The Local Semicircle Law for a Class of Random Matrices with a Fourfold Symmetry

Johannes Alt

Master Thesis
Master of Science Mathematics
LMU Munich

Advisor: Prof. László Erdős, PhD

Munich, 2014

Ich versichere hiermit,
dass ich die vorliegende Arbeit selbstständig verfasst
und keine anderen als die angegebenen Quellen
und Hilfsmittel verwendet habe.

München, den 14. August 2014

Johannes Alt

Contents

1 Introduction

A random matrix is a random variable with values in the real or complex matrices of a deterministic size. They were first introduced by Wishart in 1928. He used random matrices to model problems in mathematical statistics and data analysis [16]. In 1955, Wigner conjectured that the eigenvalues of random matrices should describe the energy levels of large atoms [15]. Since then, random matrices appeared in a number of physical models. Some applications are described in [13] and in Journal of Physics A **36**(12), 2003: *Special issue: Random Matrix Theory.*

The second example indicates that the distribution of the eigenvalues of a random matrix is a particularly interesting and often studied question in random matrix theory. This justifies the focus on square matrices, i.e. $N \times N$ matrices with $N \in \mathbb{N}$. For a random matrix with eigenvalues $(\lambda_i)_{i=1}^N$, this distribution is defined by $\mu_N := N^{-1} \sum_{i=1}^N \delta_{\lambda_i}$ and is called the *empirical spectral measure.* A first answer has already been given by Wigner in [15]. Under some assumptions about the entries, he could show by computing its moments that the empirical spectral measure follows the semicircle law induced by the density $\sqrt{(4-x^2)_+}/(2\pi)$ in the limit that the matrix size N goes to infinity.

This result, which is called Wigner's semicircle law, makes it possible to compute the part of the eigenvalues contained in a fixed interval in the limit N to infinity. In more recent results, this value is calculated for a variable interval size as well. However, the interval size is not allowed to decrease too fast to guarantee that the interval contains at least some eigenvalues for each matrix size N. Such statements are called local semicircle laws.

Different versions of local semicircle laws were proved by Erdős and coworkers during their work on the Wigner-Dyson-Gaudin-Mehta conjecture which is formulated in Conjecture 1.2.1 and Conjecture 1.2.2 in [13]. It asserts that the local statistics of the eigenvalues of a random matrix are independent of the distribution of the entries in the limit N to infinity. This independence of the actual distribution is called universality. The proof of this conjecture by Erdős and coworkers in [5, 8, 11] is build upon establishing a local semicircle law in the first step. Their solution is reviewed in [4, 9]. Their most general version of a local semicircle law is verified in [7]. Besides technical assumptions about the regularity of the entries the most important requirement is their interdependence structure. The matrix $H = (h_{xy})_{x,y}$ is supposed to be complex Hermitian (or real symmetric) i.e. $h_{xy} = \bar{h}_{yx}$ for all x and y such that $(h_{xy})_{x \le y}$ forms an independent family of random variables. This means that the entries are independent up to the hermiticity constraint.

Most of the work in random matrix theory starts with this independence assumption. Therefore, it is an interesting aim to study such questions without this assumption or with a weaker substitute. In the present thesis, we determine the limiting distribution for a class of random matrices obeying a different interdependence structure. We introduce further dependences among the entries by adding an additional symmetry to the hermiticity assumption. More precisely, we suppose that the random variables $(h_{xy}; x, y = -N/2, \dots, N/2)$ satisfy the symmetry constraint

$$h_{xy} = \bar{h}_{yx} = h_{-y,-x} = \bar{h}_{-x,-y}$$

for all x and y and are independent besides these relations. We call this symmetry *fourfold symmetry.*

The local semicircle law in Theorem 4.1, which is the main result of this thesis, has the same formulation as Theorem 5.1 in [7] which means that the same estimates hold in the current situation as well. However, several parts of the proof given there have to be adapted to the fourfold symmetry. In particular, it is necessary to transfer the fluctuation averaging which is a key tool for the approach which is pursued in [7] and the present work and which is based on Stieltjes transforms, resolvents and large deviation bounds.

We conclude this chapter with an overview of the structure of the present thesis. In the second chapter, we describe the assumptions about the random matrices necessary for our version of the local semicircle law, in particular, the fourfold symmetry. Moreover, we present the tools we use to establish this result. We prove in chapter 3 that the Fourier transform of a Gaussian orthogonal ensemble is an example of a random matrix fulfilling our assumptions. The following chapter is devoted to the proof of the local semicircle law. In the next two chapters, we verify two tools used in the proofs of the previous chapter. We establish the resolvent identities in the first of these chapters and the second chapter contains a proof of the fluctuation averaging in the present situation which is a key ingredient in our approach to the semicircle law. Finally, we finish the thesis by a collection of some well-known auxiliary results in the last chapter.

Acknowledgement: Several people supported me during the preparation of the present thesis. Herewith, I would like to thank them.

I am grateful to Prof. László Erdős. He supervised me despite our different workplaces and proposed this interesting topic to me. Continually, he replied to my emails with an amazing velocity. Moreover, I appreciate the advices of Alessandro Michelangeli concerning various decisions I had to make in the last time. Finally, I thank my parents for their constant support, advice and encouragement during my entire studies.

2 General Tools

This chapter is devoted to the setting of the local semicircle law and the tools used for its proof in chapter 4. In the first section, we explain the symmetry of the random matrix and the technical assumptions about its entries and introduce some notation.

The rest of the chapter consists of the presentation of the tools for the proof of the local semicircle law. In the second section, we introduce the stochastic domination, a relation which is used to bound the error terms, and spectral domains which are families of special subsets of the complex plane. The third section is devoted to Large Deviation Bounds with respect to the stochastic domination. Then we give some relations which connect resolvents and resolvents of minors. The fifth section contains some notation for partial expectations used in our context. The last section deals with the Fluctuation Averaging which is an important ingredient of the proof of the local semicircle law.

2.1 Setting: Fourfold Symmetry

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. For $N \in \mathbb{N}$ let $(h_{xy}^{(N)}; y = -N/2, \dots, 0,\ x = -y, \dots, y)$ be a familiy of independent complex valued random variables such that $h_{xx}^{(N)}$ is real valued for all x and $h_{xy}^{(N)}$ is centered, i.e. $\mathbb{E}[h_{xy}^{(N)}] = 0$ for $y = -N/2, \dots, 0$ and $x = -y, \dots, y$. Most of the time, the dependence on N will be suppressed in our notation. We set

$$\begin{aligned} h_{xy} &:= \bar{h}_{yx}, \qquad \text{for } x = -N/2, \dots, 0,\ y = x, \dots, -x, \\ \text{and afterwards } h_{xy} &:= h_{-y,-x}, \text{ for } x = -N/2, \dots, N/2, y = -x, \dots, N/2. \end{aligned}$$

Then the matrix $H_{(N)} = (h_{xy}^{(N)})_{x,y=-N/2}^{N/2}$ fulfills the relations

$$h_{xy} = \bar{h}_{yx} = h_{-y,-x} = \bar{h}_{-x,-y} \tag{2.1}$$

for all x, y. The following matrix illustrates these relations

$$\begin{array}{cc} & \begin{array}{cccccc} -N/2 & x & y & -y & -x & N/2 \end{array} \\ \begin{array}{c} -N/2 \\ x \\ y \\ -y \\ -x \\ N/2 \end{array} & \left(\begin{array}{ccc|ccc} & & & & & \\ & & h_{xy} & & & \\ & \bar{h}_{xy} & & & & \\ \hline & & & & h_{xy} & \\ & & & \bar{h}_{xy} & & \\ & & & & & \end{array}\right) \end{array}.$$

We call these dependences between the entries of H *Fourfold Symmetry.*

Correspondingly, the symmetry of $(h_{xy})_{x,y}$ with a family of independent complex valued random variables $(h_{xy})_{x\leq y}$ with real valued h_{xx} for all x and $h_{xy} := \bar{h}_{yx}$ for all $x > y$ is called *Twofold Symmetry*. The latter symmetry was studied in [7]. In both cases, H is a Hermitian matrix and has therefore only real eigenvalues.

Furthermore, we have to assume that $\mathbb{E}h_{xy}^2 = 0$ for all $x \neq y$ and all N for technical reasons. Note that this requirement was not necessary in [7]. For example, this assumption is fulfilled if $\operatorname{Im} h_{xy}$ and $\operatorname{Re} h_{xy}$ have the same distribution for $x \neq y$.

Our aim is to determine the limit of the empirical spectral measure

$$\mu_N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^{(N)}}$$

for $N \to \infty$ (in an appropriate sense) where $(\lambda_i^{(N)})_{i=1}^N$ are the eigenvalues of $H_{(N)}$ (counted with multiplicity). Since all eigenvalues of H are real μ_N defines a probability distribution on $\mathbb{R}$.

To expect a convergence of these measures, it is necessary to make some further mostly technical assumptions. We define the N-dependent quantity $s_{xy} := \mathbb{E}|h_{xy}|^2$ and assume that $\sum_y s_{xy} = 1$ for all x and for all N, i.e. the symmetric matrix $S = (s_{xy})_{x,y=-N/2}^{N/2}$ is stochastic. Moreover, we introduce the normalized random variables $\zeta_{xy} := s_{xy}^{-1/2} h_{xy}$ (If $s_{xy} = 0$ we set $\zeta_{xy} \sim \mathcal{N}(0,1)$.) which fulfill $\mathbb{E}\zeta_{xy} = 0$ and $\mathbb{E}|\zeta_{xy}|^2 = 1$ and we assume that there are constants μ_p for $p \in \mathbb{N}$ such that

$$\mathbb{E}|\zeta_{xy}|^p \leq \mu_p \tag{2.2}$$

for all x, y and all N.

We set $M := (\max_{x,y} s_{xy})^{-1}$ and assume that this N-dependent parameter satisfies

$$N^{\delta} \leq M \leq N \tag{2.3}$$

for some $\delta > 0$. Note that the first estimate is an assumption on the random variables whereas the bound $M \leq N$ follows from $s_{xy} \leq M^{-1}$ for all x, y and $N^{-1} \leq s_{xy}$ for at least one pair x, y.

We will see in chapter 3 that the Fourier transform of a Gaussian orthogonal ensemble fulfills all assumptions, i.e. it is an example of a random matrix which our results can be applied to.

It is helpful to consider the Stieltjes transform of the empirical spectral measure for determining its limit and establishing the convergence. Recall that the Stieltjes transform $S(\nu)$ of a finite measure ν on $\mathbb{R}$ is defined as

$$S(\nu)(z) = \int_{\mathbb{R}} \frac{1}{x-z} \mathrm{d}\nu(x)$$

for $z \in \mathbb{C}\backslash\mathbb{R}$. For $z \in \mathbb{C}\backslash\mathbb{R}$ we compute

$$m_N(z) := S(\mu_N)(z) = \int_{\mathbb{R}} \frac{1}{x-z}\mathrm{d}\mu_N = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i - z} = \frac{1}{N}\sum_{i=1}^{N}\langle e_i, (H-z)^{-1}e_i\rangle = \frac{1}{N}\mathrm{tr}(H-z)^{-1}$$

where $(e_i)_{i=1}^N$ is an orthonormal eigenbasis of H corresponding to the eigenvalues $(\lambda_i)_{i=1}^N$. Because the right hand side $\mathrm{tr}(H-z)^{-1}/N$ does not explicitely contain the eigenvalues this computation suggests to examine the Stieltjes transform of μ_N and thus the resolvent $(H-z)^{-1}$.

2.2 Stochastic Domination and Spectral Domains

We will prove that $m_N(z)$ converges to the Stieltjes transform $m(z)$ of the limiting distribution for $N \to \infty$ by estimating the error term $|m_N(z) - m(z)|$ from above in an appropriate sense. This relation which implements a notion of an event with asymptotically high probability is introduced in the following definition.

Definition 2.1 (Stochastic Domination). *Let $X = (X^{(N)}(u); u \in U^{(N)}, N \in \mathbb{N})$ and $Y = (Y^{(N)}(u); u \in U^{(N)}, N \in \mathbb{N})$ be two families of nonnegative random variables for a possibly N-dependent parameter set $U^{(N)}$. We say that X is* stochastically dominated *by Y, uniformly in u, if for all $\varepsilon > 0$ and $D > 0$ there is a $N_0(\varepsilon, D) \in \mathbb{N}$ such that*

$$\sup_{u \in U^{(N)}} \mathbb{P}\left[X^{(N)}(u) > N^{\varepsilon} Y^{(N)}(u)\right] \leq N^{-D}$$

for all $N \geq N_0$. In this case, we use the notation $X \prec Y$. If X is a family consisting of complex valued random variables and $|X| \prec Y$ then we write $X \in O_{\prec}(Y)$.

Note that it suffices to check the estimate in the previous definition for small $\varepsilon > 0$ and for large $D > 0$. Suppose that the estimate holds for $\tilde{\varepsilon} \in (0, \varepsilon)$, $\tilde{D} \in (D, \infty)$ and all $N \geq N_0$ then

$$\mathbb{P}\left[X^{(N)} > N^{\varepsilon} Y^{(N)}\right] \leq \mathbb{P}\left[X^{(N)} > N^{\tilde{\varepsilon}} Y^{(N)}\right] \leq N^{-\tilde{D}} \leq N^{-D}$$

for all $N \geq N_0$.

In this thesis the stochastic domination will be used to estimate functions of the random variables $h_{xy}^{(N)}$. These bounds are always understood to be uniform in the present parameters except the parameter δ in (2.3) and the constants μ_p in (2.2). At the end of this section, we will explain the usual form of the parameter set in our situation.

The following two Lemmas give simple examples of estimates with respect to the relation defined in the previous definition which are useful for our purposes.

Lemma 2.2. *We have $h_{xy} \prec s_{xy}^{1/2} \prec M^{-1/2}$ uniformly in x and y.*

Proof. Fix $\varepsilon > 0$ and $D > 0$. Choose $p \in \mathbb{N}$ such that $p\varepsilon > D$. Then there is $N_0 \in \mathbb{N}$ with $\mu_p N^{-p\varepsilon} \leq N^{-D}$ for all $N \geq N_0$ and we have

$$\mathbb{P}\left(|h_{xy}| > N^{\varepsilon} s_{xy}^{1/2}\right) \leq \mathbb{P}\left(|\zeta_{xy}| > N^{\varepsilon}\right) \leq N^{-p\varepsilon} \mathbb{E}[|\zeta_{xy}|^p] \leq \mu_p N^{-p\varepsilon} \leq N^{-D}$$

for all $N \geq N_0$. In the second step, equality holds if $s_{xy} \neq 0$. Otherwise, the left hand side is zero. In the third step, Chebyshev's inequality (7.1) was applied. □

Lemma 2.3. *If $\Xi^{(N)}$ is a family of events with asymptotically very high probability, i.e. for every $D > 0$ exists $N_0 \in \mathbb{N}$ such that $\mathbb{P}(\Xi^c) \leq N^{-D}$ for $N \geq N_0$ then the indicator function $\mathbf{1}(\Xi)$ of Ξ fulfills $1 - \mathbf{1}(\Xi) \prec 0$.*

Proof. Fix $\varepsilon > 0$ and $D > 0$. Then

$$\mathbb{P}\left(1 - \mathbf{1}(\Xi) > N^{\varepsilon} \cdot 0\right) = \mathbb{P}(\mathbf{1}(\Xi) = 0) = \mathbb{P}(\Xi^c) \leq N^{-D}$$

for all $N \geq N_0$. □

Note that the relation $\prec$ is reflexive and transitive. Moreover, the next Lemma summarizes the behaviour of $\prec$ under arithmetic operations of the random variables.

Lemma 2.4. *(i) Let X be a family of random variables and $c > 0$ then $cX \prec X$.*

(ii) If $X(u,v) \prec Y(u,v)$ uniformly in $u \in U$ and $v \in V$ and $|V| \leq N^c$ for some $c > 0$ then

$$\sum_{v\in V} X(u,v) \prec \sum_{v\in V} Y(u,v)$$

uniformly in $u \in U$.

(iii) For $X_1(u) \prec Y_1(u)$ uniformly in u and $X_2(u) \prec Y_2(u)$ uniformly in u we have $X_1(u)X_2(u) \prec Y_1(u)Y_2(u)$ uniformly in u.

(iv) If $X \prec YN^\varepsilon$ for all $\varepsilon > 0$ then $X \prec Y$.

(v) If $X \prec Y$ and $\alpha > 0$ then $X^\alpha \prec Y^\alpha$.

Proof. To prove the first part, let $\varepsilon > 0$ and $D > 0$. Choose $N_0 \in \mathbb{N}$ such that $c < N^\varepsilon$ for all $N \geq N_0$. Then $\mathbb{P}(cX > N^\varepsilon X) \leq \mathbb{P}(X > X) = 0 \leq N^{-D}$ for all $N \geq N_0$.

Seondly, fix $\varepsilon > 0$ and $D > 0$. Then there is $N_0 \in \mathbb{N}$ such that $\mathbb{P}(X(u,v) > N^\varepsilon Y(u,v)) \leq N^{-C-D}$ for all $u \in U$, $v \in V$ and $N \geq N_0$. Fix $u \in U$. Since $\sum_{v\in V} X(u,v) > \sum_{v \in V} N^\varepsilon Y(u,v)$ implies the existence of $v \in V$ such that $X(u,v) > N^\varepsilon Y(u,v)$ we get

$$\begin{aligned}\mathbb{P}\left(\sum_{v\in V} X(u,v) > \sum_{v\in V} N^\varepsilon Y(u,v)\right) &\leq \mathbb{P}\left(\exists v \in V \colon X(u,v) > N^\varepsilon Y(u,v)\right) \\ &\leq \sum_{v\in V} \mathbb{P}\left(X(u,v) > N^\varepsilon Y(u,v)\right) \leq N^{-D}\end{aligned}$$

for all $N \geq N_0$. As the upper bound is independent of u the statement of (ii) follows.

For fixed $\varepsilon > 0$ and $D > 0$ we find $N_0 \in \mathbb{N}$ such that $\mathbb{P}\left(X_1(u) > N^{\varepsilon/2} Y_1(u)\right) \leq N^{-2D}$, $\mathbb{P}\left(X_2(u) > N^{\varepsilon/2} Y_2(u)\right) \leq N^{-2D}$ and $N^{-D} \leq 1/2$ for all $N \geq N_0$ for all $u \in U$. Thus, we have for $u \in U$ the estimate

$$\begin{aligned}\mathbb{P}\left(X_1(u)X_2(u) > N^\varepsilon Y_1(u)Y_2(u)\right) &\leq \mathbb{P}\left(X_1(u) > N^{\varepsilon/2}Y_1(u) \text{ or } X_2(u) > N^{\varepsilon/2} Y_2(u)\right) \\ &\leq \mathbb{P}\left(X_1(u) > N^{\varepsilon/2}Y_1(u)\right) + \mathbb{P}\left(X_2(u) > N^{\varepsilon/2}Y_2(u)\right) \leq 2N^{-2D} \leq N^{-D}\end{aligned}$$

for all $N \geq N_0$ and all $u \in U$. This establishes (iii).

The claim in (iv) follows from the identity $\mathbb{P}(X > N^\varepsilon Y) = \mathbb{P}(X > N^{\varepsilon/2}N^{\varepsilon/2}Y)$ and the assumption applied to $\varepsilon/2$.

The identity $\mathbb{P}(X^\alpha > N^\varepsilon Y^\alpha) = \mathbb{P}(X > N^{\varepsilon/\alpha}Y)$ yields claim (v). ☐

When studying the distance $|m(z) - m_N(z)|$ for fixed N where m_N is the Stieltjes transform of μ_N and m the Stieltjes transform of the limiting distribution we will suppose that $\eta = \operatorname{Im} z \geq M^{-1}$. This assumption is necessary since small η drastically weakens our estimates. To have some unified notation for this requirement we introduce the next notion.

Definition 2.5. *An N-dependent family*

$$\mathbf{D} \equiv \mathbf{D}^{(N)} \subset \{z = E + \mathrm{i}\,\eta \in \mathbb{C}; E \in [-10, 10], M^{-1} \leq \eta \leq 10\}$$

is called a spectral domain*.*

Since we consider $M \equiv M_N$ as a function of N we label the elements of such a family by N as well.

The stochastic domination is usually applied to estimate functions of the entries of H or function of the resolvent entries. Thus, the parameter set $U^{(N)}$ in the definition of the stochastic domination will consist of pairs (u, i) where u is contained in a spectral domain and i is an element of some index set. Our estimates will always be uniform in these two parameters but may depend on the values of δ in (2.3) and the constants μ_p in (2.2).

2.3 Large Deviation Bound

This section provides a particularly useful result to get upper bounds on linear combinations of random variables with respect to the stochastic domination defined in the previous section.

