

UNIVERSAL EIGENVALUE STATISTICS FOR DYNAMICALLY DEFINED MATRICES

By

ARKA ADHIKARI AND MARIUS LEMM

Abstract. We consider dynamically defined Hermitian matrices generated from orbits of the doubling map. We prove that their spectra fall into the GUE universality class from random matrix theory.

1 Introduction

Eugene Wigner formulated the vision that the spectra of complex quantum systems are well described by the eigenvalues of random matrices. His work inspired the Wigner–Dyson–Mehta universality conjecture that the eigenvalues of matrices with independent entries are universal in the sense that they only depend on the symmetry type of the matrix—Hermitian (GUE) or real symmetric (GOE). The Wigner–Dyson–Mehta universality conjecture was a driving force of random matrix theory and was finally resolved around 10 years ago in a series of groundbreaking works [21, 22, 27, 28, 40]. A major focus of random matrix theory in recent years has been to broaden the reach of this vision to include various models of “not-too-random” matrices whose entries have dependencies. Among the main venues of research are the study of probabilistically generated correlations [1, 4, 5, 9, 20, 36] and graph-theoretical constraints on random adjacency matrices [10, 11, 34]. The motivation to move towards increasingly structured ensembles is the widely held belief that the universality phenomenon encompasses many other strongly correlated systems than traditional random matrices. This belief is based on many real-world examples of strongly correlated point processes which empirically reproduce random matrix statistics. Famous examples include the BGS quantum chaos conjecture and Montgomery’s pair correlation conjecture for the zeros of the Riemann zeta function. It is a major open problem to explain the scope of universal random matrix statistics.

In our ongoing research program [2, 3], we turned to dynamical systems theory as a method for producing matrices with dependencies. We consider **dynamically defined** matrices which are generated by evaluating a complex-valued function f along orbits of an ergodic transformation T and then using the sequence $f(x), f(Tx), f(T^2x), \dots$ to fill a rectangular array, e.g., as in (2.1) below. As is common in dynamical systems theory, all the randomness then comes from sampling the starting position x .

We briefly recall that mathematical physicists have long studied the spectral theory of dynamically defined matrices of Schrödinger type. These matrices are typically tridiagonal, thus sparse. In the world of Schrödinger operators (or, more generally, Jacobi operators), dynamical systems theory and spectral theory are intimately connected in deep and sometimes surprising ways. See [7, 13, 14, 19, 35] for examples of breakthrough results in this area. One finding has been that sufficiently random-like Schrödinger operators display Anderson localization and also Poisson spectral statistics [39]. Hence, roughly speaking, it can be said that the sparse dynamically defined Schrödinger operators tend to belong to the opposite of the random matrix regime.

We take a new perspective and consider dynamically defined Hermitian matrices that are full (i.e., all entries are of comparable size). This puts us in the world of random matrix theory, but with a novel kind of correlation structure. Our goal is then to establish that these dynamically defined matrices display random matrix statistics down to the smallest scale, i.e., that dynamical correlations can mimic random ones.

The main result of this work can be summarized as follows.

The GUE universality class contains dynamically defined matrices.

We emphasize that a single entry $f(T^kx)$ of our dynamically generated matrix (defined in (2.1) below) fully determines all other entries, so in some sense there is complete deterministic dependence within the matrix and this fact places the model outside of the existing techniques. From a dynamical systems perspective, it is nonetheless clear that any strongly mixing dynamical system will lead to rapid decorrelation of the entries. The latter perspective is fruitful for us, but we have to make precise how the dynamical decorrelation of matrix entries manifests on the spectral level and this requires some new techniques.

Naturally, the precise manner of dynamical decorrelation depends on the underlying dynamical system. In the prior works [2, 3], we have focused on extremely rigid dynamical systems of skew-shift type which are heavily correlated. Accordingly, their spectra are difficult to analyze down to small spectral scales and universality remains open. This is also reflected in the Schrödinger world, where the skew-shift dynamics are also a significant frontier [14, 32, 37].

The present paper starts from a natural strongly mixing dynamical system, the doubling map $T : [0, 1] \rightarrow [0, 1]$ defined by

$$(1.1) \quad T(x) = 2x \pmod{1}.$$

For this map, we can derive the desired strong conclusion, universal eigenvalue statistics.

Existing techniques on correlated matrices come in two flavors: either they use a special model-dependent correlation structure (e.g., in the random graph setting) [10, 11, 34] or they require assumptions on the correlation decay [5, 6, 16, 26] that are not fulfilled in the dynamical setting for the reason mentioned above that a single matrix entry fully determines all other ones. We address this issue by using a resampling trick and by adapting techniques developed for correlated matrices with finite range of dependence [16] to a logarithmic range of dependence. In contrast to most of the above works, we can reduce to a scalar-valued self-consistent equation by suitably reorganizing the self-consistent equations arising from the dynamical covariance structure, cf. Section 4.2.

From the random matrix theory perspective, our result broadens the scope of the GUE universality class which is seen to encompass dynamically defined matrices. From a mathematical physics perspective, the result provides a spectral-theoretic confirmation of the quasi-random nature of the doubling map; this can be seen as a “delocalization analog” of a well-known result by Bourgain–Schlag [14] establishing Anderson localization for potentials dynamically defined via strongly mixing potentials; see also [18].

We leave it as an open problem to extend the result presented here to a wider class of dynamical systems satisfying a quantitative mixing assumption, such as appropriate subshifts of finite type described by the potential formalism of Bowen [15]. Our result opens the door to developing a more general classification theory which would link properties of the underlying dynamical systems to the emergent spectra of dynamically defined full matrices. It would be interesting to see which features of such a theory (if any) mirror existing developments in the world of Schrödinger operators, especially given that those model the qualitatively different localization regime.

The paper is organized as follows.

