal{G}^{(a,\emptyset)}X\right)_{aa} \in_n \beta\gamma\mathcal{F} + O_{\prec}(N^{-D}) \tag{5.81}$$

Similarly for the third term of the r.h.s. of (5.78), using (5.2), we can write it as

$$(X^T\mathcal{G}^{(a,\emptyset)}X)_{aa} = \sum_{kl}(X^T)_{ak}\mathcal{G}_{kl}^{(a,a)} X_{la} + (\mathcal{G}_{aa}^{(a,\emptyset)})^{-1}\sum_{kl}(X^T)_{ak}\mathcal{G}_{ka}^{(a,\emptyset)}\mathcal{G}_{al}^{(a,\emptyset)} X_{la}$$

$$= (X^T\mathcal{G}^{(a,a)}X)_{aa} + (\mathcal{G}_{aa}^{(a,\emptyset)})^{-1}\left(\mathcal{G}^{(a,\emptyset)}X\right)_{aa}\left(\mathcal{G}^{(a,\emptyset)}X\right)_{aa}$$

Using (5.75), (5.61) and (5.76), we obtain

$$\chi_a (X^T\mathcal{G}^{(a,\emptyset)}X)_{aa} \in_n \alpha\mathcal{F} + \beta\gamma^2\mathcal{F} + O_{\prec}(N^{-D}), \tag{5.82}$$

For the fourth term of the r.h.s. of (5.78), using (5.2), we have

$$\mathcal{G}_{kk}^{(a,a)} = \mathcal{G}_{kk}^{(a,\emptyset)} - \frac{\mathcal{G}_{ka}^{(a,\emptyset)}\mathcal{G}_{ak}^{(a,\emptyset)}}{\mathcal{G}_{aa}^{(a,\emptyset)}}$$

Together with (5.60), it implies that

$$m_{\mathcal{G}}^{(a,\emptyset)} = m^{(a,a)} + \frac{1}{N}\mathcal{G}_{aa}^{(a,\emptyset)}\left(\left(X\tilde{S}SX^T\right)_{aa} + 1\right) \tag{5.83}$$

and

$$\chi_a m_{\mathcal{G}}^{(a,\emptyset)} \in_n \alpha\mathcal{F} + \beta\gamma^2\mathcal{F} + \frac{1}{N}\beta\mathcal{F} + O_{\prec}(N^{-D})$$

36

Now inserting these bounds back to (5.78) and using the relations between α, β and γ in (5.54) and (5.55), we obtain (5.79).

2.b Now we prove (5.80). With (5.83) and

$$(\mathcal{G}_{aa}^{(a,\emptyset)})^{-1} = -w(1 + (XG^{(a,a)}X^T)_{aa}) = -w(1 + m^{(a,a)} + Z_a^{(a)})$$

we write

$$\frac{1}{1 + m_{\mathcal{G}}^{(a,\emptyset)} + |z|^2 \mathcal{G}_{aa}^{(a,\emptyset)}} = \frac{\left(\mathcal{G}_{aa}^{(a,\emptyset)}\right)^{-1}}{\frac{1+m^{(a,a)}}{\mathcal{G}_{aa}^{(a,\emptyset)}} + \frac{1}{N}\left(\left(X\widetilde{\mathcal{S}}\mathcal{S}X^T\right)_{aa} + 1\right) + |z|^2} \tag{5.84}$$

$$= \frac{-w(1 + (XG^{(a,a)}X^T)_{aa})}{-w(1 + m^{(a,a)})(1 + m^{(a,a)} + Z_a^{(a)}) + \frac{1}{N}\left(\left(X\widetilde{\mathcal{S}}\mathcal{S}X^T\right)_{aa} + 1\right) + |z|^2}$$

We write this denominator as

$$\left(-w(1 + m^{(a,a)})(1 + m^{(a,a)}) + |z|^2\right) + \left(-w(1 + m^{(a,a)})Z_a^{(a)} + \frac{1}{N}\left(\left(X\widetilde{\mathcal{S}}\mathcal{S}X^T\right)_{aa} + 1\right)\right) \tag{5.85}$$

With (5.22), (5.43), (5.18), we can bound the first term in the second bracket as follows:

$$\chi_a|w(1 + m^{(a,a)})Z_a^{(a)}| \leqslant N^{-\varepsilon/5}$$

holds with 1-high probability. Together with (5.60) and (5.55), with 1-high probability, we can bound the second bracket of (5.85) as

$$\chi_a\left(-w(1 + m^{(a,a)})Z_a^{(a)} + \frac{1}{N}\left(\left(X\widetilde{\mathcal{S}}\mathcal{S}X^T\right)_{aa} + 1\right)\right) \leqslant N^{-\varepsilon/6} \tag{5.86}$$

On the other hand, we claim for some $C > 0$, the following inequality holds with 1-hight probability.

$$\chi_a\left|-w(1 + m^{(a,a)})(1 + m^{(a,a)}) + |z|^2\right| \geqslant \chi_a(\log N)^{-C} \tag{5.87}$$