Theorem 2.6 (Large Deviation Bounds)**.** *Let $(X_i^{(N)} : i \in \{1, \ldots, N\}, N \in \mathbb{N})$ and $(Y_i^{(N)} : i \in \{1, \ldots, N\}, N \in \mathbb{N})$ be independent families of random variables and let $(a_{ij}^{(N)} : i, j \in \{1, \ldots, N\}, N \in \mathbb{N})$ and $(b_i^{(N)} : i \in \{1, \ldots, N\}, N \in \mathbb{N})$ be constants. If all entries $X_i^{(N)}$ are independent and all entries $Y_i^{(N)}$ are independent and there are constants μ_p such that*

$$\mathbb{E}X_i^{(N)} = \mathbb{E}Y_i^{(N)} = 0, \quad \mathbb{E}|X_i^{(N)}|^2 = \mathbb{E}|Y_i^{(N)}|^2 = 1, \quad \mathbb{E}|X_i^{(N)}|^p, \mathbb{E}|Y_i^{(N)}|^p \leq \mu_p \tag{2.4}$$

for all $i \in \{1, \ldots, N\}$ and $N, p \in \mathbb{N}$ then

$$\Big|\sum_i b_i X_i\Big| \prec \Big(\sum_i |b_i|^2\Big)^{1/2}, \tag{2.5}$$

$$\Big|\sum_{i,j} a_{ij} X_i Y_j\Big| \prec \Big(\sum_{i,j} |a_{ij}|^2\Big)^{1/2}, \tag{2.6}$$

$$\Big|\sum_{i \neq j} a_{ij} X_i X_j\Big| \prec \Big(\sum_{i \neq j} |a_{ij}|^2\Big)^{1/2}. \tag{2.7}$$

If the coefficients $a_{ij}^{(N)}$ and $b_i^{(N)}$ depend on an additional parameter u, then all these estimates are uniform in u.

The intuitive idea behind the Large Deviation Bound is that these linear combinations asymptotically behave like the square root of their variances with very high probability. For example

$$\mathbb{E}\Big|\sum_{i,j} a_{ij} X_i Y_j\Big|^2 = \sum_{i,j,k,l} a_{ij}\overline{a_{kl}} \mathbb{E} X_i Y_j \overline{X_k Y_l} = \sum_{i,j} |a_{ij}|^2$$

where we used the assumptions $\mathbb{E}X_i = \mathbb{E}Y_j = 0$ and $\mathbb{E}|X_i|^2 = \mathbb{E}|Y_j|^2 = 1$ in (2.4) in the last step.

The actual proof relies on estimating high moments of the linear combinations by the Marcinkiewicz-Zygmund inequality. This is done in Lemma B.2, B.3 and B.4 in [6]. Using Chebyshev's inequality (7.1) the estimates in Theorem 2.6 can be directly deduced from these bounds on the moments.

2.4 Resolvent Identities

Computing the Stieltjes transform of the empirical spectral measure indicates the importance of resolvents for our approach. Therefore, we introduce a notation for them. For a Hermitian matrix H and $z \in \mathbb{C}\backslash\mathbb{R}$ we denote the resolvent of $H - z$ by

$$G(z) := (H - z)^{-1}.$$

Moreover, we will need resolvents of minors of H which are denoted by the following symbol.

Definition 2.7. *Let H be a Hermitian matrix. For $\mathbb{T} \subset \{-N/2, \ldots, N/2\}$ we define the $N \times N$ matrix $H^{(\mathbb{T})}$ through*

$$(H^{(\mathbb{T})})_{ij} := \mathbf{1}(i \notin \mathbb{T})\mathbf{1}(j \notin \mathbb{T})h_{ij}.$$

For $z \in \mathbb{C}\backslash\mathbb{R}$ let

$$G_{ij}^{(\mathbb{T})}(z) := (H^{(\mathbb{T})} - z)_{ij}^{-1}$$

denote the resolvent of $H^{(\mathbb{T})}$ in z. We set

$$\sum_i^{(\mathbb{T})} := \sum_{i;i\notin\mathbb{T}}.$$

Note that the inverse $G_{ij}^{(\mathbb{T})}(z)$ always exists since $H^{(\mathbb{T})}$ is again a Hermitian matrix. To simplify the notation we will sometimes suppress that the resolvent $G^{(\mathbb{T})}$ depends on the complex number z. By abuse of notation we also write $H^{(a_1,\ldots,a_n)}$ for $H^{(\mathbb{T})}$ if $\mathbb{T} = \{a_1, \ldots, a_n\}$ and $H^{(\mathbb{T}b_1,\ldots,b_m)}$ for $H^{(\mathbb{T}\cup\{b_1,\ldots,b_m\})}$. This short notation is also applied to the resolvents $G^{(\mathbb{T})}$.

The following relation is the starting point to establish a self-consistent equation which will be an important tool throughout the whole proof of the semicircle law.

Lemma 2.8 (Schur's Complement Formula)**.** *For a Hermitian matrix $(h_{xy})_{x,y=-N/2}^{N/2}$, we have the relation*

$$\frac{1}{G_{xx}} = h_{xx} - z - \sum_{a,b}^{(x)} h_{xa}G_{ab}^{(x)}h_{bx}. \tag{2.8}$$

Due to the fourfold symmetry (2.1) the entries $G_{ab}^{(x)}$ are not independent of the entries h_{xa} of H which means that the Large Deviation Bounds are not applicable. However, the following resolvent identities can be used to replace $G_{ab}^{(x)}$ by $G_{ab}^{(x,-x)}$ at the cost of an error term.

Lemma 2.9 (Resolvent Identities)**.** *For a Hermitian matrix $H = (h_{ij})_{i,j=-N/2}^{N/2}$ and $\mathbb{T} \subset \{-N/2, \ldots, N/2\}$ the following statements hold:*
If $i, j, k \notin \mathbb{T}$ and $i, j \neq k$ then

$$G_{ij}^{(\mathbb{T})} = G_{ij}^{(\mathbb{T}k)} + \frac{G_{ik}^{(\mathbb{T})}G_{kj}^{(\mathbb{T})}}{G_{kk}^{(\mathbb{T})}}, \quad \frac{1}{G_{ii}^{(\mathbb{T})}} = \frac{1}{G_{ii}^{(\mathbb{T}k)}} - \frac{G_{ik}^{(\mathbb{T})}G_{ki}^{(\mathbb{T})}}{G_{ii}^{(\mathbb{T})}G_{ii}^{(\mathbb{T})}G_{kk}^{(\mathbb{T})}}. \tag{2.9}$$

If $i, j \notin \mathbb{T}$ satisfy $i \neq j$ then

$$G_{ij}^{(\mathbb{T})} = -G_{ii}^{(\mathbb{T})}\sum_k^{(\mathbb{T}i)} h_{ik}G_{kj}^{(\mathbb{T}i)} = -G_{jj}^{(\mathbb{T})}\sum_k^{(\mathbb{T}j)} G_{ik}^{(\mathbb{T}j)}h_{kj}. \tag{2.10}$$

The proofs of (2.8) and Lemma 2.9 are contained in chapter 5.

A further important tool for estimating the error terms is the relation (2.11) which is sometimes called *Ward identity.*

Lemma 2.10. *For a Hermitian matrix H we have*

$$\sum_l |G_{kl}^{(\mathbb{T})}(E+\mathrm{i}\,\eta)|^2 = \frac{1}{\eta}\mathrm{Im}\, G_{kk}^{(\mathbb{T})}(E+\mathrm{i}\,\eta) \tag{2.11}$$

for $E \in \mathbb{R}$ and $\eta > 0$.

Proof. We set $z := E + \mathrm{i}\,\eta$. Then we have

$$\begin{aligned}\sum_l |G_{kl}^{(\mathbb{T})}(z)|^2 =& \sum_l \overline{G_{kl}^{(\mathbb{T})}(z)} G_{kl}^{(\mathbb{T})}(z) = \left(G^{(\mathbb{T})^*} G^{(\mathbb{T})}\right)_{kk} = |(H^{(\mathbb{T})} - z)^{-1}|^2_{kk} \\ =& \frac{1}{\eta}\left(\mathrm{Im}\,(H^{(\mathbb{T})} - z)^{-1}\right)_{kk} = \frac{1}{\eta}\mathrm{Im}\,(H^{(\mathbb{T})} - z)^{-1}_{kk}.\end{aligned}$$

Here, we applied in the fourth step the functional calculus for selfadjoint operators and the identity

$$\frac{1}{|x-z|^2} = \frac{1}{\eta}\mathrm{Im}\,(x-z)^{-1}$$

for $x \in \mathbb{R}$ and in the last step the definition of the imaginary part of a matrix. □

The following Lemma gives a trivial nevertheless useful bound on the modulus of resolvent entries and a nice connection between resolvent entries and spectral domains.

Lemma 2.11. *For $E \in \mathbb{R}$ and $\eta > 0$ we have*

$$|G_{ij}^{(\mathbb{T})}(E+\mathrm{i}\,\eta)| \le \eta^{-1} \tag{2.12}$$

for all i, j and $\mathbb{T}$. In particular, if $\mathbf{D}$ is a spectral domain we have

$$|G_{ij}^{(\mathbb{T})}(z)| \le M \tag{2.13}$$

for all $z \in \mathbf{D}$.

Proof. Let $(e_i)_{i=1}^N$ be the canonical orthonormal basis of $\mathbb{C}^N$. Let $z \in \mathbb{C}$ such that $\eta := \mathrm{Im}\, z > 0$. Applying the functional calculus yields

$$|G_{ij}^{(\mathbb{T})}(z)| \le \|G^{(\mathbb{T})}(z)e_j\|_2 \le \|G^{(\mathbb{T})}(z)\|_{\ell^2\to\ell^2} = \mathrm{dist}(z, \sigma(H))^{-1} \le \eta^{-1}$$

since $\sigma(H) \subset \mathbb{R}$ for a self-adjoint matrix H. The definition of a spectral domain $\eta \ge M^{-1}$ for $E + \mathrm{i}\,\eta \in \mathbf{D}$ implies the second estimate. □

2.5 Partial Expectation

For the partial expectation with respect to the σ-algebra generated by $H^{(x,-x)}$ we introduce the following notation.

Definition 2.12 (Partial Expectation)**.** *Let X be an integrable random variable. For $x \in \{-N/2, \dots, N/2\}$ we define the random variables $\mathbb{E}_x X$ and $\mathbb{F}_x X$ through*

$$\mathbb{E}_x X := \mathbb{E}[X | H^{(x,-x)}], \quad \mathbb{F}_x X := X - \mathbb{E}_x X.$$

The random variable $\mathbb{E}_x X$ is called the partial expectation *of X with respect to x.*

The symbols $\mathbb{E}_x$ and $\mathbb{F}_x$ are the analogues of P_i and Q_i in [7]. There, the idea is that P_i removes the information contained in column i and row i of H. Because of the twofold symmetry, i.e. the hermiticity column i and row i contain the same information. This is also true for the fourfold symmetry. However, due the fourfold symmetry column x and column $-x$ store the same information. By hermiticity of H this information is also included in row x and row $-x$. Therefore, excluding columns x and $-x$ and rows x and $-x$ yields the correct counterpart of P_i.

Definition 2.13 (Independence)**.** *We say that the integrable random variable X is* independent *of* $\mathbb{T} \subset \{-N/2, \dots, N/2\}$ *if* $X = \mathbb{E}_x X$ *for all* $x \in \mathbb{T}$.

Observe that if Y is independent of x we have that $\mathbb{F}_x(X)Y = XY - \mathbb{E}_x(X\mathbb{E}_x Y) = \mathbb{F}_x(XY)$. In particular,

$$\mathbb{E}\mathbb{F}_x(X)Y = \mathbb{E}\mathbb{F}_x(XY) = \mathbb{E}(XY) - \mathbb{E}\mathbb{E}_x(XY) = 0 \tag{2.14}$$

where (7.3) was used in the third step.

2.6 Fluctuation Averaging

Let $\mathbf{D}$ be a spectral domain and Ψ a deterministic (possibly z-dependent) control parameter which satisfies

$$M^{-1/2} \leq \Psi \leq M^{-c} \tag{2.15}$$

for all $z \in \mathbf{D}$ and for some $c > 0$.

The aim of the fluctuation averaging is to estimate linear combinations of the form $\sum_k t_{ik} X_k$ with special random variables X_k and a family of complex weights $T = (t_{ik})$ that satisfy

$$0 \leq |t_{ik}| \leq M^{-1}, \quad \sum_k |t_{ik}| \leq 1. \tag{2.16}$$

Note that the family T may be N-dependent. Examples of such weights are given by $t_{ik} = s_{ik}$ or $t_{ik} = N^{-1}$. It will be important that T computes with S in these two cases. We will only study the random variables $X_k = \mathbb{F}_k[(G_{kk})^{-1}]$, $X_k = \mathbb{F}_k G_{kk}$ and $X_k = G_{kk} - m$.

We define

$$\Gamma(z) := \|(1 - m^2(z)S)^{-1}\|_{\ell^\infty \to \ell^\infty} \tag{2.17}$$

for $z \in \mathbf{D}$ with a spectral domain $\mathbf{D}$. The parameter Γ will turn out to be an important control parameter and it appears in an upper bound in the next result.

Theorem 2.14 (Fluctuation Averaging)**.** *Let* $\mathbf{D}$ *be a spectral domain,* Ψ *a deterministic control parameter satisfying* (2.15) *and* $T = (t_{ik})$ *a weight satisfying* (2.16)*. If* $\Lambda \prec \Psi$ *then*

$$\left| \sum_k t_{ik} \mathbb{F}_k \frac{1}{G_{kk}} \right| \prec \Psi^2, \quad \left| \sum_k t_{ik} \mathbb{F}_k G_{kk} \right| \prec \Psi^2 \tag{2.18}$$

uniformly in i and $z \in \mathbf{D}$. If $\Lambda \prec \Psi$ and T commutes with S then for $v_k = G_{kk} - m$ we have

$$\left| \sum_k t_{ik} v_k \right| \prec \Gamma\Psi^2 \tag{2.19}$$

uniformly in i and $z \in \mathbf{D}$.

For the first estimate in (2.18), there is the following stronger bound:

Theorem 2.15. *Let* $\mathbf{D}$ *be a spectral domain,* Ψ *and* Ψ_o *deterministic control parameters satisfying* (2.15) *and* $T = (t_{ik})$ *a weight satisfying* (2.16)*. If* $\Lambda \prec \Psi$ *and* $\Lambda_o \prec \Psi_o$ *then*

$$\left| \sum_k t_{ik} \mathbb{F}_k \frac{1}{G_{kk}} \right| \prec \Psi_o^2 \tag{2.20}$$

uniformly in i *and* $z \in \mathbf{D}$.

Theorem 2.14 and Theorem 2.15 are the counterparts of Theorem 4.6 and Theorem 4.7 in [7] for the fourfold symmetry.

Intuitively, the stronger bounds on $\sum_k t_{ik} X_k$ compared to the trivial bound coming from the bounds on the random variables X_k arise as cancellations occur in these linear combinations. The origin of these cancellations is the osciallatory behaviour of the random variables X_k. The actual proof relies on a careful estimate of high moments of $\sum_k t_{ik} X_k$ by exploiting the structure of X_k via the resolvent identities. This is explained in more detail in chapter 6.

As its formulation reveals the fluctuation averaging is a tool to improve certain already established bounds. For this purpose it is applied in the end of the proof of our main result to realize the final bounds.

3 Fourier Transform of Random Matrices

As already mentioned in the introduction one motivation to study random matrices $(h_{xy})_{x,y}$ with the fourfold symmetry (2.1) and the constraint $\mathbb{E}h_{xy}^2 = 0$ is the fact that the Fourier tranform of a Gaussian orthogonal ensemble has these properties. In this chapter we state some basic definitions and prove the claims in the previous sentence.

Besides the definition of a Gaussian orthogonal ensemble we also define Gaussian unitary ensembles which are the complex valued counterparts. We use the formulation from [12].

Definition 3.1 (Gaussian Ensembles)**.** *Let $(X_i)_i$, $(Y_{ij})_{i<j}$ and $(Z_{ij})_{i<j}$ be independent families of $N(0,1)$-distributed independent random variables.*

(i) We consider the (symmetric) $N \times N$ matrix $H_N = (h_{ij}^{(N)})_{i,j}$ with

$$\begin{aligned} h_{ij}^{(N)} = h_{ji}^{(N)} &= \frac{1}{\sqrt{2N}} Y_{ij} \qquad \text{for } i < j, \\ h_{jj}^{(N)} &= \frac{1}{\sqrt{N}} X_j \end{aligned}$$

for $N \in \mathbb{N}$. Then H_N is said to be an element of GOE$_N$ and the sequence $(H_N)_{N\in\mathbb{N}}$ is called a Gaussian orthogonal ensemble.

(ii) For $N \in \mathbb{N}$ we define the (Hermitian) $N \times N$ matrix $H_N = (h_{ij}^{(N)})_{i,j}$ by

$$\begin{aligned} h_{ij}^{(N)} = \overline{h_{ji}^{(N)}} &= \frac{1}{\sqrt{2N}} (Y_{ij} + iZ_{ij}) \qquad \text{for } i < j, \\ h_{jj}^{(N)} &= \frac{1}{\sqrt{N}} X_j. \end{aligned}$$

The matrix H_N is said to belong to GUE$_N$ and the sequence $(H_N)_{N\in\mathbb{N}}$ is called a Gaussian unitary ensemble.

The adjective "orthogonal" ("unitary" respectively) in the previous definition comes from the fact that conjugating H_N with an orthogonal (unitary) $N \times N$ matrix O yields again an element in GOE$_N$ (GUE$_N$), i.e. the joint distribution of the entries of H_N is invariant under conjugation with an orthogonal (unitary) matrix. This is proved in [12].

Definition 3.2 (Fourier Transform)**.** *Let $H = (h_{xy})_{x,y}$ be a $N \times N$ matrix. The* Fourier transform *$\hat{H}$ is the $N \times N$ matrix whose entries are given by*

$$\hat{H}_{pq} = \frac{1}{N} \sum_{x,y} h_{xy} \exp\left(-i \frac{2\pi}{N} (px - qy) \right).$$

In the next Lemma we prove that the Fourier transform of a Gaussian orthogonal ensemble fulfills the conditions of Theorem 4.1 i.e. the local semicircle law holds for such random matrices.

Lemma 3.3. *Let H_N be an element of GOE_N for an odd positive integer N and $\hat{H}$ its Fourier transform. Then the entries $\hat{H}_{pq}$ and $\hat{H}_{rs}$ are independent if and only if*

$$(p,q) \notin \{(r,s),(s,r),(-r,-s),(-s,-r)\}.$$

Moreover, we have

$$\hat{H}_{pq} = \overline{\hat{H}_{qp}} = \hat{H}_{-q,-p} = \overline{\hat{H}_{-p,-q}}$$

for all p, q and

$$\mathbb{E}\hat{H}_{pq}^2 = 0$$

for all $p \neq q$.

Proof. We denote the entries of H_N by h_{ij}. To prove the if-part it suffices to establish the second statement which is a direct consequence of the fact that H_N is symmetric:

$$\begin{aligned}
\overline{\hat{H}_{qp}} =& N^{-1}\sum_{x,y} h_{xy}\exp\left(i\frac{2\pi}{N}(qx-py)\right) = N^{-1}\sum_{x,y} h_{yx}\exp\left(-i\frac{2\pi}{N}(py-qx)\right) = \hat{H}_{pq},\\
\hat{H}_{-q,-p} =& N^{-1}\sum_{x,y} h_{xy}\exp\left(-i\frac{2\pi}{N}(-qx+py)\right) = N^{-1}\sum_{x,y} h_{yx}\exp\left(-i\frac{2\pi}{N}(py-qx)\right) = \hat{H}_{pq}.
\end{aligned}$$

Combining these two relations yields the third equality.