- In Section 2, we define our model of dynamically defined matrices and list our main results on a local law for the Green’s function, the universality of the eigenvalue statistics, and the delocalization of eigenvectors.
- In Section 3, we introduce our resampling method, list analogous universality results for our resampled matrices, and show that these imply the main results of Section 2.

- In Section 4, we briefly review the central role of the self-consistent equation for the Stieltjes’ transform of our resampled matrices. We then show that our model has a scalar self-consistent equation, and we derive this scalar self-consistent equation from an associated matrix self-consistent equation.
- In Section 5, we derive the matrix self-consistent equation via appropriate concentration estimates.
- In Section 6, we prove the stability of the matrix self-consistent equation and derive associated error bounds for the Stieltjes’ transform of our resampled matrix model, as well as the Green’s function of our resampled matrix model.
- In Section 7, we use our previously derived stability bounds and error estimates to prove a local law via a standard inductive scheme.
- In Section 8, we compare our resampled matrix model to a model with a small Gaussian part via an Ornstein–Uhlenbeck process. This allows us to prove our main universality result.

2 Model and main result

2.1 Dynamically defined matrices. Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denote the standard torus. Given an evaluation (or sampling) function $f : \mathbb{T} \rightarrow \mathbb{C}$ and a fixed starting point $x \in [0, 1]$, we consider the **dynamically defined matrix**

$$X_{ij} = \frac{1}{\sqrt{N}} f(T^{(2N-1)i+j} x),$$

that is,

$$(2.1) \quad X = \frac{1}{\sqrt{N}} \begin{pmatrix} f(Tx) & f(T^2x) & \dots & f(T^N x) \\ f(T^{2N+1}x) & f(T^{2N+2}x) & \dots & f(T^{3N}x) \\ f(T^{4N+1}x) & f(T^{4N+2}x) & \dots & f(T^{5N}x) \\ \vdots & \vdots & \ddots & \vdots \\ f(T^{2N^2-N+1}x) & f(T^{2N^2-N+2}x) & \dots & f(T^{2N^2}x) \end{pmatrix}.$$

Observe that all the entries of X belong to a single dynamical orbit—the orbit of the starting point $x \in [0, 1]$ under the doubling map T . The normalization factor $\frac{1}{\sqrt{N}}$ is standard and ensures that the limiting spectral distribution is supported on an order-1 set. For convenience, we start at the top left with Tx instead of x and we skip N discrete time steps when moving from one row to the next. The method and results of this paper straightforwardly extend to the case where no discrete time steps are skipped. This only requires minor modifications, which we skip here to keep the paper short and focused on the main points of interest. These modifications are described in Remark 2.1 (ii) below.

Our results concern the (real-valued) spectrum of the following Hermitization H_X of X :

$$(2.2) \quad H_X = \begin{pmatrix} 0 & X \\ X^\dagger & 0 \end{pmatrix}$$

with X^\dagger the adjoint of X .

To turn this into a random matrix ensemble, we sample the starting position $x \in [0, 1]$ from the uniform measure on $[0, 1]$, which we call \mathbb{P} . We write \mathbb{E} for the associated expectation. The uniform measure is a natural choice because it is the equilibrium measure of the doubling map.

We now state an informal version of the main result. See Theorem 2.10 below for the formal version.

Theorem (Main result, informal version). *Suppose that f is an admissible evaluation function in the sense of Definition 2.2. Then, for every $k \geq 1$, the k -point correlation functions of the eigenvalues of H_X converge to the k -point correlation functions of the eigenvalues of an $N \times N$ GUE matrix as $N \rightarrow \infty$.*

The function $f(x) = \exp(2\pi i x)$ is an example of an admissible evaluation function in the sense of Definition 2.2.

The result can be rephrased to describe the singular values of X if desired. Generalization to non-square matrices X is straightforward.

Remark 2.1 (Choice of model). (i) Let us explain the reason why it is convenient to skip N discrete time steps in (2.1). The fact that the doubling map is exponentially mixing implies that, as we make precise later, the matrix entries are logarithmically correlated. More precisely, the correlation is negligible after roughly $O((\log_2 N)^6)$ discrete time steps. In the model (2.1), this means different rows are effectively independent (and they are fully independent for the modified ensemble Y from (3.2)). This makes the derivation of the self-consistent equation for the Stieltjes transform of the Green's function (which is at the heart of proving the local law) particularly clean because the self-consistent equation becomes scalar, cf. Section 4.2 below.

(ii) Alternatively, we could also consider the matrix

$$(2.3) \quad \frac{1}{\sqrt{N}} \begin{pmatrix} f(Tx) & f(T^2x) & \dots & f(T^N x) \\ f(T^{N+1}x) & f(T^{2N+2}x) & \dots & f(T^{2N}x) \\ f(T^{2N+1}x) & f(T^{4N+2}x) & \dots & f(T^{3N}x) \\ \vdots & \vdots & \ddots & \vdots \\ f(T^{N^2-N+1}x) & f(T^{N^2-N+2}x) & \dots & f(T^{N^2}x) \end{pmatrix}$$

where there is no skipping of steps between rows. Let us explain how the method can be adapted to this model with only minor complications. First, we define an analogous modified ensemble Y similar to (3.2). Next, notice that now only the first $(\log_2 N)^6$ many entries of a row are correlated with the last $(\log_2 N)^6$ entries of the previous row. Observe that the “bad” entries are all located in the same last $(\log_2 N)^6$ columns and the fraction of bad columns rapidly vanishes as $N \rightarrow \infty$. Therefore, we can remove them by hand to effectively break the correlation between rows. To implement this, we set the last $(\log_2 N)^6$ columns to 0. By eigenvalue interlacing, this changes the local law by at most a term $\frac{(\log_2 N)^6}{N\eta}$, which is a negligible error term. After setting the last $(\log_2 N)^6$ columns to 0, the resulting matrix will have completely independent rows. The self-consistent equation is then seen to be scalar and can be analyzed with the same methods used in this paper used for (2.1) with very minor modifications. The exact correlation structure is now given by a Toeplitz matrix rather than a repeated block of Toeplitz matrices. Deriving lower bounds on the smallest eigenvalue of the pure Toeplitz matrix would also follow from the same methods as used for the block Toeplitz matrices in this paper. Once the local law is established, the usual 3-step strategy yields universality. We leave the remaining details to the interested reader and focus on (2.1) in the following.