If (5.87) does not hold, then $\chi_a = 1$ and $1 + m^{(a,a)} = (-|z| + O(\log N)^{-C})w^{-1/2}$. With (4.3), (5.21) and $||z| - 1| \leqslant 2\varepsilon$, we obtain

$$1 + m_{\mathcal{G}}^{(a,\emptyset)} = (-|z| + O(\log N)^{-C})w^{-1/2}, \tag{5.88}$$

It follows from $1 + m^{(a,a)} = (-|z| + O(\log N)^{-C})w^{-1/2}$ and (5.23) that

$$\mathcal{G}_{aa}^{(a,\emptyset)} = (|z|^{-1} + O(\log N)^{-C})w^{-1/2} \tag{5.89}$$

Inserting them into (5.10), with (2.6), we have

$$|\mathcal{Z}_a| = O(\log N)^{-C}w^{-1/2} \tag{5.90}$$

Now insert (5.88), (5.89) and (5.90) into (5.4), we obtain $|G_{aa}| \geqslant (\log N)^{C-1}|w|^{-1/2}$ for any $C > 0$, which contradacts (5.15). Therefore, (5.87) must hold for some $C > 0$.

Recall the denominator of the r.h.s. of (5.84) equals to the sum of the l.h.s. of (5.86),(5.87) (see (5.85)). Then inserting (5.86),(5.87) into (5.85), we have that for any fixed D, there exists $C_{\varepsilon,D}$, such that with 1-high probability,

$$\frac{\chi_a}{1+m_{\mathcal{G}}^{(a,\emptyset)}+|z|^2\mathcal{G}_{aa}^{(a,\emptyset)}} = -\chi_a w(1+(XG^{(a,a)}X^T)_{aa}) \tag{5.91}$$

$$*\sum_{k=1}^{C_{\varepsilon,D}}\frac{\left((-w(1+m^{(a,a)})Z_a^{(a)}+\frac{1}{N}\left(\left(X\widetilde{S}SX^T\right)_{aa}+1\right)\right)^{k-1}}{\left(-w(1+m^{(a,a)})(1+m^{(a,a)})+|z|^2\right)^k} + O_\prec(N^{-D})$$

For the terms in (5.91), we apply (5.75) on $(XG^{(a,a)}X^T)_{aa}$ and $Z_a^{(a)}$, apply (5.24) on $(1+m^{(a,a)})$, apply (5.60) on $\left(X\widetilde{S}SX^T\right)_{aa}$, apply (5.55) on γ and apply (5.87) on the denominator of (5.91), we obtain

$$\frac{\chi_a}{1+m_{\mathcal{G}}^{(a,\emptyset)}+|z|^2\mathcal{G}_{aa}^{(a,\emptyset)}} \in_n -\chi_a\left(w\mathcal{F}+w\alpha\mathcal{F}\right)\sum_{k=1}^{C_{\varepsilon,D}}\left(-w\alpha^2\mathcal{F}+\gamma^2\mathcal{F}\right)^{k-1}+O_\prec(N^{-D})$$

With the bounds of α and γ in (5.54), (5.55) and (5.57), it implies (5.80). Combining (5.79), (5.80) and (5.77), we obtain (5.69).

3. For (5.70), it clearly follows the Def. 5.11, (5.58) and Def. 5.13.

4. Now we prove (5.71). First with (5.6) and (5.13), we have

$$(\mathcal{G}_{aa})^{-1} = -w(1+(YG^{(\emptyset,a)}Y^T)_{aa}). \tag{5.92}$$

Applying (5.3) on G_{ak} with $i=a$, recalling $Y=X-zI$, we have

$$G_{ak} = G_{ak}^{(\emptyset,a)} + w\mathcal{G}_{aa}\left((G^{(\emptyset,a)}X^T)_{aa}-z^*G_{aa}^{(\emptyset,a)}\right)\left((XG^{(\emptyset,a)})_{ak}-zG_{ak}^{(\emptyset,a)}\right) \tag{5.93}$$

$$=G_{ak}^{(\emptyset,a)}+w\mathcal{G}_{aa}G_{aa}^{(\emptyset,a)}|z|^2G_{ak}^{(\emptyset,a)}-zwG_{ak}^{(\emptyset,a)}\mathcal{G}_{aa}(G^{(\emptyset,a)}X^T)_{aa}$$

$$-z^*w\mathcal{G}_{aa}G_{aa}^{(\emptyset,a)}(XG^{(\emptyset,a)})_{ak}+w\mathcal{G}_{aa}(G^{(\emptyset,a)}X^T)_{aa}(XG^{(\emptyset,a)})_{ak}$$