Since $\hat{H}_{pq}$ and $\hat{H}_{rs}$ are again jointly normally distributed and $\mathbb{E}\hat{H}_{pq} = \mathbb{E}\hat{H}_{rs} = 0$ it suffices to prove that $\mathbb{E}\hat{H}_{pq}\overline{\hat{H}_{rs}} = 0$ and $\mathbb{E}\hat{H}_{pq}\hat{H}_{rs} = 0$ in order to show that these random variables are independent. Let $(X_j)_j$ and $(Y_{ij})_{i<j}$ be independent families of independent $N(0,1)$-distributed random variables as in the definition of an element of GOE_N. Then we have $\mathbb{E}h_{ij}^2 = (2N)^{-1}$ for $i \neq j$ and $\mathbb{E}h_{ii}^2 = N^{-1}$. Therefore, $\mathbb{E}h_{x_1y_1}h_{x_2y_2} = (2N)^{-1}(\delta_{x_1x_2}\delta_{y_1y_2} + \delta_{x_1y_2}\delta_{y_1x_2})$. We get

$$\begin{aligned}
\mathbb{E}\hat{H}_{pq}\overline{\hat{H}_{rs}} =& \frac{1}{2N^3}\sum_{x,y}\left[\exp\left(-i\frac{2\pi}{N}(px-qy+rx-sy)\right) + \exp\left(-i\frac{2\pi}{N}(px-qy+ry-sx)\right)\right]\\
=& \frac{1}{2N^3}\left(\sum_x \exp\left(-i\frac{2\pi}{N}(p+r)x\right)\right)\left(\sum_y \exp\left(i\frac{2\pi}{N}(q+s)y\right)\right)\\
&+ \frac{1}{2N^3}\left(\sum_x \exp\left(-i\frac{2\pi}{N}(p-s)x\right)\right)\left(\sum_y \exp\left(-i\frac{2\pi}{N}(r-q)y\right)\right).
\end{aligned}$$

Note that for $N = 2k+1$ the summations are supposed to start in $-k$ and end in k and we have

$$\sum_{x=-k}^{k} \exp\left(-i\frac{2\pi}{N}mx\right) = \begin{cases} N, & m = 0\\ 0, & m \neq 0 \end{cases}$$

for $m \in (-N,N) \cap \mathbb{Z}$. Thus, $\mathbb{E}\hat{H}_{pq}\overline{\hat{H}_{rs}} \neq 0$ if and only if $(p,q) \in \{(-r,-s),(s,r)\}$. Since $\mathbb{E}\hat{H}_{pq}\hat{H}_{rs} = \mathbb{E}\hat{H}_{pq}\overline{\hat{H}_{sr}}$ we have that $\mathbb{E}\hat{H}_{pq}\hat{H}_{rs} \neq 0$ if and only if $(p,q) \in \{(-s,-r),(r,s)\}$. In particular, $\mathbb{E}\hat{H}_{pq}^2 = 0$ for $p \neq q$. □

4 The Local Semicircle Law

In this chapter, we show the main result of this thesis, the local semicircle law. We keep the notation and the technical assumptions of section 2.1.

As for the twofold symmetry, the distribution of the empirical spectral measure μ_N in the limit $N \to \infty$ will be the semicircle law, i.e. the measure μ_{sc} with the density $\sqrt{(4-x^2)_+}/(2\pi)$ with respect to the Lebesgue measure on $\mathbb{R}$.

The first section of this chapter contains the precise formulation of the local semicircle law. Moreover, we describe the heuristic ideas underlying the proof presented in the remaining sections of this chapter.

4.1 Main Result and Heuristic Idea of the Proof

Our approach to verify the convergence of the empirical spectral measure μ_N in the limit $N \to \infty$ is based on studying the corresponding Stieltjes transforms. The limit will be the semicircle law whose Stieltjes transform is denoted by

$$m(z) := \frac{1}{2\pi} \int_{[-2,2]} \frac{\sqrt{4-x^2}}{x-z} \mathrm{d}x \tag{4.1}$$

for $z \in \mathbb{C}\backslash\mathbb{R}$. With this definition the complex valued function $m(z)$ is the unique solution of

$$m(z) + \frac{1}{m(z)+z} = 0. \tag{4.2}$$

such that $\operatorname{Im} m(z) > 0$ for $\operatorname{Im} z > 0$ (Compare Remark 4.10).

Recall that the Stieltjes transform m_N of the empirical spectral measure μ_N of $H_{(N)}$ satisfies $m_N(z) = \operatorname{tr} G(z)/N$.

For the definition of the spectral domain $\mathbf{S}$ underlying our estimates we need the positive real numbers:

$$\eta_E := \min\left\{\eta; \frac{1}{M\eta} \leq \min\left\{\frac{M^{-\gamma}}{\Gamma(z)^3}, \frac{M^{-2\gamma}}{\Gamma(z)^4 \operatorname{Im} m(z)}\right\} \text{ for all } z \in [E+\mathrm{i}\,\eta, E+\mathrm{i}\,10]\right\} \tag{4.3}$$

for $\gamma \in (0, 1/2)$ and $E \in \mathbb{R}$. Then, for $\gamma \in (0, 1/2)$ the spectral domain $\mathbf{S}$ is defined as

$$\mathbf{S} = \mathbf{S}^{(N)}(\gamma) := \{E + \mathrm{i}\,\eta; |E| \leq 10, \eta_E \leq \eta \leq 10\} \tag{4.4}$$

which is indeed a spectral domain by Lemma 4.13.

In the next Theorem which is the main result of the present thesis we claim that the deterministic parameter

$$\Pi(z) := \sqrt{\frac{\operatorname{Im} m(z)}{M\eta}} + \frac{1}{M\eta}$$

is an upper bound of the modulus of the off-diagonal resolvent entries and of the deviation of the diagonal resolvent entries from m. The final proof of this statement is contained in section 4.5.

Theorem 4.1 (Local Semicircle Law)**.** *For $\gamma \in (0, 1/2)$ we have*

$$|G_{ij}(z) - \delta_{ij} m(z)| \prec \Pi(z) \tag{4.5}$$

uniformly in i, j and $z \in \mathbf{S}$, as well as

$$|m_N(z) - m(z)| \prec \frac{1}{M\eta} \tag{4.6}$$

uniformly in $z \in \mathbf{S}$.

The bound in (4.6) is one order better than the naive guess motivated by (4.5) and the triangle inequality. The main origin of this stronger estimate is the Fluctuation Averaging.

To have a shorter notation in the following arguments, we introduce the z-dependent stochastic control parameter

$$\Lambda_d(z) := \max_x |G_{xx}(z) - m(z)|, \quad \Lambda_o(z) := \max_{x \neq y} |G_{xy}(z)|, \quad \Lambda(z) := \max\{\Lambda_o(z), \Lambda_d(z)\}$$

measuring the deviation of the diagonal resolvent entries from m, the modulus of the off-diagonal resolvent entries and the maximum of all these quantities. Hence, (4.5) is equivalent to $\Lambda(z) \prec \Pi(z)$ uniformly in $z \in \mathbf{S}$.

Our proof is an adaption of the proof of Theorem 5.1 in [7] which is the analogue of Theorem 4.1 for the twofold symmetry. In each of the following sections, we stress the adjustments which were necessary for the fourfold symmetry or mark sections which are basically unchanged.

Before embarking on the proof of Theorem 4.1 we explain the idea behind the proof. The explanations about the self-consistent equations are inspired by [4]. The central idea is to prove that

$$m_N(z) \approx -\frac{1}{z + m_N(z)}$$

for large N and to conclude that $m_N(z) \approx m(z)$ since $m(z)$ is the unique solution of (4.2) with $\operatorname{Im} m(z) > 0$ for $\operatorname{Im} z > 0$ and (4.2) is stable under small perturbations for $z \neq \pm 2$. More generally, we establish the equation

$$\frac{1}{G_{ii}} = -z + \Upsilon_i - \sum_k s_{ik} G_{kk} \tag{4.7}$$

with the error term Υ_i. The idea how to derive this equation is explained below. In the case $s_{ik} = 1/N$ for all i and k, i.e. the standard Wigner case where all variances are equal, we get the *scalar (or first level) self-consistent equation*

$$m_N(z) = \frac{1}{N} \sum_i G_{ii} \approx -\frac{1}{z + m_N(z)}$$

by inverting and averaging (4.7) and neglecting the error terms. As s_{ik} is not constant in general and we want to estimate Λ, i.e. single terms of the resolvent, we have to deal with (4.7). Thus, m_N approximatively fulfills (4.2) which defines the Stieltjes transform of the semicircle law. This indicates to estimate the error term Υ_i for large N in order to prove the convergence of the Stieltjes transform of the empirical spectral measure to the Stieltjes transform of the semicircle law, i.e. to prove (4.6).

To show (4.5), we need to estimate Λ and in particular Λ_d. This is done by studying the *vector (or second level) self-consistent equation*

$$-\sum_a s_{xa}v_a + \Upsilon_x - m - z = \frac{1}{v_x + m} \tag{4.8}$$

with $v_x := G_{xx} - m$. This relation is a direct consequence of (4.7). We invert this equation and assume that $|-\sum_a s_{xa}v_a + \Upsilon_x|$ is small in order to expand the left-hand side around $-m-z$ up to second order. Thus, we get

$$v_x = m^2\left(\sum_a s_{xa}v_a - \Upsilon_x\right) + \varepsilon_x$$

where we used (4.2) and denoted the error terms by ε_x. We introduce the vectors $\mathbf{v} = (v_x)_x$, $\mathcal{E} = (-m^2\Upsilon_x + \varepsilon_x)_x$ and the matrix $S = (s_{xy})_{x,y}$ to write the previous equation (combined for all x) in the form

$$\mathbf{v} = m^2 S\mathbf{v} + \mathcal{E}.$$

Solving this equation for $\mathbf{v}$ yields

$$\mathbf{v} = (1 - m^2 S)^{-1}\mathcal{E}.$$

Therefore, proving that $\Gamma = \|(1-m^2S)^{-1}\|_{\ell^\infty\to\ell^\infty}$ is bounded and that $\|\mathcal{E}\|_\infty$ is small implies that

$$\Lambda_d = \max_x |G_{xx} - m| = \|\mathbf{v}\|_\infty \le \|(1-m^2S)^{-1}\|_{\ell^\infty\to\ell^\infty}\|\mathcal{E}\|_\infty = \Gamma\|\mathcal{E}\|_\infty$$

is small as well. As these considerations indicate Γ will be an important control parameter in the sequel.

The proof that the off-diagonal terms G_{xy} are small for the fourfold symmetry differs from the proof for the twofold symmetry. In the latter case one uses (2.10) twice to replace G_{xy} by $G_{ab}^{(x,y)}$ and applies the Large Deviation Bound. For the fourfold symmetry the proof that $G_{x,-x}$ is small resembles this strategy but the representation has to be split further to remove the dependences. We use the representation

$$G_{x,-x} = -G_{xx}G_{-x,-x}^{(x)}\left(h_{x,-x} - \sum_k^{(x,-x)} h_{xk}^2 G_{k,-k}^{(x,-x)} - \sum_{k\neq l}^{(x,-x)} h_{xk}G_{k,-l}^{(x,-x)}h_{xl}\right)$$

and estimate the last two terms with the Large Deviation Bounds. The Large Deviation Bounds are only applicable to the second summand because of the assumption $\mathbb{E}h_{xy}^2 = 0$. This technical assumption is also needed at some further instances to make the application of the Large Deviation Bounds possible. To bound G_{xy} from above for $-x \neq y \neq x$ we use (2.9) and (2.10) to get a similar representation where $G^{(x,-x,y,-y)}$ appears instead of $G^{(x,-x)}$. Using that $G_{x,-x}$ is small and applying the Large Deviation Bounds yield that G_{xy} is small as well.

Next, we explain how (4.7) is established which is done in section 4.2. The starting point is Schur's complement formula

$$\frac{1}{G_{xx}} = h_{xx} - z - \sum_{a,b}^{(x)} h_{xa}G_{ab}^{(x)}h_{bx}. \tag{4.9}$$

First, we describe the computation in the case studied in [7]. For this twofold symmetry we have $\mathbb{E}[h_{xa}G_{ab}^{(x)}h_{bx}|H^{(x)}] = G_{ab}^{(x)}\mathbb{E}[h_{xa}h_{bx}] = \delta_{ab}G_{aa}^{(x)}s_{xa}$. After replacing $G_{aa}^{(x)}$ by G_{aa}, which is possible with an error by the resolvent identities, the third term on the right-hand side of (4.9) becomes the desired term $\sum_a s_{xa}G_{aa}$ up to some error terms.

Here, it is important that $H^{(x)}$ is independent of $(h_{xa})_a$ and that $G_{ab}^{(x)}$ is measureable with respect to $\sigma(H^{(x)})$.

However, in the present case h_{xa} is not independent of $H^{(x)}$ for $a \neq -x$ since the $-x$-column and the $-x$-row contain the same information as the x-column and the x-row due to the four-fold symmetry (2.1). Thus, we have to replace $H^{(x)}$ by $H^{(x,-x)}$ and consider the conditional expectation with respect to the σ-algebra generated by the latter. To get $G_{ab}^{(x)}$ measureable with respect to $H^{(x,-x)}$ we have to replace $G_{ab}^{(x)}$ by $G_{ab}^{(x,-x)}$ which is done by using a resolvent identity. Neglecting the error terms we get

$$\mathbb{E}_x \sum_{a,b}^{(x,-x)} h_{xa} G_{ab}^{(x,-x)} h_{bx} \approx \sum_a s_{xa} G_{aa}.$$

Some auxiliary estimates on the resolvent entries are collected in section 4.3. Moreover, we show that under the assumption that Λ is smaller than M^{-c} for some $c > 0$ we get a stronger bound on Λ_o and a bound on Υ_x in terms of Λ. These bounds are proved by using the resolvent identities of Lemma 2.9, the identity (2.11) and the Large Deviation Bounds of Theorem 2.6. We want to demonstrate on a typical expression how these tools are used to estimate the error terms:

$$|\sum_{i \neq j}^{(x,-x)} h_{xi} G_{ij}^{(x,-x)} h_{jx}| \prec \left(\sum_{i \neq j}^{(x,-x)} s_{xi} s_{jx} |G_{ij}^{(x,-x)}|^2 \right)^{1/2} \leq \left(\frac{1}{M\eta} \sum_i^{(x,-x)} s_{xi} \operatorname{Im} G_{ii}^{(x,-x)} \right)^{1/2} \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}.$$

In the first step, we conditioned on $G^{(x,-x)}$, exploited the independence of $(h_{xy})_y$ and $G^{(x,-x)}$ and applied the Large Deviation Bound. Next, we used the trivial bound $s_{xy} \leq M^{-1}$ and identity (2.11). The definition of Λ yields $\operatorname{Im} G_{ii} \leq \operatorname{Im} m + \Lambda$ which holds for $\operatorname{Im} G_{ii}^{(x,-x)}$ with $\leq$ replaced by $\prec$ as well. Finally, we used $\sum_i s_{xi} = 1$.

In section 4.4 we show a preliminary bound on Λ. Therefore, we prove that asymptotically there is an interval which cannot contain the value of Λ i.e. the value of Λ is either smaller than the lower bound or bigger than the upper bound. Then we verify that for large values of $\eta := \operatorname{Im} z$ for $z \in \mathbf{S}$ the parameter Λ lies below the lower bound. Since Λ depends continuously on z the function Λ cannot cross this gap and we conclude that Λ must be smaller than the lower bound for all $z \in \mathbf{S}$. This lower bound is called Ψ_0. The whole idea for establishing this preliminary bound on Λ is illustrated in Figure 4.1. The self-consistent equation (4.8) plays an important role in the proofs of these claims.

Section 4.5 contains the proof of Theorem 4.1 which uses the preparations in the previous sections. First, we observe that a given nicely behaving bound Ψ on Λ can always be improved i.e. $\Lambda \prec F(\Psi)$ for some function F. The key idea in the proof of this result is the fluctuation averaging which implies

$$|\sum_y s_{xy} v_y| \prec \Gamma \Psi^2$$

if $|v_x| \prec \Psi$ for all x. Using this result iteratively with the preliminary bound Ψ_0 on Λ from section 4.4 we get $\Lambda \prec F^{(k)}(\Psi_0)$ which implies the estimate (4.5). Moreover, applying the fluctuation averaging again yields that $\Lambda \prec \Pi$ implies $|N^{-1} \sum_i \Upsilon_i| \prec \Pi^2$. Therefore, we get

$$|m_N(z) - m(z)| = \left| \frac{1}{N} \sum_i v_i \right| \prec |1 - m^2(z)|^{-1} \left| \frac{1}{N} \sum_i \Upsilon_i \right| \prec \frac{1}{M\eta}$$

where the second step is a consequence of the self-consistent equation (4.8) and the third step follows from the above application of the fluctuation averaging and estimates on m. This establishes (4.6).

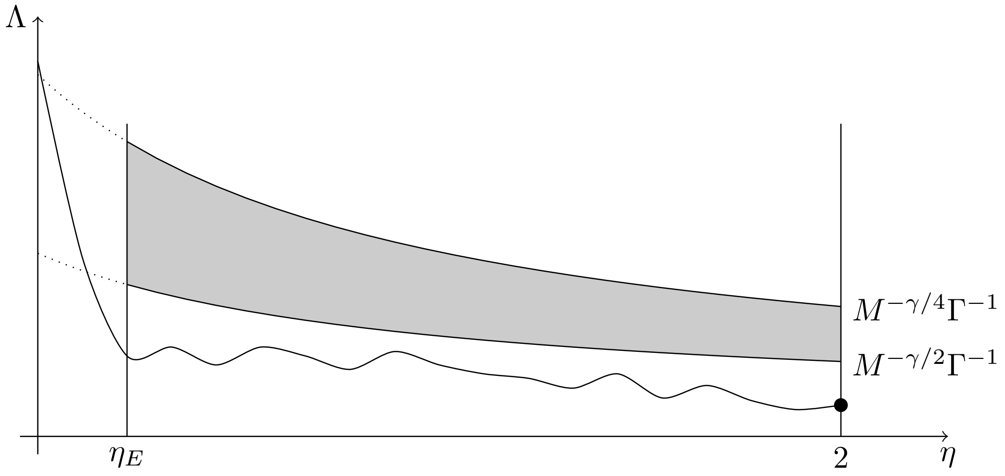

Figure 4.1: This sketch illustrates the proof of the existence of the preliminary bound Ψ_0 on Λ. For fixed E the value $\Lambda(E+\mathrm{i}\,\eta)$ is plotted as a function of η. We verify that $\Lambda(E+\mathrm{i}\,\eta)$ cannot lie in the grey region for $\eta \geq \eta_E$. For $\eta = 2$ we prove that $\Lambda(E+\mathrm{i}\,\eta)$ is smaller than the lower border of this region. Since $\Lambda(E+\mathrm{i}\,\eta)$ is a continuous function of η we conclude that $\Lambda(E+\mathrm{i}\,\eta)$ does not cross this border for $\eta \geq \eta_E$.

The proofs of both estimates crucially rely on the self-consistent equation (4.8). Although its importance was not obvious a priori the multiple usage of this self-consistent equation justifies the efforts to establish it and to control the error term Υ_x.

In section 4.6 we collect some properties of m and several estimates on m and Γ which are used in section 4.3 to 4.5.

4.2 Self-consistent Equation

The goal of this section is to establish a self-consistent equation which we will obtain from manipulating Schur's complement formula.

The fourfold symmetry causes some differences compared to section 5.1 from [7] which are explained at the end of this section.