2.2 Admissible evaluation functions. The following assumption on the evaluation function f arises naturally in the proof.

We require $f \in C^2(\mathbb{T})$ and define its Fourier coefficients by

$$(2.4) \quad c_k = \int_0^1 f(y) e^{-2\pi i k y} dy, \quad k \in \mathbb{Z}.$$

Then we associate to f the function $g_f : \mathbb{T} \rightarrow [0, \infty)$ given by

$$(2.5) \quad g_f(x) = \sum_{\substack{n \geq 0: \\ 2 \nmid n}} \left| \sum_{k=0}^{\infty} c_{n2^k} \exp[2\pi i k x] \right|^2$$

where $2 \nmid n$ means that 2 is not a divisor of n . To see that $g_f(x)$ is finite, note that $g_f(x) \leq \|c\|_{\ell^1}^2 \leq C \|f\|_{H^1}^2$ with $c = (c_k)_{k \in \mathbb{Z}}$, where the last estimate uses the Cauchy–Schwarz inequality.

Definition 2.2 (Admissible evaluation function). We say that $f \in C^2(\mathbb{T})$ is an **admissible evaluation function** if $\mathbb{E}[f] = 0$ and there exists a constant $g_{\min} > 0$ such that

$$(2.6) \quad \inf_{x \in [0, 1]} g_f(x) \geq g_{\min} > 0.$$

Example 2.3. The functions $f(x) = e^{2\pi i x}$ and $f(x) = \cos(2\pi x)$ are admissible evaluation functions.

Remark 2.4. A simple sufficient condition for $f \in C^2(\mathbb{T})$ to be an admissible evaluation function can be obtained by ignoring cancellations and requiring some concentration of Fourier modes along dyadic scales. More precisely, if there exists an integer $n \geq 1$ such that $2 \nmid n$ and

$$(2.7) \quad |c_n| > \sum_{k=1}^{\infty} |c_{n2^k}|,$$

then f is an admissible evaluation function.

To clarify the meaning of the function g_f , we express it through the following function $\phi : \mathbb{Z} \rightarrow \mathbb{C}$ which measures the correlation between f and \bar{f} among different dyadic scales:

$$(2.8) \quad \phi(j) = \mathbb{E}[f(x)\overline{f(T^j x)}] = \sum_{k=1}^{\infty} \overline{c_k} c_{k2^j}, \quad j \in \mathbb{Z}_+ \cup \{0\},$$

and we set $\phi(-j) = \overline{\phi(j)}$. We can use ϕ to define the infinite Toeplitz matrix Φ by

$$(2.9) \quad \Phi_{i,j} = \phi(i-j), \quad i, j \in \mathbb{Z}.$$

Observe that Φ is Hermitian. We will see below that Φ determines the limiting spectral distribution of H_X as $N \rightarrow \infty$.

By expanding the square in (2.5), we can now rewrite g_f as a Fourier series associated to the correlation function ϕ , i.e.,

$$(2.10) \quad g_f(x) = \sum_{j \in \mathbb{Z}} \phi(j) e^{2\pi i j x} = \phi(0) + 2\operatorname{Re} \sum_{j=1}^{\infty} \phi(j) e^{2\pi i j x}.$$

One says that formula (2.10) represents the **Fourier symbol** of the infinite Toeplitz matrix Φ . In particular, $\inf_x g_f(x) = \inf \operatorname{spec} \Phi$. We see that Assumption (2.6) is equivalent to the spectral condition that Φ is strictly positive definite. This is how the function g_f arises in the proof.

2.3 First result: local law. We begin by stating a **local law** which shows that, as $N \rightarrow \infty$, the empirical spectral distribution of H converges to a deterministic limiting density even at small scales. The local law is a key prerequisite to our main result, universality.

The local law is conveniently formulated through the Stieltjes transform. Given a measure $d\mu$ on the real axis, its Stieltjes transform m_μ is defined by

$$m_\mu(z) = \int_{\mathbb{R}} \frac{1}{x - z} d\mu(x), \quad \text{Im } z > 0.$$

We prove a formulation of the local law in which the Stieltjes transform of H_X converges to $m_\infty(z)$ which solves the following self-consistent equation

$$(2.11) \quad \int_0^1 \frac{1}{g_f(x)m_\infty(z) + z} dx = -m_\infty(z).$$

Proposition 2.5. *Let $\text{Im } z > 0$. Then, there exists a unique solution $m_\infty(z)$ to (2.11) with positive imaginary part.*

This proposition is proved in Appendix A. To state the local law, we introduce the following notation. Given $\varepsilon, \delta > 0$, we let

$$\mathcal{D}_{\varepsilon, \delta} = \left\{ E + i\eta \in \mathbb{C} : \rho_\infty(E) \geq \varepsilon, \eta \in (N^{-1+\delta}, 1) \right\}$$

where the first condition $\rho_\infty(E) \geq \varepsilon$ puts us in the bulk of the spectrum. We say that a sequence of events A_N holds **almost surely** as $N \rightarrow \infty$, if $\mathbb{P}(A_N) \rightarrow 1$ as $N \rightarrow \infty$, where we recall that \mathbb{P} is the uniform measure for $x \in [0, 1]$.

Theorem 2.6 (Local law). *Let $f \in C^2(\mathbb{T})$ be an admissible evaluation function. Then, for every $\varepsilon, \delta > 0$,*

$$(2.12) \quad \sup_{z \in \mathcal{D}_{\varepsilon, \delta}} \left| \frac{1}{2N} \text{Tr}(H_X - z)^{-1} - m_\infty(z) \right| \leq \frac{N^\delta}{N\eta}$$

holds almost surely as $N \rightarrow \infty$.