Writing the first term in the r.h.s. as $G_{ak}^{(\emptyset,a)}\mathcal{G}_{aa}(\mathcal{G}_{aa})^{-1}$ and applying (5.92) on $(\mathcal{G}_{aa})^{-1}$, we can write the first three terms in the r.h.s. of (5.93) as

$$\left(-1-(XG^{(\emptyset,a)}X^T)_{aa}+z^*(XG^{(\emptyset,a)})_{aa}\right)w\mathcal{G}_{aa}G_{ak}^{(\emptyset,a)}$$

Therefore

$$G_{ak}=\left(-1-(XG^{(\emptyset,a)}X^T)_{aa}+z^*(XG^{(\emptyset,a)})_{aa}\right)w\mathcal{G}_{aa}G_{ak}^{(\emptyset,a)}+\left(-z^*G_{aa}^{(\emptyset,a)}+(G^{(\emptyset,a)}X^T)_{aa}\right)w\mathcal{G}_{aa}(XG^{(\emptyset,a)})_{ak} \tag{5.94}$$

Inserting (5.68)-(5.70), (5.81), (5.82), the fact: $\alpha\beta=\chi_a$ and (5.61) into (5.94), we have

$$\chi_aG_{ak}\in_n\left(1+\alpha+\beta\gamma+\beta\gamma^2\right)\mathcal{F}^{[\emptyset]}\left(\gamma\mathcal{F}_{1/2}^{[k]}+\mathcal{F}_{0,X}^{[k]}\right)+(1+\gamma)\mathcal{F}^{[\emptyset]}(XG^{(\emptyset,a)})_{ak}+O_\prec(N^{-D})$$

38

More precisely, here what we used is the G-version of (5.81), (5.82), i.e.,

$$\chi_a(XG^{(\emptyset,a)})_{aa} \in_n \beta\gamma\mathcal{F} \quad \text{and} \quad \chi_a(XG^{(\emptyset,a)}X^T)_{aa} \in (\alpha + \beta\gamma^2)\mathcal{F}.$$

They follows from (5.81), (5.82) and the symmetry between G and $\mathcal{G}$.

Next using (5.57), we have

$$\chi_a G_{ak} \in_n (\alpha + \beta\gamma)\mathcal{F}^{[\emptyset]}\left(\gamma\mathcal{F}^{[k]}_{1/2} + \mathcal{F}^{[k]}_{0,X}\right) + \mathcal{F}^{[\emptyset]}(XG^{(\emptyset,a)})_{ak} + O_{\prec}(N^{-D}) \tag{5.95}$$

$$\in_n \chi_a\sqrt{\frac{\alpha}{N\eta}}\mathcal{F}^{[k]}_{1/2} \cdot \mathcal{F}^{[\emptyset]} + (\alpha + \beta\gamma)\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}^{[k]}_{0,X} + \chi_a(XG^{(\emptyset,a)})_{ak}\mathcal{F}^{[\emptyset]} + O_{\prec}(N^{-D})$$

For $(XG^{(\emptyset,a)})_{ak}$ in (5.95), using (5.2), for $k \neq a$ we have (note: s can be a)

$$G^{(a,a)}_{sk} = G^{(\emptyset,a)}_{sk} - \frac{G^{(\emptyset,a)}_{sa}G^{(\emptyset,a)}_{ak}}{G^{(\emptyset,a)}_{aa}}$$

Together with (5.61), (5.68) and (5.70), it implies that

$$\chi_a(XG^{(\emptyset,a)})_{ak} = \chi_a(XG^{(a,a)})_{ak} + \chi_a\frac{(XG^{(\emptyset,a)})_{aa}}{G^{(\emptyset,a)}_{aa}}G^{(\emptyset,a)}_{ak} \tag{5.96}$$

$$\in_n \chi_a(XG^{(a,a)})_{ak} + \beta\gamma^2\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}^{[k]}_{1/2} + \beta\gamma\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}^{[k]}_{0,X} + O_{\prec}(N^{-D})$$

It follows from (5.73), (note: $|w^{-1/2}|\gamma = \alpha^{1/2}(N\eta)^{-1/2}$) that

$$\chi_a(XG^{(a,a)})_{ak} \in_n \sqrt{\frac{\alpha}{N\eta}}\mathcal{F}^{[k]}_{1/2} + \alpha\mathcal{F}^{[k]}_0 X_{ak} \in_n \sqrt{\frac{\alpha}{N\eta}}\mathcal{F}^{[k]}_{1/2} + \alpha\mathcal{F}^{[k]}_{0,X}.$$

Inserting it into (5.96), with Lemma 5.10, we obtain

$$\chi_a(XG^{(\emptyset,a)})_{ak} \in_n \sqrt{\frac{\alpha}{N\eta}}\mathcal{F}^{[k]}_{1/2} \cdot \mathcal{F}^{[\emptyset]} + (\alpha + \beta\gamma)\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}^{[k]}_{0,X} + O_{\prec}(N^{-D}) \tag{5.97}$$

Together with (5.95), we obtain (5.71).