By applying the resolvent identity (2.9), the last term of (2.8) can be represented in the form

$$\begin{aligned}\sum_{a,b}^{(x)} h_{xa} G_{ab}^{(x)} h_{bx} =& h_{x,-x} G_{-x,-x}^{(x)} h_{-x,x} + \sum_{a}^{(x,-x)} h_{xa} G_{a,-x}^{(x)} h_{-x,x} + \sum_{b}^{(x,-x)} h_{x,-x} G_{-x,b}^{(x)} h_{bx} \\ &+ \sum_{a,b}^{(x,-x)} h_{xa} G_{ab}^{(x,-x)} h_{bx} + \left(G_{-x,-x}^{(x)}\right)^{-1} \sum_{a,b}^{(x,-x)} h_{xa} G_{a,-x}^{(x)} G_{-x,b}^{(x)} h_{bx} \end{aligned} \tag{4.10}$$

Recall our notation $\mathbb{E}_x[X] := \mathbb{E}[X|H^{(x,-x)}]$ and $\mathbb{F}_x[X] := X - \mathbb{E}_x[X]$ for an integrable random variable X. Since $G_{ab}^{(x,-x)}$ is measureable with respect to $H^{(x,-x)}$ and the random variables h_{xa} and $h_{-x,b}$ are independent of $H^{(x,-x)}$ we have

$$\mathbb{E}_x\left[h_{xa} G_{ab}^{(x,-x)} h_{bx}\right] = G_{ab}^{(x,-x)} \mathbb{E}\left[h_{xa} h_{bx}\right] = s_{xa} G_{aa}^{(x,-x)} \delta_{ab}.$$

Thus, employing this computation yields

$$\sum_{a,b}^{(x,-x)} \mathbb{E}_x \left[h_{xa} G_{ab}^{(x,-x)} h_{bx} \right] = \sum_{a}^{(x,-x)} s_{xa} G_{aa}^{(x,-x)}$$

$$= \sum_a s_{xa} G_{aa} - \sum_a s_{xa} \frac{G_{ax} G_{xa}}{G_{xx}} - s_{-x,x} G_{-x,-x}^{(x)} - \sum_{a}^{(x,-x)} s_{xa} \frac{G_{a,-x}^{(x)} G_{-x,a}^{(x)}}{G_{-x,-x}^{(x)}}$$

where we used in the second step the resolvent identity (2.9) twice. By splitting the fourth summand on the right-hand side of (4.10) in a part measureable with respect to $H^{(x,-x)}$ and the remainder, i.e. according to $\mathbb{E}_x + \mathbb{F}_x = 1$, we get the representation

$$\sum_{a,b}^{(x,-x)} h_{xa} G_{ab}^{(x,-x)} h_{bx} = \sum_{a,b}^{(x,-x)} \mathbb{E}_x \left[h_{xa} G_{ab}^{(x,-x)} h_{bx} \right] + \sum_{a,b}^{(x,-x)} \mathbb{F}_x \left[h_{xa} G_{ab}^{(x,-x)} h_{bx} \right]$$

$$= \sum_a s_{xa} G_{aa} - A_x - s_{-x,x} G_{-x,-x}^{(x)} - B_x + Z_x \tag{4.11}$$

where we used the abbreviations

$$A_x := \sum_a s_{xa} \frac{G_{ax} G_{xa}}{G_{xx}}, \qquad B_x := \sum_{a}^{(x,-x)} s_{xa} \frac{G_{a,-x}^{(x)} G_{-x,a}^{(x)}}{G_{-x,-x}^{(x)}}, \tag{4.12}$$

$$Z_x := \sum_{a,b}^{(x,-x)} \mathbb{F}_x \left[h_{xa} G_{ab}^{(x,-x)} h_{bx} \right] = \sum_{a}^{(x,-x)} \left(|h_{xa}|^2 - s_{xa} \right) G_{aa}^{(x,-x)} + \sum_{a \neq b}^{(x,-x)} h_{xa} G_{ab}^{(x,-x)} h_{bx}.$$

Therefore, the results of the equations (4.10) and (4.11) allow us to write equation (2.8) in the form

$$\frac{1}{G_{xx}} = -z - m + \Upsilon_x - \sum_a s_{xa} v_a, \tag{4.13}$$

with $v_x := G_{xx} - m$ and the error term $\Upsilon_x := h_{xx} + A_x + B_x - C_x - Y_x - Z_x$ where

$$C_x := \left(|h_{x,-x}|^2 - s_{-x,x} \right) G_{-x,-x}^{(x)} + h_{-x,x} \sum_{a}^{(x,-x)} h_{xa} G_{a,-x}^{(x)} + h_{x,-x} \sum_{b}^{(x,-x)} G_{-x,b}^{(x)} h_{bx}, \tag{4.14}$$

$$Y_x := \left(G_{-x,-x}^{(x)} \right)^{-1} \sum_{a,b}^{(x,-x)} h_{xa} G_{a,-x}^{(x)} G_{-x,b}^{(x)} h_{bx}. \tag{4.15}$$

By (4.2) equation (4.13) can be transformed into the self-consistent equation

$$-\sum_a s_{xa} v_a + \Upsilon_x = \frac{1}{v_x + m} - \frac{1}{m}. \tag{4.16}$$

This self-consistent equation has the same form as (5.9) in [7]. However, we had to replace P_i by $\mathbb{E}_x$ to deduce it and the error term Υ_x contains terms which did not appear in (5.8) from [7]. The term A_x is exactly the same as A_i in (5.8) of [7]. The term Z_x is the analogue of Z_i in [7] for the fourfold symmetry. Whereas, B_x, C_x and Y_x are completely new.

4.3 Auxiliary Estimates

The following Lemma gives bounds on the entries $G_{ij}^{(\mathbb{T})}$ of the resolvent if a preliminary bound on Λ is already known. Such a bound is described by a deterministic (possibly z-dependent) parameter Ψ such that

$$cM^{-\frac{1}{2}} \leq \Psi \leq M^{-c} \tag{4.17}$$

for some $c > 0$ and all large enough N. This Lemma combines the estimates in (2.60) of [4] and (5.12) and (B.3) of [7].

Lemma 4.2. *Let* $\mathbf{D}$ *be a spectral domain and* φ *the indicator function of a (possibly* z*-dependent) event. Let* Ψ *be a deterministic control parameter satisfying* (4.17)*. If* $\varphi\Lambda \prec \Psi$ *then for any fixed finite subset* $\mathbb{T} \subset \mathbb{N}$ *holds*

$$\varphi|G_{ij}^{(\mathbb{T})}| \prec \varphi\Lambda_o \prec \Psi, \qquad \varphi|G_{ii}^{(\mathbb{T})}| \prec 1, \qquad \frac{\varphi}{|G_{ii}^{(\mathbb{T})}|} \prec 1$$

uniformly in $z \in \mathbf{D}$ *and in* i, j *for* $i \neq j$ *and* $i, j \notin \mathbb{T}$*.*

Proof. The result is shown by induction on the number of elements of $\mathbb{T}$. For the induction basis $|\mathbb{T}| = 0$ we observe that $\varphi|G_{ij}| \leq \varphi\Lambda_o \prec \Psi$ holds by assumption for $i \neq j$. For the second estimate, let $\varepsilon > 0$ and $D > 0$. If $\varphi|G_{ii}| > N^\varepsilon$ then $\varphi\Lambda > N^\varepsilon - 1$ since $|m| \leq 1$ by (4.48). The existence of a positive integer N_0 such that

$$\mathbb{P}(\varphi|G_{ii}| > N^\varepsilon) \leq \mathbb{P}(\varphi\Lambda > N^{\delta c/2}\Psi) \leq N^{-D}$$

for all $N \geq N_0$ follows from the estimate $N^\varepsilon - 1 \geq N^{-\delta c/2} \geq N^{\delta c/2}M^{-c}$ which holds for all large N and the assumption $\varphi\Lambda \prec \Psi \leq M^{-c} \leq N^{-c\delta}$ by (2.3). Thus, $\varphi|G_{ii}| \prec 1$. Let $\varepsilon > 0$ and $D > 0$. The assumption $\varphi/|G_{ii}| > N^\varepsilon$ yields $\varphi\Lambda > d - N^\varepsilon$ where $d > 0$ is a constant such that $d \leq |m|$ (cf. (4.48)). Similarly as in the proof of the second estimate, we get

$$\mathbb{P}\left(\frac{\varphi}{|G_{ii}|} > N^\varepsilon\right) \leq \mathbb{P}(\varphi\Lambda > N^{\delta c/2}\Psi) \leq N^{-D}$$

for all $N \geq N_0$ for some positive integer N_0. This establishes the induction basis. For the induction step assume that the estimates are proved for all $\mathbb{T} \subset \mathbb{N}$ subsets with n elements. Let $\mathbb{T}' \subset \mathbb{N}$ be a subset consisting of $n+1$ elements. Then we take a $k \in \mathbb{T}'$ and set $\mathbb{T} = \mathbb{T}' \backslash \{k\}$. For $i, j \notin \mathbb{T}'$ and $i \neq j$ the resolvent identity (2.9) and the induction hypothesis implies

$$\varphi|G_{ij}^{(\mathbb{T}')}| \leq \varphi|G_{ij}^{(\mathbb{T})}| + \varphi\frac{|G_{ik}^{(\mathbb{T})}G_{kj}^{(\mathbb{T})}|}{|G_{kk}^{(\mathbb{T})}|} \prec \varphi\Lambda_o + \varphi\Lambda_o^2 \prec \varphi\Lambda_o \prec \Psi$$

where we used $\varphi\Lambda_o \prec \Psi \leq 1$ in the third and fourth step. Similarly,

$$\varphi|G_{ii}^{(\mathbb{T}')}| \leq \varphi|G_{ii}^{(\mathbb{T})}| + \varphi\frac{|G_{ik}^{(\mathbb{T})}G_{ki}^{(\mathbb{T})}|}{|G_{kk}^{(\mathbb{T})}|} \prec 1 + \Psi^2 \prec 1.$$

Using the resolvent identity (2.9) and expanding $(1-x)^{-1}$ yields

$$\frac{\varphi}{|G_{ii}^{(\mathbb{T}')}|} = \frac{\varphi}{|G_{ii}^{(\mathbb{T})}|}\left|1 - \frac{G_{ik}^{(\mathbb{T})}G_{ki}^{(\mathbb{T})}}{G_{ii}^{(\mathbb{T})}G_{kk}^{(\mathbb{T})}}\right|^{-1} \leq \frac{\varphi}{|G_{ii}^{(\mathbb{T})}|} + g_i \prec 1$$

where g_i denotes the error terms coming from the expansion. By the induction hypothesis we have $g_i \prec \Psi$ uniformly in i. In the last step, we applied the induction hypothesis again and used

$\Psi \leq 1$. □

Using this result we can now control Λ_o and the different terms in Υ_x given a preliminary bound on Λ. The following Lemma is the analogue of Lemma 5.2 from [7]. The new terms B_x, C_x and Y_x are estimated in (4.35), (4.21) and after (4.30). For the second part, these bounds are found in (4.38), after (4.36) and after (4.37).

In the proof of this Lemma, we need the assumption $\mathbb{E}h_{xy}^2 = 0$ to apply the Large Deviation Bounds in (4.27). This estimate is used in (4.29) which is in turn necessary to show the bounds on Y_x, $G_{x,-x}$ and G_{xy}, i.e. Λ_o, and therefore on A_x and B_x.

Note that we define $\eta := \operatorname{Im} z$ for a complex number z.

Lemma 4.3. *Let* $\mathbf{D}$ *be a spectral domain. Then the following statements hold*

(i) Let φ be the indicator function of a (possibly z-dependent) event such that $\varphi\Lambda \prec M^{-c}$ for some $c > 0$ then

$$\varphi(\Lambda_o + |A_x| + |B_x| + |C_x| + |Y_x| + |Z_x|) \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \tag{4.18}$$

uniformly in $z \in \mathbf{D}$.

(ii) For any N-independent $\eta > 0$ the estimate

$$\Lambda_0 + |A_x| + |B_x| + |C_x| + |Y_x| + |Z_x| \prec M^{-\frac{1}{2}} \tag{4.19}$$

holds uniformly in $z \in \{w \in \mathbf{D} : \operatorname{Im} w = \eta\}$.

Proof. In the following proof of the first statement, Lemma 4.2 will be applied several times with $\Psi = M^{-c}$. We use the resolvent identity (2.10) to get the representation

$$C_x = -|h_{x,-x}|^2 G_{-x,-x}^{(x)} - s_{-x,x} G_{-x,-x}^{(x)} - \frac{G_{x,-x}}{G_{xx}} h_{-x,x} - h_{x,-x} \frac{G_{-x,x}}{G_{xx}} \tag{4.20}$$

Thus, Lemma 2.2, $s_{-x,x} \leq M^{-1}$ and Lemma 4.2 yield

$$\varphi|C_x| \prec 2M^{-1} + 2M^{-1/2} \prec M^{-1/2}. \tag{4.21}$$

Before proceeding, we establish the auxiliary bound

$$\varphi \operatorname{Im} G_{aa}^{(\mathbb{T})} \prec \operatorname{Im} m + \Lambda \tag{4.22}$$

for a fixed finite subset $\mathbb{T} \subset \mathbb{N}$ and $a \notin \mathbb{T}$. We show this estimate by induction on the number of elements in $\mathbb{T}$. For $\mathbb{T} = \varnothing$ we have $\operatorname{Im} G_{aa} = \operatorname{Im} m + \operatorname{Im}(G_{aa} - m) \leq \operatorname{Im} m + \Lambda$. The induction step follows from the resolvent identity (2.9) and Lemma 4.2 since for a finite subset $\mathbb{T} \subset \mathbb{N}$ and $k \notin \mathbb{T}$ we have

$$\varphi \operatorname{Im} G_{aa}^{(\mathbb{T}k)} \leq \varphi \operatorname{Im} G_{aa}^{(\mathbb{T})} + \varphi \left| \frac{G_{ak}^{(\mathbb{T})} G_{ka}^{(\mathbb{T})}}{G_{kk}^{(\mathbb{T})}} \right| \prec \operatorname{Im} m + \Lambda + \varphi \Lambda_o^2 \prec \operatorname{Im} m + \Lambda$$

for $a \notin \mathbb{T}$ and $a \neq k$ where we applied the induction hypothesis in the second step.

Moreover, we remark that

$$M^{-1/2} \prec \sqrt{\frac{\operatorname{Im} m}{M\eta}} \leq \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \tag{4.23}$$

where (4.50) was used in the first step.

In the following, we will apply the Large Deviation Bounds multiple times. If not otherwise specified we first condition on $G^{(x,-x)}$ and afterwards use one of the Large Deviation Bounds of Theorem 2.6. In this case we employ the fact that the random variables $(h_{xa})_{a=-N/2}^{N/2}$ are independent of $G^{(x,-x)}$. Now, we estimate Z_x by considering its two parts separately. First, we apply the Large Deviation Bound (2.7) with $X_i = \zeta_{xi}$ and $a_{ij} = s_{xi}^{1/2} G_{ij}^{(x,-x)} s_{jx}^{1/2}$ to get

$$\begin{aligned}\varphi\Big|\sum_{i\neq j}^{(x,-x)} h_{xi}G_{ij}^{(x,-x)}h_{jx}\Big| &\prec \left(\sum_{i\neq j}^{(x,-x)} s_{xi}s_{jx}\varphi|G_{ij}^{(x,-x)}|^2\right)^{1/2}\\ &\leq \left(\frac{\varphi}{M\eta}\sum_{i}^{(x,-x)} s_{xi}\operatorname{Im} G_{ii}^{(x,-x)}\right)^{1/2} \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}}\end{aligned} \tag{4.24}$$

where we used (2.11) in the second step and (4.22) in the last step. The Large Deviation Bound (2.5) with $X_i = (|\zeta_{xi}|^2-1)(\mathbb{E}|\zeta_{xi}|^4-1)^{-1/2}$ and $a_i = (\mathbb{E}|\zeta_{xi}|^4-1)^{1/2}s_{xi}G_{ii}^{(x,-x)}$ implies

$$\begin{aligned}\varphi\Big|\sum_{i}^{(x,-x)} \left(|h_{xi}|^2 - s_{xi}\right)G_{ii}^{(x,-x)}\Big| &\prec \left(\sum_{i}^{(x,-x)} s_{xi}^2(\mathbb{E}|\zeta_{xi}|^4-1)\varphi|G_{ii}^{(x,-x)}|^2\right)^{1/2}\\ &\prec (\mu_4-1)^{1/2}M^{-1/2}\left(\sum_i s_{xi}\right)^{1/2} \prec M^{-1/2}\end{aligned} \tag{4.25}$$

where we used (2.2) and Lemma 4.2 in the second step. Therefore, using (4.23) we get the bound

$$\varphi|Z_x| \leq \varphi\Big|\sum_{i\neq j}^{(x,-x)} h_{xi}G_{ij}^{(x,-x)}h_{jx}\Big| + \varphi\Big|\sum_{i}^{(x,-x)}\left(|h_{xi}|^2-s_{xi}\right)G_{ii}^{(x,-x)}\Big| \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}}. \tag{4.26}$$

Next, we estimate $\varphi|\sum_{k,l}^{(x,-x)} h_{xk}G_{kl}^{(x,-x)}h_{l,-x}|$. This sum is split into two parts which are treated separately. Since $\mathbb{E}h_{xy}^2 = 0$ for all x,y the application of the Large Deviation Bound (2.5) with $X_k = \zeta_{xk}^2(\mathbb{E}|\zeta_{xk}|^4)^{-1/2}$ and $a_k = s_{xk}(\mathbb{E}|\zeta_{xk}|^4)^{1/2}G_{k,-k}^{(x,-x)}$ yields

$$\begin{aligned}\varphi\Big|\sum_{k}^{(x,-x)} h_{xk}^2G_{k,-k}^{(x,-x)}\Big| &\prec \left(\sum_{k}^{(x,-x)} s_{xk}^2\mathbb{E}|\zeta_{xk}|^4\varphi|G_{k,-k}^{(x,-x)}|^2\right)^{1/2}\\ &\prec \mu_4^{1/2}M^{-1/2}\left(\sum_{k}^{(x,-x)} s_{xk}M^{-2c}\right)^{1/2} \prec M^{-1/2}\end{aligned} \tag{4.27}$$

where we used (2.2) and Lemma 4.2 in the second step. The application of the Large Deviation Bound (2.7) with $X_k = \zeta_{xk}$ and $a_{kl} = s_{xk}^{1/2}G_{k,-l}^{(x,-x)}s_{xl}^{1/2}$ implies

$$\varphi\Big|\sum_{k\neq l}^{(x,-x)} h_{xk}G_{k,-l}^{(x,-x)}h_{l,-x}\Big| = \varphi\Big|\sum_{k\neq l}^{(x,-x)} h_{xk}G_{k,-l}^{(x,-x)}h_{xl}\Big| \prec \left(\sum_{k\neq l}^{(x,-x)} s_{xk}s_{xl}\varphi|G_{k,-l}^{(x,-x)}|^2\right)^{1/2} \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}} \tag{4.28}$$

where the last step is shown exactly as in (4.24). Thus, using (4.23) we get

$$\varphi\Big|\sum_{k,l}^{(x,-x)} h_{xk}G_{kl}^{(x,-x)}h_{l,-x}\Big| \leq \varphi\Big|\sum_{k}^{(x,-x)} h_{xk}^2G_{k,-k}^{(x,-x)}\Big| + \varphi\Big|\sum_{k\neq l}^{(x,-x)} h_{xk}G_{k,-l}^{(x,-x)}h_{xl}\Big| \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}}. \tag{4.29}$$

By the resolvent identity (2.10) we get the representation

$$Y_x = G^{(x)}_{-x,-x} \left(\sum_{a,k}^{(x,-x)} h_{xa} G^{(x,-x)}_{ak} h_{k,-x} \right) \left(\sum_{b,l}^{(x,-x)} h_{-x,l} G^{(x,-x)}_{lb} h_{bx} \right) \tag{4.30}$$

which implies $\varphi|Y_x| \prec \frac{\operatorname{Im} m + \Lambda}{M\eta} \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}}$ by (4.29) . We start estimating Λ_o by providing a bound on $\varphi|G_{x,-x}|$. The expansion

$$G_{x,-x} = -G_{xx} \sum_{k}^{(x)} h_{xk} G^{(x)}_{k,-x} = -G_{xx} G^{(x)}_{-x,-x} \left(h_{x,-x} - \sum_{k,l}^{(x,-x)} h_{xk} G^{(x,-x)}_{kl} h_{l,-x} \right) \tag{4.31}$$

implies

$$\varphi|G_{x,-x}| \prec M^{-1/2} + \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \tag{4.32}$$

where we applied Lemma 2.2 and (4.29) in the first step and (4.23) in the second step .

For the generic off-diagonal entry G_{xy} $(-x \neq y \neq x)$ we use the expansion

$$G_{xy} = -\,G^{(-x,-y)}_{xx} G^{(x,-x,-y)}_{yy} \left(h_{xy} - \sum_{k,l}^{(x,-x,y,-y)} h_{xk} G^{(x,-x,y,-y)}_{kl} h_{ly} \right) + \frac{G^{(-x)}_{x,-y} G^{(-x)}_{-y,y}}{G^{(-x)}_{-y,-y}} + \frac{G_{x,-x} G_{-x,y}}{G_{-x,-x}}. \tag{4.33}$$

Conditioning on $G^{(x,-x,y,-y)}$ and applying the Large Deviation Bound (2.6) with $X_k = \zeta_{xk}$, $Y_l = \zeta_{ly}$ and $a_{kl} = s^{1/2}_{xk} G^{(x,-x,y,-y)}_{kl} s^{1/2}_{ly}$ yield

$$\begin{aligned} \varphi| \sum_{k,l}^{(x,-x,y,-y)} h_{xk} G^{(x,-x,y,-y)}_{kl} h_{ly}| \prec & \varphi \left(\sum_{k,l}^{(x,-x,y,-y)} s_{xk} |G^{(x,-x,y,-y)}_{kl}|^2 s_{ly} \right)^{1/2} \\ \leq & \left(\sum_{k}^{(x,-x,y,-y)} s_{xk} \frac{\varphi \operatorname{Im} G^{(x,-x,y,-y)}_{kk}}{M\eta} \right)^{1/2} \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \end{aligned} \tag{4.34}$$

where we used (2.11) and $s_{ly} \leq M^{-1}$ in the second step and (4.22) in the last step. Using the resolvent identity (2.10), conditioning on $G^{(x,-x)}$ and applying the Large Deviation Bound (2.5) with $X_k = \zeta_{xk}$ and $a_k = s^{1/2}_{xk} G^{(x,-x)}_{k,-y}$ we get

$$\varphi|G^{(-x)}_{x,-y}| = \varphi|G^{(-x)}_{xx} \sum_{k}^{(x,-x)} h_{xk} G^{(x,-x)}_{k,-y}| \prec \varphi \left(\sum_k s_{xk} |G^{(x,-x)}_{k,-y}|^2 \right)^{1/2} = \left(\frac{\varphi \operatorname{Im} G^{(x,-x)}_{-y,-y}}{M\eta} \right)^{1/2} \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}$$

where we used (4.22) in the last step. Thus, the expansion (4.33), Lemma 2.2, (4.32) and (4.23) imply

$$\varphi|G_{xy}| \prec M^{-1/2} + 3\sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}.$$

Therefore, $\varphi|G_{xy}| \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}}$ for all $y \neq x$, i.e. $\varphi\Lambda_o \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}}$. Lemma 4.2 and the previous estimate on Λ_o imply

$$\varphi|A_x| \leq \sum_a s_{xa} \varphi|G_{ax}| \varphi|G_{xa}| \frac{\varphi}{|G_{xx}|} \prec \varphi\Lambda_o^2 \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}.$$

Similarly, the estimate on Λ_o yields

$$\varphi|B_x| \leq \sum_a^{(x,-x)} s_{xa}\varphi\frac{|G^{(x)}_{a,-x}||G^{(x)}_{-x,a}|}{|G^{(x)}_{-x,-x}|} \prec \varphi\Lambda_o^2 \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}} \tag{4.35}$$

where we used Lemma 4.2 in the second step. This concludes the proof of (4.18) as $|h_{xx}| \prec M^{-1/2}$ by Lemma 2.2.