Note that the right-hand side in (2.12) vanishes in the limit as long as $\eta \gg N^{-1+\delta}$ with $\delta > 0$ arbitrary, so the local law descends arbitrarily close to the optimal scale N^{-1} where individual eigenvalues become resolved.

We can identify $m_\infty(z)$ as the Stieltjes transform of a limiting spectral distribution which we call $\rho_\infty(x)$. This $\rho_\infty(x)$ is a deterministic probability density function which generalizes the Wigner semicircle law $\rho_{\text{sc}}(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbb{1}_{[-2, 2]}(x)$ to our correlated setting (see Example 2.8 below).

Proposition 2.7. *The function $m_\infty(z)$ is the Stieltjes transform of a continuous measure $\rho_\infty(x)dx$, i.e.,*

$$(2.13) \quad m_\infty(z) = \int_{\mathbb{R}} \frac{1}{x - z} \rho_\infty(x) dx, \quad \text{Im } z > 0.$$

By the Stieltjes inversion formula,

$$\rho_\infty(E) = \lim_{\eta \rightarrow 0} \frac{1}{\pi} \operatorname{Im} m_\infty(E + i\eta), \quad E \in \mathbb{R}.$$

The proof of this proposition is deferred to Appendix A. Notice that the density $\rho_\infty(x)$ is deterministic and only depends on the choice of evaluation function f .

Example 2.8. For the evaluation function $f(x) = \exp(2\pi i x)$, we have $g_f(x) = 1$ and $\rho_\infty(x) = \rho_{\text{sc}}(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbb{1}_{[-2,2]}(x)$, the Wigner semicircle law. See Figure 1 for a pictorial representation of the emergent global law in this special case of Theorem 2.6.

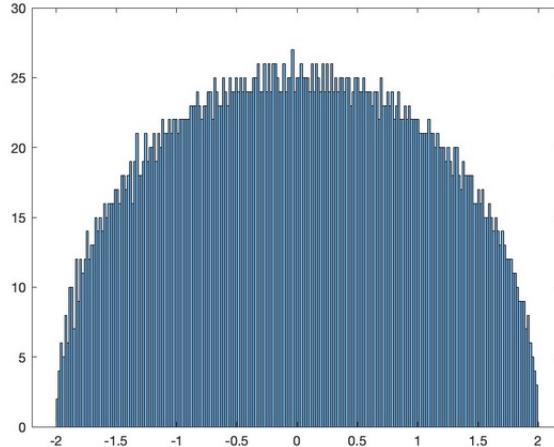


Figure 1. The empirical spectral distribution of matrices given by (2.1) with $f(x) = \exp(2\pi i x)$. This histogram depicts the emergence of the Wigner semicircle law on the global scale. Theorem 2.6 also establishes that this convergence continues to hold down to small scales.

Remark 2.9. Let us explain the form of the self-consistent equation (2.11). In the proof, it arises in the form

$$(2.14) \quad -\frac{1}{m_\infty(z)} m_\Phi\left(-\frac{z}{m_\infty(z)}\right) = m_\infty(z).$$

where $m_\Phi \equiv m_{\rho_\Phi}$ is the Stieltjes transform of the spectral density of the infinite Toeplitz matrix Φ . Since Φ is unitarily equivalent to multiplication by g_f in Fourier space, we have the explicit representation

$$m_\Phi(z) = \int_0^1 \frac{1}{g_f(x) - z} dx, \quad \operatorname{Im} z > 0,$$

which connects (2.14) to (2.11).

2.4 Main result: universality. We are now ready to state our main universality result. We write $\lambda_1, \dots, \lambda_{2N}$ for the eigenvalues of H_X and $p_N^N : \mathbb{R}^{2N} \rightarrow \mathbb{R}$ for the symmetrized eigenvalue density of H_X . The k -point correlation functions p_N^k are defined by

$$p_N^k(\lambda_1, \dots, \lambda_k) = \int p_N^N(\lambda_1, \dots, \lambda_N) d\lambda_{N-k+1} \dots d\lambda_N, \quad k \geq 1.$$

The universal objects are $p_{N,\text{GUE}}^{(k)}$, defined analogously as the k -point correlation functions of a matrix from the Gaussian Unitary Ensemble. They are explicitly computable, e.g.,

$$p_{N,\text{GUE}}^{(k)}(x_1, \dots, x_k) = \det(K_N(x_i, x_j))_{1 \leq i, j \leq k}$$

where $K_N(x, y)$ has an explicit formula in terms of Hermite polynomials [29].

We use the convention that for $\alpha = (\alpha_1, \dots, \alpha_k) \in \mathbb{R}^k$,

$$E + \alpha = (E + \alpha_1, \dots, E + \alpha_k).$$

Theorem 2.10 (Main result: Universality). *Let $f \in C^4(\mathbb{T})$ be an admissible evaluation function. Let $k \geq 1$ and let $O : \mathbb{R}^k \rightarrow \mathbb{R}$ be smooth and compactly supported. Then, for $\varepsilon > 0$ and $\rho(E) \geq \varepsilon$, there exists $\kappa > 0$ such that*

$$(2.15) \quad \int_{\mathbb{R}^k} O(\alpha) \left[p_N^k \left(E + \frac{\alpha}{2N\rho_\infty(E)} \right) - p_{N,\text{GUE}}^{(k)} \left(E + \frac{\alpha}{2N\rho_{\text{sc}}(E)} \right) \right] d^k \alpha = O(N^{-\kappa})$$

Remark 2.11. (i) The condition $\rho_f(x) \geq \varepsilon$ means that we work in the bulk of the spectrum. It may be possible to prove similar results at the spectral edge, but this is not our focus here.