5. Now we prove (5.72). With (5.97), (5.70) and Lem. 5.10, we have

$$\chi_\alpha\left((G^{(\emptyset,a)}X^T)_{ka} - z^*G^{(\emptyset,a)}_{ka}\right) \in_n \beta\gamma\mathcal{F}^{[k]}_{1/2} \cdot \mathcal{F}^{[\emptyset]} + \beta\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}^{[k]}_{0,X} + O_{\prec}(N^{-D}) \tag{5.98}$$

Together with (5.3), (5.92), (5.68) and (5.56), we can write G_{kl} as follow,

$$\chi_a\left(G_{kl} - G^{(\emptyset,a)}_{kl}\right) = \chi_a w\mathcal{G}_{aa}\left((G^{(\emptyset,a)}X^T)_{ka} - z^*G^{(\emptyset,a)}_{ka}\right)\left((XG^{(\emptyset,a)})_{al} - zG^{(\emptyset,a)}_{al}\right) \tag{5.99}$$

$$\in_n \frac{1}{\beta}\mathcal{F}^{[\emptyset]}\left(\beta\gamma\mathcal{F}^{[k]}_{1/2} \cdot \mathcal{F}^{[\emptyset]} + \beta\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}^{[k]}_{0,X}\right)\left(\beta\gamma\mathcal{F}^{[l]}_{1/2} \cdot \mathcal{F}^{[\emptyset]} + \beta\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}^{[l]}_{0,X}\right) + O_{\prec}(N^{-D})$$

$$\in_n \frac{1}{N\eta}\mathcal{F}^{[k]}_{1/2} \cdot \mathcal{F}^{[l]}_{1/2} \cdot \mathcal{F}^{[\emptyset]} + \beta\gamma\left(\mathcal{F}^{[k]}_{0,X} \cdot \mathcal{F}^{[l]}_{1/2} + \mathcal{F}^{[l]}_{0,X} \cdot \mathcal{F}^{[k]}_{1/2}\right) \cdot \mathcal{F}^{[\emptyset]} + \beta\mathcal{F}^{[k]}_{0,X} \cdot \mathcal{F}^{[l]}_{0,X} \cdot \mathcal{F}^{[\emptyset]} + O_{\prec}(N^{-D})$$

39

Furthermore, with (5.2), (5.68), (5.70) and (5.56), we can write $G_{kl}^{(\emptyset,a)}$ as

$$\chi_a\left(G_{kl}^{(\emptyset,a)} - G_{kl}^{(a,a)}\right) = \chi_a \frac{G_{ka}^{(\emptyset,a)} G_{al}^{(\emptyset,a)}}{G_{aa}^{(\emptyset,a)}} = \chi_a G_{aa}^{(\emptyset,a)} \frac{G_{ka}^{(\emptyset,a)}}{G_{aa}^{(\emptyset,a)}} \frac{G_{al}^{(\emptyset,a)}}{G_{aa}^{(\emptyset,a)}}$$

$$\in_n \frac{\chi_a}{N\eta}\mathcal{F}_{1/2}^{[k]}\mathcal{F}_{1/2}^{[l]}\mathcal{F}^{[\emptyset]} + \beta\gamma\left(\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{1/2}^{[l]} + \mathcal{F}_{0,X}^{[l]}\mathcal{F}_{1/2}^{[k]}\right)\mathcal{F}^{[\emptyset]} + \beta\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{0,X}^{[l]}\mathcal{F}^{[\emptyset]} + O_\prec(N^{-D})$$

Therefore, together with (5.99), we obtain (5.72). $\qquad\square$

Next, we write the terms appeared in the Lemma 4.9 as polynomials in $\mathcal{F}$, $\mathcal{F}_{1/2}$ and $\mathcal{F}_{1/2}\cdot\mathcal{F}$ (with proper coefficients and ignorable error terms).

Lemma 5.15. *Let $w \in I_\varepsilon$ and $||z| - 1| \leqslant 2\varepsilon$. Under the assumption of Lemma 3.2, for any fixed large $D > 0$, with χ_a defined in (4.17) and $F_{0,X}^{[k]}$ defined in (5.67), we have that for $k \neq a$*

$$\chi_a(m - m^{(a,a)}) \in_n \frac{1}{N\eta}\mathcal{F} + O_\prec(N^{-D}) \tag{5.100}$$

$$\chi_a G_{bb} \in_n \beta\mathcal{F} + O_\prec(N^{-D}) \tag{5.101}$$

$$\chi_a(YG)_{aa} \in_n \mathcal{F} + O_\prec(N^{-D}) \tag{5.102}$$

$$\chi_a(YG)_{ak} \in_n \gamma\mathcal{F}_{1/2}^{[k]} \cdot \mathcal{F}^{[\emptyset]} + \mathcal{F}_{0,X}^{[k]}\mathcal{F}^{[\emptyset]} + O_\prec(N^{-D}), \tag{5.103}$$

$$\chi_a(YG^2)_{ab} \in_n \frac{\chi_a}{\eta}\mathcal{F}_{1/2} \cdot \mathcal{F} + X_{ba}\mathcal{F} + O_\prec(N^{-D}). \tag{5.104}$$

$$\chi_a(G^2)_{aa} \in_n \frac{\alpha}{\eta}\mathcal{F} + O_\prec(N^{-D}) \tag{5.105}$$

$$\chi_a(G^2)_{bb} \in_n \beta\eta^{-1}\mathcal{F} + O_\prec(N^{-D}) \tag{5.106}$$

$$\chi_a(YG^2Y^*)_{aa} \in_n \frac{w\alpha}{\eta}\mathcal{F} + O_\prec(N^{-D}) \tag{5.107}$$