We are now showing the second statement. In order to prove (4.19), we mainly proceed as in the proof of (4.18). However, instead of using the bounds of Lemma (4.2) we estimate the resolvent entries applying (2.12).

This estimate yields $|Z_x| \prec M^{-1/2}$ for constant η by following the same steps as in the proof of (4.26) but using (2.12) instead of Lemma 4.2. Before estimating C_x we prove an auxiliary bound. Applying the Cauchy-Schwarz inequality, $|h_{xy}| \prec s_{xy}^{1/2}$, (2.11) and (2.12) we get

$$|\sum_a^{(\mathbb{T})} h_{xa}G^{(\mathbb{T}')}_{ab}| \leq \left(\sum_a |h_{xa}|^2\right)^{1/2}\left(\sum_a |G^{(\mathbb{T}')}_{ab}|^2\right)^{1/2} \prec \left(\sum_a s_{xa}\right)^{1/2}\left(\eta^{-1}\operatorname{Im} G^{(\mathbb{T}')}_{bb}\right)^{1/2} \leq \eta^{-1}. \tag{4.36}$$

This bound applied to the second and third term in the definition of C_x (4.14) and Lemma 2.2 imply $|C_x| \prec M^{-1/2}$ for constant η. Using (2.12) and proceeding as in the proof of (4.29) we get

$$|\sum_{i,j}^{(x,-x)} h_{xi}G^{(x,-x)}_{ij}h_{j,-x}| \prec M^{-1/2}. \tag{4.37}$$

Thus, by (4.30) we have $|Y_x| \prec M^{-1}$ for constant η. The bound

$$\frac{|G^{(x)}_{a,-x}|}{|G^{(x)}_{-x,-x}|} = |\sum_k^{(x,-x)} G^{(x,-x)}_{ak}h_{k,-x}| \prec \left(\sum_k s_{ak}|G^{(x,-x)}_{ak}|^2\right)^{1/2} \leq \left(\frac{\operatorname{Im} G^{(x,-x)}_{aa}}{M\eta}\right)^{1/2} \prec M^{-1/2}$$

which follows from the Large Deviation Bound (2.5) with $X_k = \zeta_{k,-x}$ and $a_k = s^{1/2}_{k,-x}G^{(x,-x)}_{ak}$, (2.11) and (2.12) yields

$$|B_x| \leq \sum_a^{(x,-x)} s_{xa}\frac{|G^{(x)}_{a,-x}|}{|G^{(x)}_{-x,-x}|}|G^{(x)}_{-x,a}| \prec M^{-1/2} \tag{4.38}$$

where (2.12) was applied in the second step.

Before estimating A_x for constant η we consider Λ_o. Applying (2.12), Lemma 2.2 and (4.37) to the expansion (4.31) implies $|G_{x,-x}| \prec M^{-1/2}$. We use (4.33) to estimate the generic off-diagonal term, i.e. to prove $|G_{xy}| \prec M^{-1/2}$ for $-y \neq x \neq y$. A similar argument as in (4.34) where instead of Lemma 4.2 the estimate (2.12) is used and Lemma 2.2 yield that the first term in (4.33) is of order $O_\prec(M^{-1/2})$. Expanding $G_{x,-x}$ as in (4.31) with the roles of x and $-x$ reversed implies $|G_{x,-x}/G_{-x,-x}| \prec M^{-1/2}$ for constant η . Because of

$$\frac{G^{(-x)}_{-y,y}}{G^{(-x)}_{-y,-y}} = -G^{(-x,-y)}_{yy}\left(h_{-y,y} - \sum_{a,b}^{(-x,y,-y)} h_{-y,a}G^{(-x,y,-y)}_{ab}h_{by}\right)$$

we get $|G^{(-x)}_{-y,y}/G^{(-x)}_{-y,-y}| \prec M^{-1/2}$ by Lemma 2.2, (2.12) and proceeding similarly to the proof of (4.37) . Combining these three bounds and using the expansion (4.33) yield $|G_{xy}| \prec M^{-1/2}$.

Hence, $\Lambda_o \prec M^{-1/2}$. Finally, we get

$$|A_x| \leq |s_{xx}G_{xx}| + \sum_a^{(x)} s_{xa}|G_{ax}||\sum_b^{(x)} h_{xb}G_{ba}^{(x)}| \prec M^{-1} + \sum_a^{(x)} s_{xa}M^{-1/2}\eta^{-1} \prec M^{-1/2}$$

where in the second step we used $s_{xx} \leq M^{-1}$, $\Lambda_o \prec M^{-1/2}$ and (4.36). This finishes the proof of the Lemma. □

4.4 Preliminary Bound on Λ

This section is devoted to the proof of the following first preliminary bound on Λ. Its content is the basically unchanged section 5.3 of [7].

Proposition 4.4. *We have $\Lambda \prec M^{-\gamma/3}\Gamma^{-1}$ uniformly in* $\mathbf{S}$.

To show this result, we need two preparatory Lemmas. First, we prove that either $\Lambda > M^{-\gamma/4}\Gamma^{-1}$ or $\Lambda \leq M^{-\gamma/2}\Gamma^{-1}$ with very high probability. Then, we verify in the second Lemma that Λ fulfills the second bound for large η. Finally, we employ the continuity of Λ as a function of z and the connectedness of $\mathbf{S}$ to conclude that Λ cannot cross the gap between the two estimates above. The whole idea of this argument is illustrated in Figure 4.1.

Lemma 4.5. *We have the estimate $\mathbf{1}(\Lambda \leq M^{-\gamma/4}\Gamma^{-1})\Lambda \prec M^{-\gamma/2}\Gamma^{-1}$ uniformly in* $\mathbf{S}$.

Proof. The definition $\varphi := \mathbf{1}(\Lambda \leq M^{-\gamma/4}\Gamma^{-1})$ yields

$$\varphi\Lambda \leq M^{-\gamma/4}\Gamma^{-1} \leq CM^{-\gamma/4}$$

by (4.56). Therefore, (4.18) holds. In order to estimate Λ_d we use the expansion

$$\frac{\varphi}{v_x + m} = \varphi\left(\frac{1}{m} - m^{-2}v_x + g_x\right)$$

which holds because of $|v_x| \leq CM^{-\gamma/4}$ on $\{\varphi = 1\}$ and the first bound in (4.48). The error term g_x fulfills $g_x \in O_\prec(\Lambda^2)$ uniformly in z and x. Inserting the expansion into (4.16) implies

$$\varphi(-\sum_a s_{xa}v_a + \Upsilon_x) = \varphi(-m^{-2}v_x + g_x)$$

which is equivalent to

$$\varphi(v_x - m^2\sum_a s_{xa}v_a) = \varphi m^2(g_x - \Upsilon_x). \tag{4.39}$$

Introducing $\mathbf{v} = (v_x)_{x=-N/2}^{N/2}$, $S = (s_{xa})_{x,a=-N/2}^{N/2}$ and $\mathbf{V} = m^2(g_x - \Upsilon_x)_{x=-N/2}^{N/2}$ we can write these equations in the form

$$\varphi(1 - m^2S)\mathbf{v} = \varphi\mathbf{V}.$$

Inverting $(1 - m^2S)$ yields

$$\varphi\Lambda_d = \varphi\|\mathbf{v}\|_{\ell^\infty} = \varphi\|(1 - m^2S)^{-1}\mathbf{V}\|_{\ell^\infty} \leq \varphi\Gamma(z)\|\mathbf{V}\|_{\ell^\infty}.$$

Since trivially $g_x \in O_\prec(\Lambda^2)$ and $\varphi|\Upsilon_x| \prec \sqrt{\frac{\operatorname{Im} m+\Lambda}{M\eta}}$ by (4.18) we have

$$\varphi\Lambda_d \prec \varphi\Gamma(z)\left(\Lambda^2 + \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}\right).$$

Due to (4.18) and (4.56) the following estimate holds

$$\varphi\Lambda_o \prec \varphi\sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \prec \varphi\Gamma(z)\left(\Lambda^2 + \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}\right).$$

Altogether,

$$\varphi\Lambda \prec \varphi\Gamma(z)\left(\Lambda^2 + \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}\right).$$

Thus, the definition of φ implies

$$\varphi\Gamma(z)\Lambda^2 \le \varphi M^{-\gamma/2}\Gamma^{-1} \le M^{-\gamma/2}\Gamma^{-1}$$

and

$$\varphi\Gamma(z)\sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \le \Gamma\sqrt{\frac{\varphi\operatorname{Im} m}{M\eta}} + \Gamma\sqrt{\frac{\varphi\Lambda}{M\eta}} \le \Gamma\sqrt{\frac{\varphi\operatorname{Im} m}{M\eta}} + \Gamma\sqrt{\frac{\Gamma^{-1}}{M\eta}} \le 2M^{-\gamma/2}\Gamma^{-1}$$

where in the first step (7.4), in the second step the definition of φ and in the third step the definition of $\mathbf{S}$ was used. Therefore, we conclude that

$$\varphi\Lambda \prec \varphi\Gamma(z)\left(\Lambda^2 + \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}}\right) \prec M^{-\gamma/2}\Gamma^{-1}$$

which is the estimate we claimed. □

Lemma 4.6. *We have $\Lambda \prec M^{-1/2}$ uniformly in $z \in [-10, 10] + 2\mathrm{i}$.*

Proof. We use the bounds $|G_{ij}^{(\mathbb{T})}| \le 1/\eta = 1/2$ from (2.12) and $|m| \le 1/\eta = 1/2$ from (4.49). In particular, they imply $|v_x| = |G_{xx} - m| \le 1$ and $|m^{-1}| \ge 2$. The self-consistent equation (4.16) can be rewritten in the form

$$v_x = \frac{m(\sum_k s_{xk}v_k - \Upsilon_x)}{m^{-1} - \sum_k s_{xk}v_k + \Upsilon_x}. \tag{4.40}$$

Using $|m^{-1}| \ge 2$ and $|v_x| \le 1$ yields

$$\left|m^{-1} - \sum_k s_{xk}v_k + \Upsilon_x\right| \ge |m^{-1}| - \sum_k s_{xk}|v_k| - |\Upsilon_x| \ge 1 - |\Upsilon_x|.$$

Therefore, (4.40) implies

$$\Lambda_d = \max_x|v_x| \le \frac{\Lambda_d + \max_x|\Upsilon_x|}{2 - 2\max_x|\Upsilon_x|} = \frac{1}{2}\Lambda_d + g$$

where we used $|m| \le 1/2$ in the second step and expanded $(1-x)^{-1}$ in the last step. The function g is the error term coming from this expansion. As $|v_x| \le 1$ and thus $\Lambda_d \le 1$ we have $g \in O_\prec(\max_x|\Upsilon_x|) = O_\prec(M^{-1/2})$ by (4.19). Hence, $\Lambda_d \prec M^{-1/2}$. This estimate together with $\Lambda_o \prec M^{-1/2}$ by (4.19) establishes the assertion of the Lemma. □

Proof of Proposition 4.4. We fix $D > 14$. Then by Lemma 4.5 there is a positive integer $N_0 \equiv N_0(\gamma, D)$ (independent of z) such that

$$\mathbb{P}\left(M^{-\gamma/2}\Gamma(z)^{-1} \leq \Lambda(z) \leq M^{-\gamma/4}\Gamma(z)^{-1}\right) \leq N^{-D}$$

for all $z \in \mathbf{S}$ and for all $N \geq N_0$. As $\gamma/3 \leq \gamma/2$ we have

$$\mathbb{P}\left(M^{-\gamma/3}\Gamma(z)^{-1} \leq \Lambda(z) \leq M^{-\gamma/4}\Gamma(z)^{-1}\right) \leq N^{-D}$$

for all $z \in \mathbf{S}$ and for all $N \geq N_0$.

Set $d \coloneqq \sqrt{2}N^{-6}$ and consider

$$\begin{aligned} L \coloneqq &\{(-10 + kd, ld), (10, ld), (-10 + kd, \eta_{kd-10}), (-10 + kd, 10); \\ &0 \leq k \leq 10\sqrt{2}N^6, 0 \leq l \leq 5\sqrt{2}N^6\} \cup \{(10, \eta_{10}), (10, 10)\}. \end{aligned}$$

Then, $\Delta \coloneqq L \cap \mathbf{S} \subset \mathbf{S}$ with $|\Delta| \leq N^{14}$. Furthermore, for every $z \in \mathbf{S}$ there is a $w \in \Delta$ with $|z - w| \leq N^{-6}$. The above estimate implies

$$\begin{aligned} &\mathbb{P}\left(\exists w \in \Delta\colon M^{-\gamma/3}\Gamma(w)^{-1} \leq \Lambda(w) \leq M^{-\gamma/4}\Gamma(w)^{-1}\right) \\ \leq\ &\mathbb{P}\left(\bigcup_{w\in\Delta}\left\{M^{-\gamma/3}\Gamma(w)^{-1} \leq \Lambda(w) \leq M^{-\gamma/4}\Gamma(w)^{-1}\right\}\right) \leq N^{-D+14} \end{aligned}$$

for every $N \geq N_0$.

Next, we want to extend this estimate to $\mathbf{S}$. To this end, we use a continuity argument to show

$$\begin{aligned} &A \coloneqq \left\{\exists z \in \mathbf{S}\colon 2M^{-\gamma/3}\Gamma(z)^{-1} < \Lambda(z) < 2^{-1}M^{-\gamma/4}\Gamma(z)^{-1}\right\} \\ \subset\ &B \coloneqq \left\{\exists w \in \Delta\colon M^{-\gamma/3}\Gamma(w)^{-1} \leq \Lambda(w) \leq M^{-\gamma/4}\Gamma(w)^{-1}\right\}. \end{aligned}$$

Assume that there is a $z \in \mathbf{S}$ such that the property defining A is fulfilled. By the definition of Δ we may choose $w \in \Delta$ such that $|z - w| \leq N^{-6}$. We prove that w satisfies the condition in the definition of B. Since $\mathbf{S}$ is a spectral domain by Lemma 4.13 Lemma 4.14 implies that Λ and Γ are Lipschitz continuous with Lipschitz constant at most $2M^2$, $2c^{-2}M^4$ respectively. Thus, there is a $N_1 \in \mathbb{N}$ (independent of z and w) such that $3/2\Gamma(w)^{-1} \geq \Gamma(z)^{-1} \geq 3/4\Gamma(w)^{-1}$ for all $N \geq N_1$ since $c \leq \Gamma$ by (4.56). By possibly increasing N_1 we can assume that $M^{\gamma/4-3} \leq c^3/8$ and $M^{\gamma/3-3} \leq c^3/4$ for all $N \geq N_1$. Hence, the second bound in (4.57) and $M^{-1} \leq \eta$ imply

$$M^{-\gamma/3}\Gamma(w)^{-1} \leq \Lambda(w) \leq M^{-\gamma/4}\Gamma(w)^{-1}$$

for $N \geq N_1$. In the following, we assume $N_0 \geq N_1$.

Thus, we have $\mathbb{P}(A) \leq \mathbb{P}(B) \leq N^{-D+14}$ for every $N \geq N_0$. Thus, $\Xi \coloneqq A^c$ fulfills $\mathbb{P}(\Xi) \geq 1 - N^{-D+14}$ and for every $z \in \mathbf{S}$ either $\mathbf{1}(\Xi)\Lambda(z) < 2M^{-\gamma/3}\Gamma(z)^{-1}$ or $\Lambda(z) > \mathbf{1}(\Xi)2^{-1}M^{-\gamma/4}\Gamma(z)^{-1}$ hold. The inclusion $[-10 + \mathrm{i}\,10, 10 + \mathrm{i}\,10] \subset \mathbf{S}$ yields the connectedness of $\mathbf{S}$. This fact and the continuity of Λ and Γ imply that either

$$\mathbf{1}(\Xi)\Lambda(z) < 2M^{-\gamma/3}\Gamma(z)^{-1} \text{ for every } z \in \mathbf{S}$$

or

$$\Lambda(z) > \mathbf{1}(\Xi)2^{-1}M^{-\gamma/4}\Gamma(z)^{-1} \text{ for every } z \in \mathbf{S}.$$

Finally, we exclude the second case by applying Lemma 4.6. We assume that this case holds and fix a $z \in \mathbf{S}$ with $\operatorname{Im} z = 2$. Then, we have

$$2^{-1}M^{-\gamma/4}\Gamma(z)^{-1} \geq M^{-\gamma/4}c \geq M^{\varepsilon-1/2}$$

for $\varepsilon := (2-\gamma)/8$ and every sufficiently large N due to the second estimate in (4.57) and $\eta = 2$. Hence,

$$\mathbb{P}\left(\Lambda(z) < 2^{-1}M^{-\gamma/4}\Gamma(z)^{-1}\right) \geq \mathbb{P}\left(\Lambda(z) < M^{\varepsilon-1/2}\right) \geq \mathbb{P}\left(\Lambda(z) < N^{\varepsilon}M^{-1/2}\right) \geq 1 - N^{-E}$$

by Lemma 4.6 for every sufficiently large N and every $E > 0$.

Therefore, $\mathbb{P}\left(\Lambda(z) < 2^{-1}M^{-\gamma/4}\Gamma(z)^{-1}\right) \geq 1/2$ for every sufficiently large N. However, this contradicts $\mathbb{P}(\Xi) \geq 1 - N^{-D+14}$. This finishes the proof. □

4.5 Iteration Step and Proof of the main Result

In the whole section which is the analogue of section 5.3 in [7], let Ψ be a deterministic control parameter satisfying

$$cM^{-1/2} \leq \Psi \leq M^{-\gamma/3}\Gamma^{-1}. \tag{4.41}$$

Note that Ψ satisfies (4.17) if it fulfills (4.41) due to (4.56). Moreover, we have

$$M^{-\gamma/3}\Gamma^{-1} \geq (M\eta)^{-1/3} \geq \left(\frac{\operatorname{Im} m(z)}{M\eta}\right)^{1/3} \geq cM^{-1/3} \geq cM^{-1/2}$$

where we used the definition of $\mathbf{S}$ in the first step, $\operatorname{Im} m(z) \leq 1$ by (4.48) in the second step and (4.50) in the third step. Thus, the upper bound in (4.41) is bigger than the lower bound.

Now, we are proving that a given bound on Λ in terms of a deterministic control parameter can be iteratively improved by applying the fluctuation averaging mechanism. The analogue of the following result is Proposition 5.6 in [7] whose proof is basically the same.