- (ii) While the results are stated as holding almost surely as $N \rightarrow \infty$, the proof quantifies the convergence speed as polynomial. Since this point is not our focus here, we use the qualitative notion for simplicity of presentation.
- (iii) There exist several universality results for correlated matrices in the literature [6, 16, 26]. However, the stated results do not apply to our dynamically defined model. For instance, the condition on correlation decay in [6, 26] requires that for any pair of bounded functions F and G applied to entries H_{ij} and H_{kl}

$$(2.16) \quad \mathbb{E}[F(H_{ij})G(H_{kl})] \leq \frac{C_\nu \|F\|_\infty \|G\|_\infty}{1 + (|i - k| + |j - l|)^\nu}, \quad \nu \in \mathbb{Z}_+.$$

Now, in our model, fixing i, j, k, l and taking $F = \text{id}$ and $G = T^p$ for an appropriate power p , the left-hand side equals $\mathbb{E}[|H_{kl}|^2] = N^{-1}$, a constant.

Since this does not decay at all, the correlation condition (2.16) breaks down in our model. The reason is that very general test functions F and G need to be controlled, which include digit shifts. A similar argument applies to the post-resampling matrix Y defined in (3.2) below; taking the power p to be $\lfloor (\log_2 N)^6 \rfloor$ the correlation again does not decay. So the assumptions of the results in [6, 26] are also not satisfied for Y . The matrix Y satisfies (2.16) with $C_v = (\log_2 N)^{6v}$. We have been kindly informed by researchers closely familiar with [6] that by tracing the proof of the local law in that work such a logarithmic dependence of the constant can be accommodated. We elect here to go via the rather direct route of studying the much simpler scalar self-consistent equation instead.

Remark 2.12 (Open problem). The formulae (2.1) and (2.2) defining our matrix ensemble can be generalized to dynamically define a matrix ensemble starting from any dynamical system (X, T, μ) . It is an interesting open problem, lying on the interface of dynamical systems theory and random matrix theory, to extend the present universality result to a much broader class of dynamical systems. Natural candidates are dynamical systems satisfying a quantitative mixing assumption, e.g., subshifts of finite type building on the potential formalism developed in [15]; see also [8].

2.5 Delocalization of eigenvectors. In physics, the delocalization of eigenvectors is seen as a hallmark of the random matrix theory phase. In a standard way, the precise control on Green's function established in proving Theorem 2.6 implies eigenvector delocalization bounds in the bulk of the spectrum.

We use the convention that eigenvectors $u_\alpha \in \mathbb{R}^N$ are ℓ^2 -normalized and so a completely delocalized eigenvector, which is roughly equally supported on all coordinates, would have $\max_{1 \leq i \leq N} |u_\alpha(i)|^2 \sim N^{-1}$.

Theorem 2.13 (Delocalization bound for eigenvectors). *Let $H_X u_\alpha = Eu_\alpha$ for some $E \in \mathbb{R}$ satisfying $\rho_\infty(E) \geq \epsilon$ for some $\epsilon > 0$. Then, there exists $\delta > 0$ such that*

$$\max_{1 \leq i \leq N} |u_\alpha(i)|^2 \leq N^{\delta-1}$$

holds almost surely as $N \rightarrow \infty$.

2.6 Proof strategy. Overall, our proof relies on the famous 3-step strategy for proving universality developed in the past years. However, the dynamical nature of correlations in our matrix ensemble requires some new ideas which can be summarized as follows.

- The first novel feature of our proof is a **dynamical resampling trick** which we use to “preprocess” the dynamically defined ensemble H_X into another one, called H_Y , whose entries become independent after $(\log N)^6$ many discrete time steps. The resampling procedure is based on the binary expansion because the doubling map then acts as a digit shift. Importantly, the construction is such that the Green’s functions before and after resampling agree on scales $\text{Im}[z] \ll N^{-1}$. This means that their spectra are essentially the same and so it suffices to prove local law and universality for the ensemble H_Y with its logarithmic range of dependence.
- In a second step, we prove local law and universality for the eigenvalues of H_Y with its logarithmic range of dependence by extending Che’s analysis for finite range of dependence [16]. The possibility that these methods could be extended to a logarithmic range of dependence to be logarithmic, even to a logarithmic power of a logarithm, was already raised by Che.
- A problem compared to the situation studied in [16] is that we no longer have strict positivity of the correlation matrix and Lipschitz continuity of the correlations. We address these issues by leveraging the block structure and symmetries to reduce to a scalar self-consistent equation for the Stieltjes transform. For this scalar equation, we can recover positive definiteness by using the Ward identity, the assumption that f is an admissible evaluation functions and techniques in the spectral theory of Toeplitz matrices, e.g., turning a banded Toeplitz matrix into a circulant matrix (whose spectrum is explicitly computable) by a finite-rank perturbation.
- Other building blocks of the proofs, e.g., the fast equilibration of Dyson Brownian Motion, are standard in the field and are accordingly only summarized briefly.

We use the convention that C denotes a generic positive constant that is independent of N and whose value may change from line to line.

3 Step I: Resampling to logarithmic range of dependence

3.1 Resampling binary digits. Recall that x is the starting point of the doubling map dynamics in (2.1). We write

$$x = \sum_{n \geq 1} d_n 2^{-n}, \quad d_n \in \{0, 1\},$$

for its binary expansion. The doubling map acts as a digit shift in the binary basis,

$$T^k x = \sum_{n \geq 1} d_{n+k} 2^{-n}, \quad k \geq 0.$$

We let $\{b_n^k\}_{n,k \geq 1}$ be a two-parameter family of independent coin flips with outcomes $\{0, 1\}$ occurring with probability $\frac{1}{2}$ each. We then define the resampling of $T^k x$ by

$$(3.1) \quad y_k = \sum_{1 \leq n \leq (\log N)^6} d_{n+k} 2^{-n} + \sum_{n \geq (\log N)^6 + 1} b_n^k 2^{-n}, \quad \text{where } \log N \equiv \log_2 N.$$

(The choice of 6 is somewhat arbitrary; any sufficiently large power of a logarithm will do.)