Proof of Lemma 5.15: **1.** For (5.100), using (5.72) and (5.69), we have

$$\chi_a(m - m^{(a,a)}) = \frac{\chi_a}{N}G_{aa} + \chi_a\frac{1}{N}\sum_{k\neq a}\left(G_{kk} - G_{kk}^{(a,a)}\right)$$

$$\in_n \frac{\alpha}{N}\mathcal{F} + \frac{1}{N}\sum_{k\neq a}\left(\frac{\chi_a}{N\eta}\mathcal{F}_{1/2}^{[k]}\mathcal{F}_{1/2}^{[k]} + \beta\gamma\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{1/2}^{[k]} + \beta\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{0,X}^{[k]}\right)\mathcal{F}^{[\emptyset]} + O_\prec(N^{-D})$$

$$\in_n \left(\frac{\alpha}{N} + \frac{1}{N\eta} + \frac{\beta\gamma}{N} + \frac{\beta}{N}\right)\mathcal{F} + O_\prec(N^{-D})$$

Here for the last $\in_n$, we used

$$\sum_{k\neq a}\frac{1}{N}\mathcal{F}_{1/2}^{[k]}\mathcal{F}_{1/2}^{[k]} \in \mathcal{F}, \quad \sum_{k\neq a}\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{1/2}^{[k]} \in \mathcal{F}, \quad \sum_{k\neq a}\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{0,X}^{[k]} \in \mathcal{F}. \tag{5.108}$$

Then with (5.54) and (5.55), we obtain (5.100).

2. For (5.101), it follows from (5.72), $\mathcal{F}_{1/2} \cdot \mathcal{F}_{1/2} \subset \mathcal{F}$ and the fact: $X_{ab} = 0$ that

$$\chi_a G_{bb} \in_n \chi_a G_{bb}^{(a,a)} + \frac{\chi_a}{N\eta}\mathcal{F} + \beta\gamma X_{ba}\mathcal{F}_{1/2} \cdot \mathcal{F} + \beta X_{ba}X_{ba}\mathcal{F}$$

$$\in_n \alpha\mathcal{F} + \frac{\chi_a}{N\eta}\mathcal{F} + (\beta + \gamma\beta)\mathcal{F}$$

where we used (5.25) on $G_{bb}^{(a,a)}$, $X_{ba} \in \mathcal{F}_{1/2}$. Now using Lemma 5.10, we obtain (5.101).

3. For (5.102), with (5.3) and (5.92), we can write it as

$$\chi_a(YG)_{aa} = -\chi_a w \mathcal{G}_{aa}(YG^{(\emptyset,a)})_{aa} = -\chi_a w \mathcal{G}_{aa}\left((XG^{(\emptyset,a)})_{aa} - zG_{aa}^{(\emptyset,a)}\right) \tag{5.109}$$

Then with (5.69), (5.68) and (5.81), we obtain (5.102).

4. Now we prove (5.103), with (5.3) and (5.92) again, we write it is

$$(YG)_{ak} = -w\mathcal{G}_{aa}(YG^{(\emptyset,a)})_{ak} = -w\mathcal{G}_{aa}\left((XG^{(\emptyset,a)})_{ak} - zG_{ak}^{(\emptyset,a)}\right) \tag{5.110}$$

Then using (5.98) and (5.69), we obtain (5.103).

5. For (5.104), by definition, we write $(YG^2)_{ab}$ as

$$(YG^2)_{ab} = \sum_{k\neq a}(YG)_{ak}G_{kb} + (YG)_{aa}G_{ab}$$

Then using (5.103), (5.72), with $X_{ab} = 0$, we get

$$\chi_a \sum_{k\neq a}(YG)_{ak}G_{kb} \tag{5.111}$$

$$\in_n \sum_{k\neq a}\mathcal{F}^{[\emptyset]}\left(\gamma\mathcal{F}_{1/2}^{[k]} + \mathcal{F}_{0,X}^{[k]}\right)\left(G_{kb}^{(a,a)} + \frac{\chi_a}{N\eta}\mathcal{F}_{1/2}^{[k]}\mathcal{F}_{1/2}^{[b]} + \beta\gamma\left(\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{1/2}^{[b]} + \mathcal{F}_{0,X}^{[b]}\mathcal{F}_{1/2}^{[k]}\right) + \beta\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{0,X}^{[b]}\right)$$