Proposition 4.7. *Let Ψ be a deterministic control parameter satisfying (4.41) and fix $\varepsilon \in (0, \gamma/3)$. If $\Lambda \prec \Psi$ then $\Lambda \prec F(\Psi)$ with*

$$F(\Psi) := M^{-\varepsilon}\Psi + \sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{M^{\varepsilon}}{M\eta}.$$

Proof. Because of $\Psi \leq CM^{-\gamma/3}$ by (4.56) we may apply (4.18) with $\varphi = 1$. This yields

$$\Lambda_o + |\Upsilon_x| \prec \sqrt{\frac{\operatorname{Im} m + \Lambda}{M\eta}} \prec \sqrt{\frac{\operatorname{Im} m + \Psi}{M\eta}}. \tag{4.42}$$

Next, we want to estimate Λ_d. Therefore, we consider the event $E := \{\Lambda(z) \leq M^{-\gamma/4}\}$ and show that E has asymptotically a high probability. Fix $D > 0$. Due to (4.41), (2.3) and $\Lambda \prec \Psi$ there is $\tilde{\varepsilon} > 0$ such that the inequalities

$$\mathbb{P}(E^c) \leq \mathbb{P}(\Lambda > N^{\tilde{\varepsilon}}\Psi) \leq N^{-D}$$

hold for all $N \geq N_0$. Thus, Lemma 2.3 implies $1 - \psi \prec 0$ for $\psi := \mathbf{1}(E)$. Now, (4.39) yields

$$\psi|v_x| \leq C\psi|\sum_y s_{xy}v_y| + C\psi|\Upsilon_x| + C\psi|g_x|.$$

Thus, we get

$$\psi|v_x| \prec \Gamma\Psi^2 + \sqrt{\frac{\operatorname{Im} m + \Psi}{M\eta}}$$

uniformly in x and z where we applied the fluctuation averaging (2.19) with $t_{ik} = s_{ik}$ and (4.56) to the first term, (4.42) to the second term and used $|g_x| \prec \Lambda^2 \prec \Psi^2$ uniformly in x and (4.56) to estimate the third term. As $1 - \psi \prec 0$ we get

$$\Lambda_d = \psi\Lambda_d + (1-\psi)\Lambda_d \prec \Gamma\Psi^2 + \sqrt{\frac{\operatorname{Im} m + \Psi}{M\eta}}.$$

Combining this estimate with (4.42) yields

$$\Lambda \prec \Gamma\Psi^2 + \sqrt{\frac{\operatorname{Im} m + \Psi}{M\eta}} \prec M^{-\varepsilon}\Psi + \sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{M^{\varepsilon}}{M\eta} + M^{-\varepsilon}\Psi \prec M^{-\varepsilon}\Psi + \sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{M^{\varepsilon}}{M\eta}.$$

In the second step, we used that $\Psi \leq M^{-\gamma/3}\Gamma^{-1}$ implies $\Gamma\Psi^2 \leq M^{-\varepsilon}\Psi$, (7.4) and

$$\frac{\Psi}{M\eta} \leq \frac{M^{2\varepsilon}}{2(M\eta)^2} + \frac{1}{2}M^{-2\varepsilon}\Psi^2$$

which is a consequence of Young's inequality (7.5). This concludes the proof. □

The following Lemma corresponds to Lemma 5.7 in [7]. Due to the fourfold symmetry the error term Υ_x contains the additional terms B_x, $s_{x,-x}\mathbb{E}_x G^{(x)}_{-x,-x}$ and $\mathbb{E}_x Y_x$ compared to the twofold symmetry.

The proof of (4.44) in the next Lemma requires $\mathbb{E}h_{xy}^2 = 0$ in order to apply the Large Deviation Bounds.

For the average of a vector $w = (w_i)_i$ with N entries we use the notation

$$[w] = \frac{1}{N}\sum_i w_i.$$

Lemma 4.8. *If Ψ is a deterministic control parameter such that $\Lambda \prec \Psi$ then we have $[\Upsilon] \in O_{\prec}(\Psi^2)$.*

Proof. By Schur's complement formula (2.8) and the definition of Υ_x we have

$$\Upsilon_x = A_x + B_x - s_{x,-x}\mathbb{E}_x G^{(x)}_{-x,-x} - \mathbb{E}_x Y_x + \mathbb{F}_x \frac{1}{G_{xx}}. \tag{4.43}$$

The fluctuation averaging (2.18) with $t_{ik} = 1/N$ yields $[\mathbb{F}_x G_{xx}^{-1}] \in O_{\prec}(\Psi^2)$. Obviously, we have $|A_x| \prec \Psi^2$ and $|B_x| \prec \Psi^2$ by Lemma 4.2. Since $\mathbf{S}$ is a spectral domain by Lemma 4.13 the estimate (2.12) implies $\mathbb{E}|G^{(x)}_{-x,-x}|^p \leq N^p$. Therefore, Lemma 6.1 is applicable which implies together with Lemma 4.2 and $s_{-x,x} \leq M^{-1}$ that $|s_{x,-x}\mathbb{E}_x G^{(x)}_{-x,-x}| \prec \Psi^2$ since $\Psi \leq CM^{-\gamma/3}$ by (4.41) and (4.56). Using (4.27) and the first two steps in (4.28) with $\varphi = 1$ we get

$$\Big|\sum_{k,l}^{(x,-x)} h_{xk} G^{(x,-x)}_{kl} h_{l,-x}\Big| \prec M^{-1/2} + \Psi \prec \Psi \tag{4.44}$$

since Ψ satisfies (4.41). Thus, the representation of Y_x in (4.30) and applying Lemma 4.2 yields $|Y_x| \prec \Psi^2$. Hence, Lemma 6.1 whose requirement on the moments of $|Y_x|$ is fulfilled because of (2.2) and (2.13) implies $|\mathbb{E}_x Y_x| \prec \Psi^2$ and $|[\Upsilon]| \prec \Psi^2$ follows from (4.43). □

In order to prove the main result we apply Proposition 4.7 iteratively starting with the bound in Proposition 4.4. The following proof of Theorem 4.1 essentially agrees with the proof of Theorem 5.1 in [7].

Proof of Theorem 4.1. First, we proof that if Ψ is a deterministic control parameter which fulfills (4.41) then $F(\Psi)$ satisfies (4.41) as well. Due to (4.50) we have $F(\Psi) \geq \sqrt{c}M^{-1/2}$. As $\Psi \leq M^{-\gamma/3}\Gamma^{-1}$ we have $M^{-\varepsilon}\Psi \leq M^{-\gamma/3-\varepsilon}\Gamma^{-1}$. By definition of $\mathbf{S}$ we have

$$\sqrt{\frac{\operatorname{Im} m}{M\eta}} \leq M^{-\gamma}\Gamma^{-2} \leq M^{-\gamma/3-\varepsilon}\Gamma^{-1}$$

for large N. Moreover, the definition of $\mathbf{S}$ yields $M^{\varepsilon}(M\eta)^{-1} \leq c^{-2}M^{\varepsilon-\gamma} \leq M^{-\gamma/3-\varepsilon}\Gamma^{-1}$ for large N. Thus, $F(\Psi) \leq M^{-\gamma/3-\varepsilon}\Gamma^{-1}$ for all large N and $F(\Psi)$ fulfills (4.41).

Obviously, the deterministic control parameter $\Psi_0 := M^{-\gamma/3}\Gamma^{-1}$ satisfies (4.41). Proposition 4.4 states the validity of $\Lambda \prec \Psi_0$. With the recursive definition $\Psi_{k+1} := F(\Psi_k)$ we get the estimate $\Lambda \prec \Psi_k$ for all $k \in \mathbb{N}$ which follows from Proposition 4.7 by induction. Moreover, we have the explicite formula

$$\Psi_k = M^{-k\varepsilon}M^{-\gamma/3}\Gamma^{-1} + \left(\sum_{i=0}^{k-1} M^{-i\varepsilon}\right)\sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{1}{M\eta}\left(\sum_{i=-1}^{k-2} M^{-i\varepsilon}\right).$$

The estimate $M^{-k\varepsilon}M^{-\gamma/3}\Gamma^{-1} \prec \sqrt{\frac{\operatorname{Im} m}{M\eta}}$ for $k := \lceil \varepsilon^{-1} \rceil$ which is a consequence of (4.56) and (4.50) and the above representation of Ψ_k imply

$$\Lambda \prec \sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{M^{\varepsilon}}{M\eta} \leq \sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{N^{\varepsilon}}{M\eta}$$

for all $\varepsilon > 0$. Therefore, we conclude by Lemma 2.4 (iv) that

$$\Lambda \prec \sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{1}{M\eta} = \Pi(z)$$

which is the bound claimed in (4.5).

Next, we show (4.6). We set $\psi := \mathbf{1}(\Lambda \leq M^{-\gamma/4})$. Averaging equation (4.39) implies

$$\psi m^2\left(-[v] + [\Upsilon]\right) = -\psi[v] + g$$

where $g := N^{-1}\sum_x g_x \in O_{\prec}(\psi\Lambda^2)$. Since Π satisfies (4.41) and $\Lambda \prec \Pi$ it follows from (4.42) with $\Psi = \Pi$ that

$$|\Upsilon_x| \prec \sqrt{\frac{\operatorname{Im} m + \Pi}{M\eta}} \prec \sqrt{\frac{\operatorname{Im} m}{M\eta}} + \frac{1}{M\eta} + \Pi \prec \Pi$$

where we applied (7.4) and Young's inequality (7.5) in the second step and the definition of Π in the last step. Thus, Lemma 4.8 implies $|[\Upsilon]| \prec \Pi^2$. Therefore, we get

$$|1-m^2||[v]| = \psi|1-m^2||[v]| + (1-\psi)|1-m^2||[v]| \prec |[\Upsilon]| + |g| \prec \Pi^2$$

since we have $1-\psi \prec 0$ as in the proof of Proposition 4.7. We conclude

$$|[v]| \prec \frac{\Pi^2}{|1-m^2|} \leq \left(\frac{\operatorname{Im} m}{|1-m^2|} + \frac{1}{|1-m^2|M\eta}\right)\frac{2}{M\eta} \leq \left(C + \frac{\Gamma}{M\eta}\right)\frac{2}{M\eta} \leq \frac{C}{M\eta}$$

where we used Young's inequality (7.5) in the second step and (4.51) and (4.55) in the third step. In the last step the definition of $\mathbf{S}$ and $\Gamma^{-2} \leq c$ by (4.56) was applied. □

4.6 Properties of m and Γ

In this section, we collect some properties of m and Γ which were used in the previous sections and prove these results.

Recall that m is the Stieltjes transform of the semicircle law, i.e.

$$m(z) = \frac{1}{2\pi} \int_{[-2,2]} \frac{\sqrt{4 - x^2}}{x - z} \mathrm{d}x$$

for $z \in \mathbb{C}\backslash\mathbb{R}$. In the next Lemma we establish an explicit representation of $m(z)$ for $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$. Since $m(z) = \overline{m(\overline{z})}$ this determines $m(z)$ for all $z \in \mathbb{C}\backslash\mathbb{R}$.

Lemma 4.9. *For $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$ we have*

$$m(z) = \frac{-z + \sqrt{z^2 - 4}}{2}.$$

Before embarking on the proof we explain our convention concerning the square root of a complex number. We define the square root $\sqrt{w}$ of a complex number which is not real to be the unique solution of $z^2 - w = 0$ with positive imaginary part, explicitely

$$\sqrt{w} = \operatorname{sign}(\operatorname{Im} w) \frac{|w| + w}{\sqrt{2(|w| + \operatorname{Re} w)}}. \tag{4.45}$$

Moreover, for its real and imaginary part, respectively, we have

$$\operatorname{Re} \sqrt{w} = \frac{1}{\sqrt{2}} \operatorname{sign}(\operatorname{Im} w) \sqrt{|w| + \operatorname{Re} w} = \frac{\operatorname{Im} w}{\sqrt{2(|w| - \operatorname{Re} w)}}, \tag{4.46}$$

$$\operatorname{Im} \sqrt{w} = \frac{|\operatorname{Im} w|}{\sqrt{2(|w| + \operatorname{Re} w)}} = \frac{1}{\sqrt{2}} \sqrt{|w| - \operatorname{Re} w}. \tag{4.47}$$

The proof we give here follows [2]. The main idea is to convert the integral into a contour integral in the complex plane and apply the residue theorem. An alternative proof which is based on representing $(x - z)^{-1}$ in the integrand as a power series, interchanging summation and integration and computing the moments of the semicircle law μ_{sc} is given in [12].

Proof. Using the substitution $x = 2\cos t$ we get

$$m(z) = \frac{2}{\pi} \int_0^{\pi} \frac{\sin^2 t}{2\cos t - z} \mathrm{d}t.$$

The substitution $\varphi = 2\pi - t$ implies

$$\int_0^{\pi} \frac{\sin^2 t}{2\cos t - z} \mathrm{d}t = \int_{\pi}^{2\pi} \frac{\sin^2 \varphi}{2\cos \varphi - z} \mathrm{d}\varphi.$$

Thus, we have

$$\begin{aligned} m(z) =& \frac{1}{\pi}\int_0^{2\pi} \frac{\sin^2\varphi}{2\cos\varphi - z}\mathrm{d}\varphi \\ =& \frac{1}{\pi}\int_0^{2\pi} \frac{1}{2\frac{\mathrm{e}^{i\varphi}+\mathrm{e}^{-i\varphi}}{2} - z}\left(\frac{\mathrm{e}^{i\varphi}-\mathrm{e}^{-i\varphi}}{2i}\right)^2\mathrm{d}\varphi \\ =& -\frac{1}{4\pi i}\int_{S^1}\frac{(\zeta-\zeta^{-1})^2}{\zeta+\zeta^{-1}-z}\zeta^{-1}\mathrm{d}\zeta \\ =& -\frac{1}{4\pi i}\int_{S^1}\frac{(\zeta^2-1)^2}{\zeta^2(\zeta^2-z\zeta+1)}\mathrm{d}\zeta. \end{aligned}$$

We will apply the residue theorem to the last integral. We set

$$f(\zeta) = -\frac{(\zeta^2-1)^2}{4\pi i\zeta^2(\zeta^2-z\zeta+1)}.$$

The poles of f are $\zeta_0 = 0$, $\zeta_+ = (z+\sqrt{z^2-4})/2$ and $\zeta_- = (z-\sqrt{z^2-4})/2$. As $\operatorname{Re}\sqrt{z^2-4}$ and $\operatorname{Re} z$ have the same sign and $\operatorname{Im}\sqrt{z^2-4}$ and $\operatorname{Im} z$ are nonnegative we get $|\operatorname{Re}\zeta_+| > |\operatorname{Re}\zeta_-|$ and $|\operatorname{Im}\zeta_+| > |\operatorname{Im}\zeta_-|$. Thus, $|\zeta_+| > |\zeta_-|$ which implies together with $\zeta_+\zeta_- = 1$ that $0, \zeta_- \in D_1(0)$ and $\zeta_+ \notin D_1(0)$ where $D_1(0) := \{w \in \mathbb{C}; |w| < 1\}$. At these poles the residues are

$$\operatorname{Res}(f,0) = -\frac{z}{4\pi i}, \text{ and } \operatorname{Res}(f,\zeta_-) = -\frac{(\zeta_- - 1)^2}{4\pi i\zeta_-^2(\zeta_- - \zeta_+)} = -\frac{\zeta_-^2(\zeta_- - \zeta_+)^2}{4\pi i\zeta_-^2(\zeta_- - \zeta_+)} = \frac{\sqrt{z^2-4}}{4\pi i}.$$

Hence,

$$m(z) = 2\pi i(\operatorname{Res}(f,0) + \operatorname{Res}(f,\zeta_-)) = \frac{-z+\sqrt{z^2-4}}{2}.$$

□

Remark 4.10. *Lemma 4.9 implies that* $m(z)^2 + zm(z) + 1 = 0$ *which is equivalent to* (4.2). *Together with* $\operatorname{Im} m(z) \geq c\eta$ *for some* $c > 0$ *by* (4.50) *this yields that* $m(z)$ *is the unique solution of* (4.2) *with* $\operatorname{Im} m(z) > 0$ *for* $\operatorname{Im} z > 0$.

In the following we denote

$$\mathbf{R} := \{E + \mathrm{i}\eta; E \in [-10, 10], \eta \in (0, 10]\}.$$

The following Lemma contains some basic estimates on m.

Lemma 4.11. *There are constants* $c > 0$ *and* $C > 0$ *such that for* $z \in \mathbf{R}$ *we have*

$$c \leq |m(z)| \leq 1 - c\eta, \tag{4.48}$$

$$|m(z)| \leq \eta^{-1}, \tag{4.49}$$

$$\operatorname{Im} m(z) \geq c\eta, \tag{4.50}$$

$$\operatorname{Im} m(z) \leq C|1 - m^2(z)|. \tag{4.51}$$

Moreover, we have

$$|m(z) - m(w)| \leq (\operatorname{Im} z)^{-1}(\operatorname{Im} w)^{-1}|z - w| \tag{4.52}$$

for $z, w \in \mathbb{C}$ *with* $\operatorname{Im} z, \operatorname{Im} w > 0$.

Note that decreasing c in any of the above estimates does not affect the validity of these bounds. Therefore, we can prove each inequality with a different constant c and we finally take c to be the minimum of these constants.

Proof. We start with proving (4.50). Since $|(x-z)^{-1}| = ((x-E)^2+\eta^2)^{-1/2} \le \eta^{-1}$ and the semicircle law defines a probability measure we have

$$|m(z)| \le \frac{1}{2\pi}\int_{[-2,2]} \frac{\sqrt{4-x^2}}{|x-z|}\mathrm{d}x \le \eta^{-1}.$$

Note that there is a $d>0$ such that $(x-E)^2+\eta^2 \le d$ for $x \in [-2,2]$, $E \in [-10,10]$ and $\eta \in (0,10]$. Thus,

$$\operatorname{Im} m(z) \ge \frac{d^{-1}\eta}{2\pi}\int_{[-2,2]} \sqrt{4-x^2}\mathrm{d}x = d^{-1}\eta.$$

Taking the imaginary part of (4.2) we get $\operatorname{Im} m(z) = |m(z)+z|^{-2}(\operatorname{Im} m(z)+\eta) = |m(z)|^2(\operatorname{Im} m(z)+\eta)$ where we employed (4.2) also in the second step. Equivalently,

$$(1-|m(z)|^2)\operatorname{Im} m(z) = |m(z)|^2\eta. \tag{4.53}$$

This equation yields $1-|m(z)|^2 \ge 0$ since $\operatorname{Im} m(z) > 0$ for $z \in \mathbf{R}$ by (4.50). Thus, $|m(z)| \le 1$ for all $z \in \mathbf{R}$ and there is $D>0$ such that $|m(z)+z| \le D$ for all $z \in \mathbf{R}$. This implies

$$|m(z)| = \frac{1}{|m(z)+z|} \ge D^{-1} =: c$$

for all $z \in \mathbf{R}$ where we used (4.2) in the first step. This shows the first bound in (4.48). Moreover, (4.53) implies

$$1-|m(z)|^2 \ge (1-|m(z)|^2)\operatorname{Im} m(z) = |m(z)|^2\eta \ge c^2\eta$$

for all $z \in \mathbf{R}$ where we used $\operatorname{Im} m(z) \le |m(z)| \le 1$ in the first step and the first estimate in (4.48) in the last step. Thus, $|m(z)|^2 \le 1-c^2\eta \le (1-\eta c^2/2)^2$ which shows the second bound in (4.48).

Using temporarily the notation $x := \operatorname{Re} m(z)$ and $y := \operatorname{Im} m(z)$ we have

$$|1-m^2(z)|^2 = (1-x^2+y^2)^2+(2xy)^2 = (1-x^2)^2+2x^2y^2+y^4+2y^2 \ge 2y^2 = 2(\operatorname{Im} m(z))^2.$$

This yields together with $\operatorname{Im} m(z) > 0$ by (4.50) the estimate $|1-m^2(z)| \ge \sqrt{2}\operatorname{Im} m(z)$ which establishes (4.51).

The last claim follows from

$$|(x-z)^{-1}-(x-w)^{-1}| \le (\operatorname{Im} z)^{-1}(\operatorname{Im} w)^{-1}|z-w| \tag{4.54}$$

and the fact that the semicircle law defines a probability measure. □

An important parameter of our system is the quantity

$$\Gamma(z) := \|(1-m^2(z)S)^{-1}\|_{\ell^\infty\to\ell^\infty}$$

for $z \in \mathbb{C}$ with $\operatorname{Im} z > 0$. Some useful bounds on Γ are collected in the following Lemma.

Lemma 4.12. *There is a constant $c>0$ such that*

$$\Gamma(z) \ge |1-m^2(z)|^{-1}, \tag{4.55}$$
$$\Gamma(z) \ge 1/2, \tag{4.56}$$
$$c\eta \le \Gamma(z) \le c^{-1}\eta^{-1} \tag{4.57}$$

for all $z \in \mathbf{R}$.

Proof. First, note that

$$\|S\|_{\ell^\infty\to\ell^\infty} = \max_x \sum_y |S_{xy}| = 1. \tag{4.58}$$

Consider the vector $\mathbf{e} := (1,\dots,1) \in \mathbb{C}^N$. Then $S\mathbf{e} = \mathbf{e}$. In particular, $(1-m^2(z)S)^{-1}\mathbf{e} = (1-m^2(z))^{-1}\mathbf{e}$. This implies (4.55). Combining this with $|m(z)|^2 \le 1$ yields

$$\Gamma(z) = \|(1-m^2(z)S)^{-1}\| \ge (1+|m(z)|^2)^{-1} \ge 1/2.$$

Similarly, since $\|m^2(z)S\| < 1$ by the second estimate in (4.48) and (4.58) we get by applying the Neumann series

$$\Gamma(z) = \|(1-m^2(z)S)^{-1}\| \le (1-|m(z)|^2)^{-1} \le c^{-1}\eta^{-1}$$

where we used $|m(z)|^2 \le (1-c\eta)^2 \le 1-c\eta$ in the third step. As $\sigma(A^{-1}) = \sigma(A)^{-1}$ for an invertible A we have

$$\|(1-m^2(z)S)^{-1}\| \ge r\left((1-m^2(z)S)^{-1}\right) = \inf_{\lambda\in\sigma(S)} |1-m^2(z)\lambda| \ge 1-|m(z)|^2 \ge 1-(1-c\eta) = c\eta$$

where $r\left((1-m^2(z)S)^{-1}\right)$ denotes the spectral radius of $(1-m^2(z)S)^{-1}$. We used $\|S\|_{\ell^\infty\to\ell^\infty} = 1$ in the third step and $|m(z)| \le (1-c\eta)^2 \le 1-c\eta$ in the fourth step. □

Now, we can show that $\mathbf{S}$ is a spectral domain which is a consequence of the first relation in the definition of η_E and the estimate (4.56).