The resampled matrix ensemble is given by

$$(3.2) \quad Y = \frac{1}{\sqrt{N}} \begin{pmatrix} f(y_1) & f(y_2) & \dots & f(y_N) \\ f(y_{2N+1}) & f(y_{2N+2}) & \dots & f(y_{3N}) \\ \vdots & \vdots & \ddots & \vdots \\ f(y_{2N^2-N+1}) & f(y_{2N^2-N+2}) & \dots & f(y_{2N^2}) \end{pmatrix}$$

with Hermitization

$$(3.3) \quad H_Y = \begin{pmatrix} 0 & Y \\ Y^\dagger & 0 \end{pmatrix}.$$

3.2 Local law and universality for the resampled ensemble. The advantage of the resampled ensemble H_Y is that in contrast to the original ensemble H_X it only has logarithmic range of dependence: for $|a - b| > (\log N)^6$, the random variables y_a and y_b are independent. In Step 2, we will use this fact to prove that the analogs of our main results (local law and universality) indeed hold for H_Y . For completeness, they are stated here.

Theorem 3.1 (Local law after resampling). *Let $f \in C^2(\mathbb{T})$ be an admissible evaluation function. Then, for every $\varepsilon, \delta > 0$,*

$$(3.4) \quad \sup_{z \in \mathcal{D}_{\kappa, \delta}} \left| \frac{1}{2N} \text{Tr}(H_Y - z)^{-1} - m_\infty(z) \right| \leq \frac{N^\delta}{N\eta}$$

holds almost surely as $N \rightarrow \infty$.

We write \tilde{p}_N^k for the k -point correlation function associated to the eigenvalues of H_Y .

Theorem 3.2 (Universality after resampling). *Let $f \in C^4(\mathbb{T})$ be an admissible evaluation function. Let $k \geq 1$ and let $O : \mathbb{R}^k \rightarrow \mathbb{R}$ be smooth and compactly supported. Then, for $\varepsilon > 0$ and $\rho(E) \geq \varepsilon$, there exists $\kappa > 0$ such that*

$$(3.5) \quad \int_{\mathbb{R}^k} O(\alpha) \left[\tilde{p}_N^k \left(E + \frac{\alpha}{2N\rho_\infty(E)} \right) - p_{N,\text{GUE}}^k \left(E + \frac{\alpha}{2N\rho_{\text{sc}}(E)} \right) \right] d^k \alpha = \mathcal{O}(N^{-\kappa})$$

as $N \rightarrow \infty$.

Before we prove these theorems, we use them to derive Theorems 2.6 and 2.10.

3.3 Green's function comparison lemma. Here and in the following we denote

$$G_X(z) = (H_X - z)^{-1}, \quad G_Y(z) = (H_Y - z)^{-1}, \quad \text{Im } z > 0$$

and

$$m_X(z) = \frac{1}{2N} \text{Tr} G_X(z), \quad m_Y(z) = \frac{1}{2N} \text{Tr} G_Y(z).$$

From now on, we always assume that $f \in C^2(\mathbb{T})$ is an admissible evaluation function.

Lemma 3.3 (Green's Function Comparison Lemma). *Let $n \geq 1$, $\varepsilon > 0$ and let E_1, \dots, E_n satisfy $\rho_\infty(E_i) \geq \varepsilon$ and ρ_∞ is the measure associated to m_∞ . Given $\sigma_1, \dots, \sigma_n \leq \sigma$, we set*

$$z_j = E_j + i\eta_j, \quad \eta_j = N^{-1-\sigma_j}.$$

Then, there exists $C_\sigma > 0$ depending only on σ such that

$$\left| \prod_{k=1}^n \text{Im } m_X(z_k) - \prod_{k=1}^n \text{Im } m_Y(z_k) \right| \leq \frac{C_\sigma}{N}.$$

Proof of Lemma 3.3. Let $\sigma > 0$. By the resolvent identity, we have that

$$(3.6) \quad G_X(z) - G_Y(z) = G_Y(z)(X - Y)G_X(z).$$

The entries of the matrix $(X - Y)$ are of the form $f(T^k x) - y_k$. For these, we have the bound

$$(3.7) \quad |f(T^k x) - f(y_k)| \leq \|f'\|_\infty |T^k x - y_k| \leq \|f'\|_\infty 2^{-(\log N)^6}.$$

The outside factors are controlled by $|G_{X,ij}(z_k)|$, $|G_{Y,ij}(z_k)| \leq \frac{1}{\text{Im}[z_k]} \leq N^{1+\sigma}$. After taking the trace, this shows that

$$|m_X(z_k) - m_Y(z_k)| \leq \|f'\|_\infty N^{4+2\sigma} 2^{-(\log N)^6}.$$

Using that $|m_X(z_k)|, |m_Y(z_k)| \leq N^{1+\sigma}$, we conclude that

$$\left| \prod_{k=1}^n \operatorname{Im} m_X(z_k) - \prod_{k=1}^n \operatorname{Im} m_Y(z_k) \right| \leq \|f'\|_\infty N^{(n-1)\sigma} N^{4+2\sigma} 2^{-(\log N)^6} \leq \frac{C_\sigma}{N}$$

since the logarithm comes with a power. \square

3.4 Proof of the main result assuming Theorems 3.1 and 3.2.

Proof of the local law, Theorem 2.6. This follows by combining Lemma 3.3 and Theorem 3.1. \square

For Theorem 2.10, we need another technical comparison result.