With (5.108) and Lemma 5.10, we obtain

$$\chi_a \sum_{k\neq a}(YG)_{ak}G_{kb} \in_n \mathcal{F}^{[\emptyset]}\sum_{k\neq a}\left(\gamma\mathcal{F}_{1/2}^{[k]} + \mathcal{F}_{0,X}^{[k]}\right)G_{kb}^{(a,a)} + \frac{\chi_a}{\eta}\left(\mathcal{F}_{1/2}^{[b]}\mathcal{F}^{[\emptyset]} + \mathcal{F}^{[b]}X_{ba}\right)$$

Then applying (5.73) on $G_{kb}^{(a,a)}$, we obtain $\chi_a G_{kb}^{(a,a)} \in \left(|w|^{-1/2}\gamma + \delta_{kb}\alpha\right)\mathcal{F}_0^{[k,b]}$. Now with

$$\sum_{k\neq a}\mathcal{F}_{1/2}^{[k]} \cdot \mathcal{F}_0^{[k,b]} \in N\mathcal{F}_{1/2}, \qquad \sum_{k\neq a}\mathcal{F}_{0,X}^{[k]} \cdot \mathcal{F}_0^{[k,b]} \in \mathcal{F}_{1/2}$$

and Lemma 5.10 again, we get

$$\chi_a \sum_{k\neq a}(YG)_{ak}G_{kb} \in_n \frac{\chi_a}{\eta}\left(\mathcal{F}_{1/2}^{[b]}\mathcal{F}^{[\emptyset]} + \mathcal{F}^{[b]}X_{ba}\right)$$

Similarly, with (5.102), (5.71) and Lemma 5.10 again, we obtain

$$\chi_a(YG)_{aa}G_{ab} \in_n \frac{\chi_a}{\eta}\left(\mathcal{F}_{1/2}^{[b]}\mathcal{F}^{[\emptyset]} + \mathcal{F}^{[b]}X_{ba}\right)$$

41

and we obtain (5.104).

6. For (5.105), we write $(G^2)_{aa}$ as

$$\chi_a(G^2)_{aa} = \chi_a \sum_{k\neq a} G_{ak}G_{ka} + \chi_a(G_{aa})^2 \in_n \sum_{k\neq a} \chi_a G_{ak}G_{ka} + \frac{\alpha}{\eta}\mathcal{F}$$

where for the second $\in_n$, we used (5.69) and (5.54). As in (5.111), using (5.71) and (5.108), we have

$$\chi_a \sum_{k\neq a} G_{ak}G_{ka} = \sum_{k\neq a} \left(\sqrt{\frac{\alpha}{N\eta}}\mathcal{F}_{1/2}^{[k]} \cdot \mathcal{F}^{[\emptyset]} + (\alpha + \beta\gamma)\mathcal{F}^{[\emptyset]} \cdot \mathcal{F}_{0,X}^{[k]} \right)^2$$

$$\in_n \left(\frac{\alpha}{\eta} + (\alpha + \beta\gamma)^2 \right)\mathcal{F}$$

Then with Lemma 5.10, we obtain (5.105).

7. (5.106), we write it as

$$\chi_a(G^2)_{bb} = \sum_{k\neq a,b} G_{bk}G_{kb} + (G_{bb})^2 + (G_{ab})^2$$

With (5.72), (5.108) and $\sum_k \mathcal{F}_\alpha^{[k,b]} \in N\mathcal{F}_\alpha^{[b]}$, $(\alpha = 0, 1/2, \emptyset)$, after a tedious calculation, we get

$$\chi_a \sum_{k\neq a,b} G_{bk}G_{kb}$$

$$= \sum_{k\neq a,b} \left(\sqrt{\frac{\alpha}{N\eta}}\mathcal{F}_0^{[k,b]} + \frac{\chi_a}{N\eta}\mathcal{F}_{1/2}^{[k]}\mathcal{F}_{1/2}^{[b]} + \beta\gamma \left(\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{1/2}^{[b]} + \mathcal{F}_{0,X}^{[b]}\mathcal{F}_{1/2}^{[k]} \right) + \beta\mathcal{F}_{0,X}^{[k]}\mathcal{F}_{0,X}^{[b]} \right)^2 \mathcal{F}^{[\emptyset]}$$

$$\in_n \frac{\alpha}{\eta}\mathcal{F} + \left(\frac{1}{\eta}\sqrt{\frac{\alpha}{N\eta}} + \beta\gamma\sqrt{\frac{\alpha}{N\eta}} + \frac{1}{N\eta^2} + \frac{\beta\gamma}{N\eta} + \beta^2\gamma^2 \right)\mathcal{F}_{1/2}^{[b]}\mathcal{F}_{1/2}^{[b]}\mathcal{F} + O_\prec(N^{-D})$$

$$+ \left(N\beta\gamma\sqrt{\frac{\alpha}{N\eta}} + \beta\sqrt{\frac{\alpha}{N\eta}} + \frac{\beta\gamma}{\eta} + \frac{\beta}{\eta} + \beta^2\gamma^2 + \beta^2\gamma \right) X_{ba}\mathcal{F}_{1/2}^{[b]}\mathcal{F} + \left(\beta^2\gamma^2 N + \beta^2\gamma + \beta^2 \right) X_{ba}^2 \mathcal{F}$$