Lemma 4.13. *The family* $\mathbf{S}$ *of subsets of* $\mathbb{C}$ *defined in* (4.4) *is a spectral domain.*

Proof. Let $z \in \mathbb{C}$ and $\eta := \operatorname{Im} z$. Then $\eta \ge \eta_E$ where η_E was defined in (4.3). Thus,

$$\frac{1}{M\eta} \le \frac{M^{-\gamma}}{\Gamma(z)^3} \le 8M^{-\gamma} \le 1.$$

Here, we used (4.3) in the first step and (4.56) in the second step. The last estimate holds for all large N because of (2.3). This implies $\eta \ge M^{-1}$ for all large N, i.e. $\mathbf{S}$ is a spectral domain. □

Lemma 4.14. *For every spectral domain* $\mathbf{D}$*, the maps* Λ *and* Γ *are Lipschitz-continuous on* $\mathbf{D}$ *with*

$$|\Lambda(z) - \Lambda(w)| \le 2M^2|z-w|,$$
$$|\Gamma(z) - \Gamma(w)| \le 2c^{-2}M^4|z-w|$$

for $z, w \in \mathbf{D}$.

Proof. We have $\Lambda(z) = \|G(z) - m(z)1\|_{\max}$ where $\|\cdot\|_{\max}$ denotes the matrix norm of elementwise maximum. Thus,

$$|\Lambda(z)-\Lambda(w)| \le \|G(z)-m(z)1-G(w)+m(w)1\|_{\max} \le \|G(z)-G(w)\|_{\max} + |m(z)-m(w)|.$$

On the other hand, we have

$$\|G(z)-G(w)\|_{\ell^2\to\ell^2} = \|(\cdot-z)^{-1} - (\cdot-w)^{-1}\|_{\sigma(H)} \le (\operatorname{Im} z)^{-1}(\operatorname{Im} w)^{-1}|z-w|.$$

where we applied the functional calculus in the first step and the estimate (4.54) in the second step.

Because of $\|A\|_{\max} \leq \|A\|_{\ell^2\to\ell^2}$ and (4.52) we get

$$|\Lambda(z) - \Lambda(w)| \leq 2M^2|z-w|$$

since $\operatorname{Im} z \geq M^{-1}$ for $z \in \mathbf{D}$ by the defining property of a spectral domain.

For the second claim we use the resolvent identity $(1-A)^{-1} - (1-B)^{-1} = (1-A)^{-1}(A-B)(1-B)^{-1}$ for $1 \in \rho(A)$ and $1 \in \rho(B)$. The inverse triangular inequality and this resolvent identity imply

$$\begin{aligned}|\Gamma(z) - \Gamma(w)| \leq & \|(1-m^2(z)S)^{-1}(m^2(z)-m^2(w))S(1-m^2(w)S)^{-1}\|_{\ell^\infty\to\ell^\infty} \\ \leq & \Gamma(z)|m^2(z)-m^2(w)|\Gamma(w) \leq 2c^{-2}M^4|z-w|\end{aligned}$$

where we used the submultiplicativity of $\|\cdot\|_{\ell^\infty\to\ell^\infty}$ and $\|S\|_{\ell^\infty\to\ell^\infty} = 1$ in the second step and the second bound in (4.57) and in (4.48) and $M^{-1} \leq \eta_E$ in the last step. $\square$

5 Proof of the Resolvent Identities

The aim of this chapter is the verification of Schur's complement formula (2.8) and the resolvent identities given in Lemma 2.9. The approach we are pursuing here is guided by [10]. The proofs are based on the following auxiliary Lemma which is pure linear algebra.

Its formulation requires the following notation: For $A \in \mathbb{C}^{n\times n}$ and $\mathbb{T} \subset \{1,\dots,n\}$ we define $A^{[\mathbb{T}]} \in \mathbb{C}^{l\times l}$ with $l := n - |\mathbb{T}|$ through

$$(A^{[\mathbb{T}]})_{ij} := A_{ij}$$

for $i,j \in \{1,\dots,n\}\backslash\mathbb{T}$. Note that $A^{[\mathbb{T}]}$ is the minor of A arising by removing the rows and columns with index in $\mathbb{T}$ and keep the numbering of the rows and columns. The difference to $H^{(\mathbb{T})}$ defined in Definition 2.7 is that the rows and columns are removed and not solely replaced by zeros.

Lemma 5.1. *Let $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{m\times n}$ and $C \in \mathbb{C}^{m\times m}$ be matrices such that C is invertible. Set*

$$D := \begin{pmatrix} A & B^* \\ B & C \end{pmatrix} \in \mathbb{C}^{(m+n)\times(m+n)}$$

and

$$\hat{D} := A - B^*C^{-1}B \in \mathbb{C}^{n\times n}.$$

If D is invertible then $\hat{D}$ is invertible with

$$(D^{-1})_{ij} = (\hat{D}^{-1})_{ij}$$

for $i,j \in \{1,\dots,n\}$. Furthermore, if $D^{[\mathbb{T}]}$ is invertible then $\hat{D}^{[\mathbb{T}]}$ is invertible for $\mathbb{T} \subset \{1,\dots,n\}$ and

$$((D^{[\mathbb{T}]})^{-1})_{ij} = ((\hat{D}^{[\mathbb{T}]})^{-1})_{ij}$$

for $i,j \in \{1,\dots,n\}\backslash\mathbb{T}$.

Proof. First, we want to verify that it suffices to prove the first claim. Therefore, we write

$$D^{[\mathbb{T}]} = \begin{pmatrix} A^{[\mathbb{T}]} & (\tilde{B})^* \\ \tilde{B} & C \end{pmatrix}$$

where $\tilde{B} \in \mathbb{C}^{(m-l)\times n}$ ($l := |\mathbb{T}|$) is the matrix B with the columns with index in $\mathbb{T}$ removed. Since

$$(B^*C^{-1}B)_{ij} = \sum_{k,l} B^*_{ik}(C^{-1})_{kl}B_{lj} = \sum_{k,l} (\tilde{B})^*_{ik}(C^{-1})_{kl}\tilde{B}_{lj}$$

for $i,j \notin \mathbb{T}$ we have $(B^*C^{-1}B)^{[\mathbb{T}]} = (\tilde{B})^*C^{-1}\tilde{B}$. Thus, the second claim follows from the first by replacing D with $D^{[\mathbb{T}]}$.

In order to prove the first claim we write D^{-1} in block form:

$$D^{-1} = \begin{pmatrix} S & T \\ U & V \end{pmatrix}$$

where $S \in \mathbb{C}^{n\times n}$, $T, U \in \mathbb{C}^{m\times n}$ and $V \in \mathbb{C}^{m\times m}$.

In block form the identities $DD^{-1} = D^{-1}D = 1$ have the following form:

$$DD^{-1} = \begin{pmatrix} AS + B^*U & AT + B^*V \\ BS + CU & BT + CV \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{5.1}$$

$$D^{-1}D = \begin{pmatrix} SA + TB & SB^* + TC \\ UA + VB & UB^* + VC \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{5.2}$$

Using the relations from the upper left block and the lower left block in (5.1) and the definition of $\hat{D}$ yields $\hat{D}S = 1$. Similarly, the relations from the upper left block and the lower left block in (5.2) and the definition of $\hat{D}$ imply $S\hat{D} = 1$. Thus, $\hat{D}$ is invertible and $(\hat{D})^{-1} = S = ((D^{-1})_{ij})_{i,j=1}^n$. □

A direct consequence of the previous Lemma is Schur's complement formula which we prove next.

Proof of (2.8). By replacing H by UHU conjugated with a unitary matrix U interchanging e_1 and e_i we may without loss of generality assume that $i = 1$.

To show (2.8) in the case $i = 1$ we apply Lemma 5.1 with $n = 1$, $m = N - 1$ and $A = h_{11} - z$, $B = (h_{i1})_{i=2}^N$ and $C = H^{[1]} - z$. Since $\hat{D} \in \mathbb{C}^{1\times 1}$ we get

$$G_{11} = \left((H - z)^{-1}\right)_{11} = \left(h_{11} - z - \sum_{k,l}^{(1)} h_{1k}((H^{[1]} - z)^{-1})_{kl} h_{l1}\right)^{-1} = \left(h_{11} - z - \sum_{k,l}^{(1)} h_{1k} G_{kl}^{(1)} h_{l1}\right)^{-1}.$$

In the third step, we used

$$(H^{(1)} - z)^{-1} = \begin{pmatrix} z^{-1} & 0 \\ 0 & (H^{[1]} - z)^{-1} \end{pmatrix}.$$

□

Lemma 2.9 follows from Lemma 5.1 by explicitely inverting 2×2- and 3×3-matrices as we will see next.

Proof of Lemma 2.9. First, we show that

$$G_{ii}^{(\mathbb{T})} - G_{ii}^{(k\mathbb{T})} = G_{ik}^{(\mathbb{T})} G_{ki}^{(\mathbb{T})} \left(G_{kk}^{(\mathbb{T})}\right)^{-1} \tag{5.3}$$

for $i \neq k$, $i, k \notin \mathbb{T}$. By replacing H by $H^{(\mathbb{T})}$ and possibly interchanging the rows and columns of H it suffices to prove this identity for $\mathbb{T} = \varnothing$, $i = 1$ and $k = 2$.

In this situation, we apply Lemma 5.1 with $n = 2$, $m = N - 2$ and $D = H - z$. We set

$$\tilde{H}^* = \begin{pmatrix} h_{13} & \dots & h_{1N} \\ h_{23} & \dots & h_{2N} \end{pmatrix}.$$

Then Lemma 5.1 yields

$$G_{rs} = (\hat{D}^{-1})_{rs} \qquad \text{for } r, s \in \{1, 2\}, \tag{5.4}$$

$$G_{11}^{(2)} = ((\hat{D}^{(2)})^{-1})_{11} = \frac{1}{h_{11} - z - (\tilde{H}^*(H^{[1,2]} - z)^{-1}\tilde{H})_{11}} \tag{5.5}$$

where we used $\hat{D}^2 \in \mathbb{C}^{1\times 1}$ in the second line.

Since $\hat{D} \in \mathbb{C}^{2\times 2}$ we use

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

to compute the inverse of $\hat{D}$. Note that the determinat of the matrix being inverted is $ad - bc$. Using this formula, (5.4) and (5.5) we get

$$\begin{aligned} G_{11} - G_{11}^{(2)} =& \frac{\left(h_{12} - z - (\tilde{H}^*(H^{[1,2]} - z)^{-1}\tilde{H})_{12}\right)\left(h_{21} - z - (\tilde{H}^*(H^{[1,2]} - z)^{-1}\tilde{H})_{21}\right)}{\det(\hat{D})\left(h_{11} - z - (\tilde{H}^*(H^{[1,2]} - z)^{-1}\tilde{H})_{11}\right)} \\ =& \frac{\left(-h_{12} + z + (\tilde{H}^*(H^{[1,2]} - z)^{-1}\tilde{H})_{12}\right)\left(-h_{21} + z + (\tilde{H}^*(H^{[1,2]} - z)^{-1}\tilde{H})_{21}\right)}{\det(\hat{D})\left(h_{11} - z - (\tilde{H}^*(H^{[1,2]} - z)^{-1}\tilde{H})_{11}\right)} \\ =& \frac{G_{12}G_{21}}{G_{22}}. \end{aligned}$$

This verifies (5.3) and thus the first identity in (2.9) for $i = j$ which trivially implies the second identity.

Next, we establish (2.9) for $i \neq j$, i.e.

$$G_{ij}^{(\mathbb{T})} - G_{ij}^{(k\mathbb{T})} = G_{ik}^{(\mathbb{T})} G_{kj}^{(\mathbb{T})} \left(G_{kk}^{(\mathbb{T})}\right)^{-1} \tag{5.6}$$

for $i \neq j \neq k \neq i$ and $i, j, k \notin \mathbb{T}$. As befroe we may assume that $\mathbb{T} = \varnothing$ and $i = 1$, $j = 2$ and $k = 3$. We apply Lemma 5.1 with $n = 3$, $m = N - 3$ and $D = H - z$. Thus, $\hat{D} \in \mathbb{C}^{3\times 3}$ and

$$\begin{aligned} G_{rs} =& \quad (\hat{D}^{-1})_{rs} && \text{for } r, s \in \{1, 2, 3\}, \\ G_{rs}^{(3)} =& \quad ((\hat{D}^{(3)})^{-1})_{rs} && \text{for } r, s \in \{1, 2\}. \end{aligned}$$

To invert $\hat{D}^{(3)}$ we use above formula for the inverse of a 2×2 matrix. The inverse of $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ is given by

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{pmatrix}$$

with $\det A = aei + bfg + cdh - afh - bdi - ceg$. Using this formula we can calculate the inverse of the 3×3 matrix $\hat{D}$. These formulas for the inverses of $\hat{D}^{(3)}$ and $\hat{D}$ and a similar computation as in the proof of (5.3) yields (5.6).

Finally, we establish (2.10). We only prove the first identity, the second is proved completely analogously. Since $(H^{(\mathbb{T})} - z)G^{(\mathbb{T})} = 1$ we have

$$-G_{ii}^{(\mathbb{T})} \sum_k^{(\mathbb{T}i)} h_{ik} G_{kj}^{(\mathbb{T}i)} = -G_{ii}^{(\mathbb{T})} \sum_k^{(\mathbb{T}i)} h_{ik} G_{kj}^{(\mathbb{T})} + \sum_k^{(\mathbb{T}i)} h_{ik} G_{ki}^{(\mathbb{T})} G_{ij}^{(\mathbb{T})} = G_{ii}^{(\mathbb{T})} h_{ii} G_{ij}^{(\mathbb{T})} + (1 - h_{ii} G_{ii}^{(\mathbb{T})}) G_{ij}^{(\mathbb{T})} = G_{ij}^{(\mathbb{T})}$$

where we used (2.9) in the first step. This implies (2.10). □

The proof of (2.10) is taken from [5].

6 Proof of the Fluctuation Averaging

In this chapter, we verify the fluctuation averaging, i.e. Theorem 2.14 and Theorem 2.15. To this end, we transfer the proof of the fluctuation averaging given in [7] to our setting, i.e. to random matrices with the fourfold symmetry.

We start with several preparatory lemmas. The following result is the analogue of Lemma B.1 in [7].

Lemma 6.1. *Let Ψ be a deterministic control parameter satisfying $\Psi \geq N^{-C}$ and let $X(u)$ be nonnegative random variables such that for every $p \in \mathbb{N}$ there exists a constant c_p with $\mathbb{E}[X(u)^p] \leq N^{c_p}$ for all large N. If $X(u) \prec \Psi$ uniformly in u then*

$$\begin{aligned} \mathbb{E}_x X(u)^n &\prec \Psi^n \\ \mathbb{F}_x X(u)^n &\prec \Psi^n \\ \mathbb{E} X(u)^n &\prec \Psi^n \end{aligned}$$

uniformly in u and in x.

Before embarking on the proof we want to stress that $\Psi \geq N^{-C}$ is in particular fulfilled by deterministic control parameters satisfying (4.17) due to (2.3). Moreover, the last estimate is deterministic since $\mathbb{E}X^n \prec \Psi^n$ means by the definition of $\prec$ that for every $\varepsilon > 0$ there is $N_0 \in \mathbb{N}$ such that $\mathbb{E}X^n \leq N^\varepsilon \Psi^n$ for all $N \geq N_0$.

Proof. We check directly that the definition of $\prec$ is fulfilled. Since $X \prec \Psi$ implies $X^n \prec \Psi^n$ by Lemma 2.4 (v) it suffices to prove the claim for $n = 1$. Let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra. Fix $\varepsilon > 0$ and $D > 0$. For $p \in \mathbb{N}$ we have

$$\begin{aligned} \mathbb{P}(\mathbb{E}[X|\mathcal{F}] > N^\varepsilon \Psi) \leq & N^{-p\varepsilon}\Psi^{-p}\mathbb{E}[(\mathbb{E}[X|\mathcal{F}])^p] \leq N^{-p\varepsilon}\Psi^{-p}\mathbb{E}[X^p] \\ =& N^{-p\varepsilon}\Psi^{-p}(\mathbb{E}[X^p \mathbf{1}(X \leq N^{\varepsilon/2}\Psi)] + \mathbb{E}[X^p \mathbf{1}(X > N^{\varepsilon/2}\Psi)]) \\ \leq & N^{-p\varepsilon/2} + N^{-p\varepsilon + Cp}\sqrt{\mathbb{E}X^{2p}}\sqrt{\mathbb{P}(X > N^{\varepsilon/2}\Psi)} \\ \leq & N^{-p\varepsilon/2} + N^{-p\varepsilon + Cp + c_{2p}/2 - \tilde{D}/2} \leq 2N^{-2D} \leq N^{-D} \end{aligned}$$

where we applied (7.1) in the first step, (7.2) and (7.3) in the second step, $\Psi \geq N^{-C}$ in the fourth step and chose $p > 4D\varepsilon^{-1}$, $\tilde{D} > 4D + 2(C - \varepsilon)p + c_{2p}$ and $N_0 \in \mathbb{N}$ such that $\mathbb{P}(X > N^{\varepsilon/2}\Psi) \leq N^{-\tilde{D}}$ for $N \geq N_0$ in the sixth step. There is $N_1 \in \mathbb{N}$ which is independent of u and $\mathcal{F}$ such that $N_1 \geq N_0$ the last inequality holds for all $N \geq N_1$.

Choosing $\mathcal{F} = \{\varnothing, \Omega\}$ yields the last estimate, and the first estimate follows from taking $\mathcal{F} = \sigma(H^{(x,-x)})$. The last conclusion together with the assumption $X(u) \prec \Psi$ implies

$$|\mathbb{F}_x X(u)| \leq |X(u)| + |\mathbb{E}_x X(u)| \prec \Psi$$

uniformly in u and x. □

The previous Lemma will be applied to resolvent entries. Its applicability follows from the next statement which corresponds to Lemma B.2 in [7].

Lemma 6.2. *Let* $\mathbf{D}$ *be a spectral domain, and let* Ψ_o *and* Ψ *be deterministic control parameters satisfying* (4.17) *such that* $\Lambda \prec \Psi$ *and* $\Lambda_o \prec \Psi_o$*. For a fixed finite subset* $\mathbb{T} \subset \mathbb{N}$*,* $i \neq j$ *and* $i, j \notin \mathbb{T}$ *we have*

$$|G_{ij}^{(\mathbb{T})}| \prec \Psi_o, \quad \left|\frac{1}{G_{ii}^{(\mathbb{T})}}\right| \prec 1. \tag{6.1}$$

Furthermore, $|G_{ij}^{(\mathbb{T})}| \leq M$ *and for every* $n \in \mathbb{N}$ *and* $\varepsilon > 0$ *there is* $N_0 \in \mathbb{N}$ *such that*

$$\mathbb{E}|G_{ii}^{(\mathbb{T})}|^{-n} \leq N^{\varepsilon} \tag{6.2}$$

for all $N \geq N_0$*.*

Proof. The bounds in (6.1) follow from Lemma 4.2 with $\varphi = 1$ since $|G_{ij}^{(\mathbb{T})}| \prec \Lambda_o \prec \Psi_o$. The estimates $|G_{ij}^{(\mathbb{T})}| \leq \eta^{-1} \leq M$ have already been proved in (2.12) and (2.13).