Lemma 3.4 (Correlation Function Comparison Lemma). *Let $k \geq 1$ and let $O : \mathbb{R}^k \rightarrow \mathbb{R}$ be smooth and compactly supported. Then, for $\varepsilon > 0$ and $\rho(E) \geq \varepsilon$, there exists $\kappa > 0$ such that*

$$(3.8) \quad \int_{\mathbb{R}^k} O(\alpha) \left(p_N^{(k)} \left(E + \frac{\alpha}{2N} \right) - \tilde{p}_N^{(k)} \left(E + \frac{\alpha}{2N} \right) \right) d^k \alpha = \mathcal{O}(N^{-\kappa}),$$

as $N \rightarrow \infty$.

Proof. The derivation of Lemma 3.4 from Lemma 3.3 and the local laws is standard for generalized Wigner matrices; see [29, Thm. 15.3] and [24, Thm. 6.4]. In essence, the proof idea is to mollify the observable O in each argument through approximate δ -functions of the form $\operatorname{Im} \frac{1}{x-E-i\eta}$ with $\eta \sim N^{-1-\sigma}$, which is a good approximation because these small scales resolve individual eigenvalues. This allows to express the tested correlation function using polynomials in $\operatorname{Im} m(z_1), \dots, \operatorname{Im} m(z_k)$ to which one applies Lemma 3.3. An inclusion-exclusion argument takes care of the possibility of eigenvalues matching. The local laws for $m_Y(z)$ and $m_X(z)$ play the role of a priori bound on the eigenvalue density in the form of [29, eq. (15.7)].

Our contribution here is solely to observe that the arguments carry over verbatim to the ensemble defined by H_Y . For the details, we refer the interested reader to [29, Thm. 15.3] and [24, Thm. 6.4]. \square

Proof of the main result, Theorem 2.10. This follows by combining Lemma 3.4 and Theorem 3.2. \square

It remains to prove Theorems 3.1 and 3.2. The main work is to establish Theorem 3.1, the rest is handled by the well-developed Dyson Brownian Motion machinery. From now on, we only consider the resampled matrices Y and we denote $H \equiv H_Y$, $m \equiv m_Y$, and $G \equiv G_Y$.

4 Derivation of the scalar self-consistent equation

This section is devoted to studying the self-consistent equation satisfied by the Stieltjes transform

$$(4.1) \quad m(z) = \frac{1}{2N} \text{Tr}[(H - z)^{-1}].$$

We will see that it is essentially given by (2.14) up to error terms that are subleading in N .

The standard approach for matrices with independent entries is to remove a single row and column via the Schur complement formula. With dependencies, this is no longer sufficient. The solution found in [16] is to expand the rows and columns even further until sufficient decoupling is achieved. In this section, we adapt this approach to the dynamically defined ensemble.

4.1 Preliminaries. We introduce some notation. Following [16], we define the $2N^2 \times 2N^2$ matrix Σ , called the **full correlation matrix of H** ,

$$(4.2) \quad \Sigma_{(i,j),(k,l)} = \xi_{ijkl} = N \text{Cov}[H_{ij}, \overline{H_{kl}}], \quad 1 \leq i, j, k, l \leq 2N.$$

One can immediately see that the matrix Σ is not strictly positive semidefinite in the sense of [16, Def. 2.2] since $\Sigma_{(i,j)(i,j)} = 0$ for $1 \leq i, j \leq N$.

To quantify the error terms in the self-consistent equation, we need some notation.

Definition 4.1 (Notation for index reduction). Let A be a matrix indexed by a subset of the integers.

- (i) For T and S subsets of \mathbb{Z} , we write $A|_{T,S}$ for the submatrix of A whose rows are given by T and whose columns are given by S .
- (ii) For U a subset of \mathbb{Z} , we denote by $A^{(U)}$ the matrix

$$A_{ij}^{(U)} = (A_{ij} \mathbb{1}[i \in U] \mathbb{1}[j \in U]).$$

This means that we set each entry of A whose row or column index belongs to U to 0 and do not alter the matrix A otherwise.

- (iii) For $U \subset \{1, \dots, 2N\}$, we also set $G^{(U)}(z) = (Y^{(U)} - z)^{-1}$. When U is a single element $\{k\}$, then we write G^k instead of $G^{(\{k\})}$.

Definition 4.2 (Stochastic control parameters). In order to establish various error bounds, we would require the following two stochastic control parameters.

- (i) We define

$$\Gamma := 1 \vee \sup_{i,j \in \{1, \dots, 2N\}} |G_{ij}|.$$

(ii) Given $i \in \{1, \dots, 2N\}$, let I and J be disjoint subsets of

$$[i - 2(\log N)^6, i + 2(\log N)^6] \cap \mathbb{Z}.$$

Let $\|\cdot\|$ denote the operator norm on $\mathbb{C}^{|I|}$ with respect to the Euclidean norm. We define

$$\gamma := 1 \vee \sup_{1 \leq i \leq 2N} \sup_{I,J} \|(G_{I,I}^{(J)})^{-1}\|.$$

4.2 The self-consistent equation. We first state the general form of the self-consistent equation satisfied by the Green's function.

Lemma 4.3 (Self-consistent equation for the Green's function). *Let $z = E + i\eta$ and $\sigma > 0$. Then*

$$(4.3) \quad -\frac{1}{N} \sum_{1 \leq k,l,m \leq 2N} G_{ik} \xi_{kljm} G_{lm} - z G_{ij} = \delta_{ij} + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right).$$

This lemma is proved in the next section. The proof is similar to that of [16, Lemma 3.10] due to the fact that the relevant concentration estimates for quadratic forms can be extended to a logarithmic range of dependence.

We are able to leverage the special structural properties of the dynamically defined matrix model to reduce to a scalar self-consistent equation for the Stieltjes transform $m(z)$.

The relevant finite- N analog of Φ from (2.9) is the banded $N \times N$ Toeplitz matrix Φ^N defined as follows, for any $1 \leq i,j \leq N$,

$$(4.4) \quad \Phi_{i,j}^N = \begin{cases} \phi(i-j), & \text{if } |i-j| \leq (\log N)^6, \\ 0, & \text{if } |i-j| \geq (\log N)^6 + 1. \end{cases}$$

We note that Φ^N is Hermitian because the function ϕ satisfies $\phi(j) = \overline{\phi(-j)}$.