Then using Lemma 5.10, $\mathcal{F}_{1/2}^{[b]}\mathcal{F}_{1/2}^{[b]} \in \mathcal{F}$, $X_{ba}\mathcal{F}_{1/2}^{[b]} \in \mathcal{F}$ and $X_{ba}^2 \in \mathcal{F}$, we obtain

$$\chi_a \sum_{k\neq a,b} G_{bk}G_{kb} \in_n \frac{\beta}{\eta}\mathcal{F} + O_\prec(N^{-D})$$

Similarly, using (5.72), and Lemma 5.10 we have

$$\chi_a G_{bb}G_{bb} \in_n \left(\alpha + \frac{1}{N\eta} + \beta\gamma + \beta \right)^2 \mathcal{F} + O_\prec(N^{-D}) \in_n \frac{\beta}{\eta}\mathcal{F} + O_\prec(N^{-D})$$

Using (5.69), and Lemma 5.10 we have

$$\chi_a G_{ba}G_{ab} \in_n \frac{\beta}{\eta}\mathcal{F} + O_\prec(N^{-D})$$

which completes the proof of (5.106).

8. For (5.107), it follows from

$$(YG^2Y^T)_{aa} = \mathcal{G}_{aa} + w(\mathcal{G}^2)_{aa}$$

and (5.69) and (5.105).

□

Now we are ready to prove Lemma 4.9, which is the key lemma in the proof of our main result.

5.4 Proof of lemma 4.9. First with $m - m^{(a,a)} = O(N\eta)^{-1}$ (see (4.3)) and the definition of χ_a, for any fixed $D > 0$, with 1-high probability, we can write the $h(t_X)$ as

$$h(t_X) = \chi_a h(t_X) = \sum_{k=0}^{C_{\varepsilon,D}} \frac{1}{k!} h^{(k)}(t_{X^{(a,a)}})\chi_a \left(\frac{\operatorname{Re} m - \operatorname{Re} m^{(a,a)}}{N^{\varepsilon}(N\eta)^{-1}}\right)^k + O(N^{-D})$$

where constant $C_{\varepsilon,D}$ depends on ε on D, and $h^{(k)}$ is the $k - th$ derivative of h. Using (5.100) and the fact that h is smooth and supported in $[1, 2]$, we obtain

$$h(t_X) \in_n \mathcal{F} + O_{\prec}(N^{-D}) \tag{5.112}$$

and

$$h(t_X) - h(t_{X^{(a,a)}}) \in_n N^{-\varepsilon}\mathbf{1}(|t_{X^{(a,a)}}| \leqslant 2)\mathcal{F} + O_{\prec}(N^{-D}) \tag{5.113}$$

Note: $\mathbf{1}(|t_{X^{(a,a)}}| \leqslant 2) = \mathbf{1}(|\operatorname{Re} m^{(a,a)}| \leqslant 2N^{\varepsilon}(N\eta)^{-1})$. Similarly, one can prove

$$h'(t_X), \quad h''(t_X), \quad h'''(t_X) \in_n \mathbf{1}(|t_{X^{(a,a)}}| \leqslant 2)\mathcal{F} + O_{\prec}(N^{-D}) \tag{5.114}$$

Using (5.112), (5.113) and (5.100), we have

$$\left(h(t_X)\operatorname{Re} m - h(t_{X^{(a,a)}})\operatorname{Re} m^{(a,a)}\right) \in_n \left(h(t_X)\operatorname{Re} m^{(a,a)} - h(t_{X^{(a,a)}})\operatorname{Re} m^{(a,a)}\right) + \frac{1}{N\eta}\mathcal{F} + O(N^{-D})$$

$$\in_n \frac{1}{N\eta}\mathcal{F} + O_{\prec}(N^{-D}) \tag{5.115}$$

It implies (4.19).

For (4.20), recall $B_m(X)$ is defined as

$$B_m(X) := \frac{1}{m!}(N^{1-\varepsilon}\eta)^{(m-1)}\left(mh^{(m-1)}(t_X) + h^{(m)}(t_X)t_X\right)$$

Then using (5.112), (5.114) and (5.100), we obtain (4.20).

Similarly, for (4.21), the terms appearing in the definition (3.20) have been all bounded in (5.104), (5.69), (5.106), (5.104), (5.107) and (5.101). With a simple calculation, one can obtain (4.21) and complete the proof.

□

References

[1] Y. Ameur, H. Hedenmalm, and N. Makarov, *Fluctuations of eigenvalues of random normal matrices*, Duke Mathematical Journal **159** (2011), 31–81.