In order to prove (6.2) we use Lemma 6.1 and check its assumptions: Applying Schur's complement formula (2.8) to $G_{ii}^{(\mathbb{T})}$ and using the triangle inequality we get

$$\begin{aligned}(\mathbb{E}|G_{ii}^{(\mathbb{T})}|^{-p})^{1/p} \leq & (\mathbb{E}|h_{ii}|^p)^{1/p} + |z| + \sum_{a,b}^{(\mathbb{T}i)} (\mathbb{E}|h_{ia}|^p |G_{ab}^{(\mathbb{T}i)}|^p |h_{bi}|^p)^{1/p} \\ \leq & \mu_p^{1/p} M^{-1/2} + |z| + M \sum_{a,b}^{(\mathbb{T}i)} (\mathbb{E}|h_{ia}|^{2p} \mathbb{E}|h_{bi}|^{2p})^{1/(2p)} \\ \leq & \mu_p^{1/p} M^{-1/2} + |z| + \mu_{2p}^{1/(2p)} N^2\end{aligned}$$

where we used (2.2), Lemma 2.2, (2.13) and the Cauchy-Schwarz inequality in the second step and $s_{ia}, s_{ib} \leq M^{-1}$ and (2.3) in the third step. Thus, $\mathbb{E}|G_{ii}^{(\mathbb{T})}|^{-p} \leq N^{2p+1}$ for all N large enough. Since $|G_{ii}^{(\mathbb{T})}|^{-1} \prec 1$ by (6.1) Lemma 6.1 with $\Psi = 1$ yields (6.2). □

The following Lemma which replaces (B.5) in [7] gives an auxiliary bound for estimating high moments of $|\sum_k t_{ik} \mathbb{F}_k G_{kk}^{-1}|$. In its proof we use the assumption $\mathbb{E}h_{xy}^2 = 0$ in the first estimate in (6.5) which is needed in (6.6).

Lemma 6.3. *Let* $\mathbf{D}$ *be a spectral domain. Suppose* $\Lambda \prec \Psi$ *and* $\Lambda_o \prec \Psi_o$ *for some deterministic control parameters* Ψ *and* Ψ_o *which satisfy* (4.17)*. Then for fixed* $p \in \mathbb{N}$ *we have*

$$\left|\mathbb{F}_x \left(G_{xx}^{(\mathbb{T})}\right)^{-1}\right| \prec \Psi_o \tag{6.3}$$

uniformly in $\mathbb{T} \subset \mathbb{N}$*,* $|\mathbb{T}| \leq p$*,* $x \notin \mathbb{T} \cup -\mathbb{T}$ *and* $z \in \mathbf{D}$*.*

Proof. Since $x, -x \notin \mathbb{T}$ we get as in the proof of (4.10) by using the resolvent identity (2.9) that

$$\sum_{a,b}^{(\mathbb{T}x)} h_{xa} G_{ab}^{(\mathbb{T}x)} h_{bx} = C_x^{(\mathbb{T})} + \sum_{a,b}^{(\mathbb{T}x,-x)} h_{xa} G_{ab}^{(\mathbb{T}x,-x)} h_{bx} + \left(G_{-x,-x}^{(\mathbb{T}x)}\right)^{-1} \sum_{a,b}^{(\mathbb{T}x,-x)} h_{xa} G_{a,-x}^{(\mathbb{T}x)} G_{-x,b}^{(\mathbb{T}x)} h_{bx} \tag{6.4}$$

where we used the definition

$$C_x^{(\mathbb{T})} := h_{x,-x} G_{-x,-x}^{(\mathbb{T}x)} h_{-x,x} + \sum_{a}^{(\mathbb{T}x,-x)} h_{xa} G_{a,-x}^{(\mathbb{T}x)} h_{-x,x} + \sum_{b}^{(\mathbb{T}x,-x)} h_{x,-x} G_{-x,b}^{(\mathbb{T}x)} h_{bx}.$$

The same argument as in the proof of Lemma 6.2 shows that the assumptions of Lemma 6.1 are fulfilled for each term of the expansion in (6.4).

Similar to the proof of (4.21) we get $|C_x^{(\mathbb{T})}| \prec M^{-1/2} \leq \Psi_o$ since Ψ_o satisfies (4.17). Using the first step in (4.24) and the argument in (4.25) we get

$$\begin{aligned}|\mathbb{F}_x \sum_{a,b}^{(\mathbb{T}x,-x)} h_{xa} G_{ab}^{(\mathbb{T}x,-x)} h_{bx}| \leq & |\sum_{a \neq b}^{(\mathbb{T}x,-x)} h_{xa} G_{ab}^{(\mathbb{T}x,-x)} h_{bx}| + |\sum_{a}^{(\mathbb{T}x,-x)} \Big(|h_{xa}|^2 - s_{xa}\Big) G_{aa}^{(\mathbb{T}x,-x)}| \\ \prec & \Psi_o + M^{-1/2} \prec \Psi_o\end{aligned}$$

where we used that Ψ_o fulfills (4.17). Adapting the proof of (4.29) we get

$$|\sum_{k,l}^{(x,-x)} h_{xk} G_{kl}^{(x,-x)} h_{l,-x}| \prec M^{-1/2} + \Psi_o \prec \Psi_o. \tag{6.5}$$

Hence, as $|G_{-x,-x}^{(\mathbb{T}x)}| \prec 1$ by Lemma 4.2 we get

$$|\Big(G_{-x,-x}^{(\mathbb{T}x)}\Big)^{-1} \sum_{a,b}^{(\mathbb{T}x,-x)} h_{xa} G_{a,-x}^{(\mathbb{T}x)} G_{-x,b}^{(\mathbb{T}x)} h_{bx}| \prec \Psi_o^2 \prec \Psi_o \tag{6.6}$$

using a similar representation as in (4.30). By Lemma 6.1 these estimates imply

$$|\mathbb{F}_x \sum_{a,b}^{(\mathbb{T}x)} h_{xa} G_{ab}^{(\mathbb{T}x)} h_{bx}| \prec \Psi_o.$$

Thus, we get the claim by applying Schur's complement formula (2.8) to $G_{xx}^{(\mathbb{T})}$ and observing that $|\mathbb{F}_x(h_{xx} - z)| = |h_{xx}| \prec M^{-1/2} \leq \Psi_o$ as h_{xx} is independent of $H^{(x,-x)}$ and $\mathbb{E}h_{xx} = 0$. $\square$

Next, we prove Theorem 2.15 by describing the changes needed to transfer the proof of Theorem 4.7 on pages 48 to 53 in [7] to its version for the fourfold symmetry.

First, we use Lemma 6.3 instead of (B.5). Moreover, we have to change some notions introduced in the proof of Theorem 4.7. In the middle of page 49, an equivalence relation on the set $\{1, \ldots, p\}$ is introduced which has to be substituted by the following equivalence relation. Starting with $\mathbf{k} := (k_1, \ldots, k_p) \in \{-N/2, \ldots, N/2\}^p$ and $r, s \in \{1, \ldots, p\}$ we define $r \sim s$ if and only if $k_r = k_s$ or $k_r = -k_s$. As in [7] the summation over all $\mathbf{k}$ is regrouped with respect to this equivalence relation and the notion of "lone" labels has to be understood with respect to this equivalence relation. We use the same notation $\mathbf{k}_L$ for the set of summation indices corresponding to lone labels. Differing from the definition in [7] we call a resolvent entry $G_{xy}^{(\mathbb{T})}$ with $x, y \notin \mathbb{T}$ *maximally expanded* if $\mathbf{k}_L \cup -\mathbf{k}_L \subset \mathbb{T} \cup \{x, y\}$. Correspondingly, we denote by $\mathcal{A}$ the set of monomials in the off-diagonal entries $G_{xy}^{(\mathbb{T})}$ with $\mathbb{T} \subset \mathbf{k}_L \cup -\mathbf{k}_L$, $x \neq y$ and $x, y \in \mathbf{k} \backslash \mathbb{T}$ (considering $\mathbf{k}$ as a subset of $\{-N/2, \ldots, N/2\}$) and the inverses of diagonal entries $1/G_{xx}^{(\mathbb{T})}$ with $\mathbb{T} \subset \mathbf{k}_L \cup -\mathbf{k}_L$ and $x \in \mathbf{k} \backslash \mathbb{T}$. With these alterations the algorithm can be applied as in [7]. In the proof of (B.15) the assertion $(*)$ has to be replaced by

$(*)$ For each $s \in L$ there exists $r = \tau(s) \in \{1, \ldots, p\} \backslash \{s\}$ such that the monomial $A_{\sigma_r}^r$ contains a resolvent entry with lower index k_s or $-k_s$.

To prove this claim, we suppose by contradiction that there is $s \in L$ such that $A_{\sigma_r}^r$ does not contain k_s and $-k_s$ as lower index for all $r \in \{1, \ldots, p\} \backslash \{s\}$. Without loss of generality we assume $s = 1$. This implies that each resolvent entry in $A_{\sigma_r}^r$ contains k_1 and $-k_1$ as upper index since $A_{\sigma_r}^r$ is maximally expanded for all $r \in \{2, \ldots, p\}$. Therefore, $A_{\sigma_r}^r$ is independent of k_1 as defined in Definition 2.13. Using (2.14) and proceeding as in [7] concludes the proof of $(*)$.

Following verbatim the remaining steps in the proof of Theorem 4.7 in [7] establishes the assertion of Theorem 2.15.

Now, we deduce Theorem 2.14 from Theorem 2.15. As we use (6.5) of (6.12) in the following proof of (2.19) this result also makes use of the assumption $\mathbb{E}h_{xy}^2 = 0$.

Proof of Theorem 2.14. The first estimate in (2.18) follows from Theorem 2.15 directly by setting $\Psi_o := \Psi$ and using $\Lambda_o \leq \Lambda \prec \Psi_o$.

To verify the second estimate in (2.18) we need the auxiliary estimate

$$|G_{xx}^{(\mathbb{T})} - m| \prec \Lambda$$

which can be proved by induction in the same way as (4.22). This bound implies

$$|\mathbb{F}_x G_{xx}^{(\mathbb{T})}| = |\mathbb{F}_x \left(G_{xx}^{(\mathbb{T})} - m \right)| \prec \Psi \tag{6.7}$$

by Lemma 6.1 since $|G_{xx}^{(\mathbb{T})} - m| \leq M + 1$. Now, following the proof of Theorem 2.15 verbatim with $\Psi_o := \Psi$ and replacing the usage of Lemma 6.3 by (6.7) yield the second estimate in (2.18).

Next, we establish (2.19). We manipulate Schur's complement formula (2.8) using (4.2) to get

$$\frac{1}{G_{xx}} = \frac{1}{m} + h_{xx} - \left(\sum_{k,l}^{(x)} h_{xk} G_{kl}^{(x)} h_{lx} - m \right). \tag{6.8}$$

Using Lemma 4.2 and the first estimate in (4.48) we get

$$\left| \frac{1}{G_{xx}} - \frac{1}{m} \right| = \left| \frac{G_{xx} - m}{G_{xx} m} \right| \prec |G_{xx} - m| \prec \Psi$$

Thus, $|h_{xx} - \left(\sum_{k,l}^{(x)} h_{xk} G_{kl}^{(x)} h_{lx} - m \right)| \prec \Psi$ as well. Therefore, we can expand the inverse of the right-hand side of (6.8) around $1/m$ which yields

$$v_x = G_{xx} - m = m^2 \left(-h_{xx} + \sum_{k,l}^{(x)} h_{xk} G_{kl}^{(x)} h_{lx} - m \right) + g_x \tag{6.9}$$

with error terms g_x such that $|g_x| \prec \Psi^2$ uniformly in x. By (4.10), (4.11), (4.14) and (4.15) we have the representation

$$\sum_{k,l}^{(x)} h_{xk} G_{kl}^{(x)} h_{lx} = \sum_a s_{xa} G_{aa} - A_x - B_x - s_{-x,x} G_{-x,-x}^{(x)} + Z_x + Y_x + C_x + s_{-x,x} G_{-x,-x}^{(x)}. \tag{6.10}$$

Using (6.9) we want to prove that

$$\mathbb{E}_x v_x = m^2 \sum_a s_{xa} v_a + f_x \tag{6.11}$$

where $|f_x| \prec \Psi^2$ uniformly in x. From (4.11) we get that the sum of the first four summands on the right-hand side of (6.10) is $H^{(x,-x)}$-measureable. Therefore, it suffices to show that all summands except the first on the right-hand side of (6.10) are bounded by Ψ^2 uniformly in x. For A_x and B_x this follows directly from their definitions in (4.12). Since $Z_x = \mathbb{F}_x X_x$ for some random variable X_x we get $\mathbb{E}_x Z_x = 0$. The representation (4.20) for C_x and Lemma 4.2 yield

$$|C_x| \prec M^{-1} + M^{-1/2} \Psi \prec \Psi^2$$

by (4.17).

The bound (6.6) with $\mathbb{T} = \varnothing$ gives

$$|Y_x| \prec M^{-1} \prec \Psi^2 \tag{6.12}$$

uniformly in x . This finishes the proof of (6.11).

Therefore, since $\mathbb{E}_x + \mathbb{F}_x = 1$ we have

$$\begin{aligned} w_a := \sum_x t_{ax} v_x =& \sum_x t_{ax} \mathbb{E}_x v_x + \sum_x t_{ax} \mathbb{F}_x v_x = m^2 \sum_{x,y} t_{ax} s_{xy} v_y + F_a \\ =& m^2 \sum_{x,y} s_{ax} t_{xy} v_y + F_a = m^2 \sum_x s_{ax} w_x + F_a \end{aligned} \tag{6.13}$$

where we used (6.11) with the notation $F_a := \sum_x t_{ax}(f_x + \mathbb{F}_x v_x)$ in the third step and in the fourth step that T and S commute. Note that $|F_a| \prec \Psi^2$ uniformly in a as $|\sum_x t_{ax} \mathbb{F}_x v_x| = |\sum_x t_{ax} \mathbb{F}_x G_{xx}| \prec \Psi^2$ by the second estimate in (2.18). Introducing the vectors $\mathbf{w} := (w_a)_{a=-N/2}^{N/2}$ and $\mathbf{F} := (F_a)_{a=-N/2}^{N/2}$ and writing (6.13) in matrix form we get

$$\mathbf{w} = m^2 S \mathbf{w} + \mathbf{F}.$$

Inverting the last equation yields

$$\mathbf{w} = (1 - m^2 S)^{-1} \mathbf{F}.$$

Recalling definition (2.17) and applying $\|\cdot\|_\infty$ to the last equation we have

$$\|\mathbf{w}\|_\infty \leq \Gamma \|\mathbf{F}\|_\infty \prec \Gamma \Psi^2$$

since $|F_a| \prec \Psi^2$ uniformly in a is equivalent to $\|\mathbf{F}\|_\infty \prec \Psi^2$. This proves (2.19). $\square$

7 Further Tools

In this chapter, we collect some well-known results used in the previous arguments. Thus, we hope to avoid any confusion which may be caused by citing a result just by a name.

7.1 Tools from Probability Theory

This section contains some results from probability theory, namely Chebyshev's inequality and some properties of the partial expectation.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space.

Lemma 7.1 (Chebyshev's inequality)**.** *If X is a nonnegative random variable and $\delta > 0$ then*

$$\mathbb{P}(X > \delta) \leq \delta^{-p}\mathbb{E}[X^p] \tag{7.1}$$

for all $p \in \mathbb{N}$.

Proof. Using the monotonicity of the integral we get

$$\delta^p\mathbb{P}(X > \delta) = \delta^p\mathbb{E}[\mathbf{1}_{\{X>\delta\}}] \leq \mathbb{E}[X^p]. \qquad \square$$

Let $X \in \mathscr{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ and let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra. If there is a $Y : \Omega \to \mathbb{R}$ such that Y is $\mathcal{F}$-measureable and that

$$\mathbb{E}[X\mathbf{1}_A] = \mathbb{E}[Y\mathbf{1}_A]$$

for all $A \in \mathcal{F}$ then Y is called *partial expectation* of X with respect to $\mathcal{F}$. Since Y always exists and is almost surely unique we define $\mathbb{E}[X|\mathcal{F}] := Y$. Note that the partial expectation is linear in X and that $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$ almost surely if $\mathcal{F} = \{\varnothing, \Omega\}$. Moreover, $\mathbb{E}[XZ|\mathcal{F}] = X\mathbb{E}[Z|\mathcal{F}]$ almost surely if X is $\mathcal{F}$-measureable and $Z \in \mathscr{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ such that XZ is integrable. Further properties of the partial expectation and proofs of the properties we stated can be found in [3].

Lemma 7.2 (Jensen's inequality for partial expectation)**.** *Let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra, $I \subset \mathbb{R}$ an interval, $\varphi : I \to \mathbb{R}$ a convex function and $X \in \mathscr{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ such that $X(\omega) \in I$ for all $\omega \in \Omega$. Then we have*

$$\varphi(\mathbb{E}[X|\mathcal{F}]) \leq \mathbb{E}[\varphi(X)|\mathcal{F}]. \tag{7.2}$$

This result is proved in Theorem 5.1.3 of [3].

Moreover, the partial expectation has the following property. If $X \in \mathscr{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ and $\mathcal{G} \subset \mathcal{F} \subset \mathcal{A}$ are sub-σ-algebras then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{F}] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] \tag{7.3}$$

almost surely [3].

The next proposition is a useful consequence of this property.

Proposition 7.3. *Let $X, Y \in \mathscr{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ with $XY \in \mathscr{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ and let $\mathcal{F} \subset \mathcal{A}$ be a sub-σ-algebra. If $\sigma(X)$ and $\sigma(Y, \mathcal{F})$ are independent, then*

$$\mathbb{E}[XY|\mathcal{F}] = \mathbb{E}[X]\mathbb{E}[Y|\mathcal{F}].$$

Proof. Using standard properties of the partial expectation we get

$$\mathbb{E}[XY|\mathcal{F}] = \mathbb{E}[\mathbb{E}[XY|\sigma(Y,\mathcal{F})]|\mathcal{F}] = \mathbb{E}[Y\mathbb{E}[X|\sigma(Y,\mathcal{F})]|\mathcal{F}] = \mathbb{E}[X]\mathbb{E}[Y|\mathcal{F}]. \qquad \square$$

7.2 Inequalities

Here, we state two simple inequalities for real numbers which have been frequently used in the foregoing proofs of several estimates.

Using the monotonicity of the square root function, we get

$$\sqrt{a+b} \le \sqrt{a+2\sqrt{a}\sqrt{b}+b} = \sqrt{a}+\sqrt{b} \tag{7.4}$$

for $a, b \geq 0$ where we used a binomial formula in the third step.

Similarly, $0 \le (a-b)^2/2$ yields Young's inequality

$$ab \le a^2/2 + b^2/2 \tag{7.5}$$

for $a, b \in \mathbb{R}$.

Bibliography

[1] G. Anderson, A. Guionnet, O. Zeitouni: *An Introduction to Random Matrices.* Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2010.

[2] Z. Bai, J. Silverstein: *Spectral Analysis of Large Dimensional Random Matrices.* Springer Series in Statistics, Springer, 2009.

[3] R. Durrett: *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010.

[4] L. Erdős: *Universality for random matrices and log-gases.* *ArXiv e-prints*, (2012), 1212.0839.

[5] L. Erdős, A. Knowles, H.-T. Yau, J. Yin: *Spectral Statistics of Erdős-Rényi Graphs II: Eigenvalue Spacing and the Extreme Eigenvalues.* *Communications in Mathematical Physics*, **314**(2012), no. 3, 587–640.

[6] L. Erdős, A. Knowles, H.-T. Yau, J. Yin: *Delocalization and Diffusion Profile for Random Band Matrices.* *Communications in Mathematical Physics*, **323**(2013), no. 1, 367–416.

[7] L. Erdős, A. Knowles, H.-T. Yau, J. Yin: *The local semicircle law for a general class of random matrices.* *Electron. J. Probab.*, **18**(2013), no. 59, 1–58.

[8] L. Erdős, B. Schlein, H.-T. Yau: *Universality of random matrices and local relaxation flow.* *Inventiones mathematicae*, **185**(2011), no. 1, 75–119.

[9] L. Erdős, H.-T. Yau: *Universality of local spectral statistics of random matrices.* *Bull. Amer. Math. Soc.*, **49**(2012), no. 3, 377–414.

[10] L. Erdős, H.-T. Yau, J. Yin: *Bulk universality for generalized Wigner matrices.* *Probability Theory and Related Fields*, **154**(2012), no. 1-2, 341–407.

[11] L. Erdős, H.-T. Yau, J. Yin: *Rigidity of eigenvalues of generalized Wigner matrices.* *Advances in Mathematics*, **229**(2012), no. 3, 1435 – 1515.

[12] T. Kemp: *Introduction to Random Matrices.* Lecture Notes, 2013. URL `http://www.math.ucsd.edu/~tkemp/247A/247A.Notes.pdf`.

[13] M. Mehta: *Random Matrices.* Pure and Applied Mathematics, Elsevier Science, 2004.

[14] T. Tao: *Topics in random matrix theory.* Graduate studies in mathematics ; v. 132, American Mathematical Society, Providence, R.I, 2012.

[15] E. P. Wigner: *Characteristic Vectors of Bordered Matrices With Infinite Dimensions.* *Annals of Mathematics*, **62**(1955), no. 3, 548–564.

[16] J. Wishart: *The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population.* *Biometrika*, **20A**(1928), no. 1-2, 32–52.