[2] Y. Ameur and J. Ortega-Cerdà, *Beurling-Landau densities of weighted Fekete sets and correlation kernel estimates*, preprint `arXiv:1110.0284` (2011).

[3] Z. D. Bai, *Circular law*, Ann. Probab. **25** (1997), no. 1, 494–529.

[4] Z. D. Bai and J. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, Mathematics Monograph Series, vol. 2, Science Press, Beijing, 2006.

[5] Patrick Billingsley, *Probability and measure*, 3rd ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1995. A Wiley-Interscience Publication.

[6] A. Bloemendal, L. Erdoes, A. Knowles, H.T. Yau, and J. Yin, to appear (2013).

[7] Pavel Bleher and Robert Mallison Jr., *Zeros of sections of exponential sums*, Int. Math. Res. Not. (2006), Art. ID 38937, 49.

[8] A. Borodin and C. D. Sinclair, *The Ginibre ensemble of real random matrices and its scaling limits*, Comm. Math. Phys. **291** (2009), no. 1, 177–224.

[9] P. Bourgade, H.-T. Yau, and J. Yin, *Local circular law for random matrices*, preprint `arXiv:1206.1449` (2012).

[10] ———, *The local circular law II: the edge case*, preprint `arXiv:1206.3187` (2012).

[11] R. Boyer and W. Goh, *On the zero attractor of the Euler polynomials*, Adv. in Appl. Math. **38** (2007), no. 1, 97–132.

[12] O. Costin and J. Lebowitz, *Gaussian fluctuations in random matrices*, Phys. Rev. Lett. **75** (1995), no. 1, 69–72.

[13] E. B. Davies, *The functional calculus*, J. London Math. Soc. (2) **52** (1995), no. 1, 166–176.

[14] A. Edelman, *The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law*, J. Multivariate Anal. **60** (1997), no. 2, 203–232.

[15] L. Erdős, H.-T. Yau, and J. Yin, *Bulk universality for generalized Wigner matrices*, to appear in PTRF, preprint: `arXiv:1001.3453` (2010).

[16] ———, *Rigidity of Eigenvalues of Generalized Wigner Matrices*, to appear in Adv. Mat., preprint `arXiv:1007.4652` (2010).

[17] P. J. Forrester and T. Nagao, *Eigenvalue Statistics of the Real Ginibre Ensemble*, Phys. Rev. Lett. **99** (2007).

[18] J. Ginibre, *Statistical ensembles of complex, quaternion, and real matrices*, J. Mathematical Phys. **6** (1965), 440–449.

[19] V. L. Girko, *The circular law*, Teor. Veroyatnost. i Primenen. **29** (1984), no. 4, 669–679 (Russian).

[20] F. Götze and A. Tikhomirov, *The circular law for random matrices*, Ann. Probab. **38** (2010), no. 4, 1444–1491.

[21] T. Kriecherbauer, A. B. J. Kuijlaars, K. D. T.-R. McLaughlin, and P. D. Miller, *Locating the zeros of partial sums of e^z with Riemann-Hilbert methods*, Integrable systems and random matrices, Contemp. Math., vol. 458, Amer. Math. Soc., Providence, RI, 2008, pp. 183–195.

[22] Eugene Lukacs, *Characteristic functions*, Griffin's Statistical Monographs& Courses, No. 5. Hafner Publishing Co., New York, 1960.

[23] G. Pan and W. Zhou, *Circular law, extreme singular values and potential theory*, J. Multivariate Anal. **101** (2010), no. 3, 645–656.

[24] N. Pillai and J. Yin, *Universality of Covariance matrices*, preprint `arXiv:1110.2501` (2011).

[25] B. Rider and B. Virág, *The noise in the circular law and the Gaussian free field*, Int. Math. Res. Not. IMRN **2** (2007).

[26] M. Rudelson, *Invertibility of random matrices: Norm of the inverse*, Ann. of Math. **168** (2008), no. 2, 575–600.

[27] M. Rudelson and R. Vershynin, *The Littlewood-Offord problem and invertibility of random matrices*, Adv. Math. **218** (2008), no. 2, 600–633.

[28] C. D. Sinclair, *Averages over Ginibre's ensemble of random real matrices*, Int. Math. Res. Not. IMRN **5** (2007).

[29] A. Soshnikov, *Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields*, J. Statist. Phys. **100** (2000), no. 3-4, 491–522.

[30] T. Tao and V. Vu, *Random matrices: the circular law*, Commun. Contemp. Math. **10** (2008), no. 2, 261–307.

[31] _____, *Random matrices: universality of ESDs and the circular law*, Ann. Probab. **38** (2010), no. 5, 2023–2065. With an appendix by Manjunath Krishnapur.

[32] _____, *Random matrices: Universality of local spectral statistics of non-Hermitian matrices*, preprint `arXiv:1206.1893` (2012).

[33] P. Wood, *Universality and the circular law for sparse random matrices*, The Annals of Applied Probability **22** (2012), no. 3, 1266 - 1300.