With the identification of the most-important parts of the covariance, we can simplify the self-consistent equation from 4.3 to have a nice block structure.

Corollary 4.4. *Let*

$$G^1 = G|_{[1,N] \times [1,N]}, \quad G^2 = G|_{[N+1,2N] \times [N+1,2N]}, \quad G^3 = G|_{[1,N] \times [N+1,2N]}, \quad \text{and} \\ G^4 = G|_{[N+1,2N] \times [1,N]}.$$

We have the following equations on the blocks of G^i :

$$\begin{aligned}
 (4.5) \quad G^1 \left(-\frac{1}{N} \text{Tr}[G^2 \Phi^N] - z \right) &= \mathbb{1} + E_1 = \mathbb{1} + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right), \\
 G^2 \left(\frac{1}{N} \text{Tr}[G^1] \Phi^N - z \right) &= \mathbb{1} + E_2 = \mathbb{1} + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right), \\
 G^3 \left(-\Phi^N \frac{1}{N} \text{Tr}[G^1] - z \right) &= E_3 = \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right), \\
 G^4 \left(-\frac{1}{N} \text{Tr}[G^2 \Phi^N] - z \right) &= E_4 = \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right),
 \end{aligned}$$

where the \mathcal{O}_σ errors are elementwise.

By setting the errors to zero, we can derive the limiting matrix self-consistent equation. We notice vast simplifications from the equations in (4.5) compared to studying the full matrix-valued self-consistent equation

$$(4.6) \quad \mathcal{M}^{-1} = z + \Xi(\mathcal{M})$$

for the operator $\Xi(\mathcal{M})$ on $2N$ by $2N$ matrices

$$(4.7) \quad \Xi(\mathcal{M}) = \frac{1}{N} \sum_{\substack{(k,l,j,m)=(a,b+N,c,d+N) \\ \text{or } (a+N,b,c+N,d) \\ 1 \leq a,b,c,d \leq N}} \mathcal{M}_{ik} \check{\zeta}_{kljm} \mathcal{M}_{lm},$$

where the block structure can be seen from

$$\Xi \left(\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \right) = N^{-1} \begin{pmatrix} \text{Tr}(M_{22} \Phi^N) & 0 \\ 0 & \Phi^N \text{Tr} M_{11} \end{pmatrix}.$$

Indeed, the equations for G^1 and G^2 are closed and this is all that is necessary to determine the behavior of the trace of the Green's function. Furthermore, we would also expect G^3 , G^4 and the off-diagonal entries of G^1 to concentrate around zero. With these simplifications, we can present a scalar self-consistent equation. This form of the self-consistent equation is the most important for deriving stability and error propagation bounds in Section 6. Indeed, one sees that $M_{11} = m\text{Id}$ where m solves the scalar self-consistent equation $-m = N^{-1} \text{Tr}(m\Phi^N + z)^{-1}$ which is essentially what we analyze in the following, cf. the next proposition.

We write m_{Φ^N} for the Stieltjes transform of the (no longer explicit) spectral measure of Φ^N . Recall the definition of $m(z)$ from equation (4.1).

Proposition 4.5 (Scalar self-consistent equation for $m(z)$). *We have*

$$(4.8) \quad -\frac{1}{m(z)} m_{\Phi^N} \left(-\frac{z}{m(z)} \right) = m(z) + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^6 \gamma^3}{\sqrt{N\eta}} \right).$$

Remark 4.6. Formally taking $N \rightarrow \infty$, we see that we can expect the above equation to approximate the limiting self-consistent equation (2.14).

4.3 Proof of Proposition 4.5. We recall the Ward identity (or rather Ward identities since they helpfully decompose into two relations for sample covariance matrices such as H). These give us control on the sizes of the off-diagonal Green's function elements in terms of its diagonal elements.

Lemma 4.7 (Ward Identities). *The Green's function satisfies*

$$(4.9) \quad \begin{aligned} \sum_{l=1}^N |G_{i,l}|^2 &= \frac{1}{2} \frac{\text{Im}[G_{ii}]}{\eta}, \\ \sum_{l=1}^N |G_{i,l+N}|^2 &= \frac{1}{2} \frac{\text{Im}[G_{ii}]}{\eta}. \end{aligned}$$

Here and in the following use both $A_{i,j}$ or A_{ij} for entries of a matrix A , depending on convenience.

Proof of Proposition 4.5.

Step 1. Correlation matrix. Let $1 \leq i, j, k, l \leq N$. Using (3.1), (2.8) and the mean-value theorem, we find for the correlation matrix

$$(4.10) \quad \begin{aligned} \Sigma_{(i,j+N),(k,l+N)} &= \mathbb{E}[f(y_{(2N-1)i+j})\overline{f(y_{(2N-1)k+l})}] = \delta_{i,k}\Phi_{i,j}^N + \mathcal{O}(2^{-(\log N)^6}), \\ \Sigma_{(i+N,j),(k+N,l)} &= \mathbb{E}[f(y_{(2N-1)l+k})\overline{f(y_{(2N-1)j+i})}] = \delta_{j,l}\Phi_{i,k}^N + \mathcal{O}(2^{-(\log N)^6}), \\ \Sigma_{(i,j+N),(k+N,l)} &= \mathbb{E}[f(y_{(2N-1)i+j})f(y_{(2N-1)l+k})] = \delta_{i,l}\Psi_{k,j}^N + \mathcal{O}(2^{-(\log N)^6}), \\ \Sigma_{(i+N,j),(k,l+N)} &= \mathbb{E}[\overline{f(y_{(2N-1)j+i})}f(y_{(2N-1)k+l})] = \delta_{j,k}\overline{\Psi_{i,l}^N} + \mathcal{O}(2^{-(\log N)^6}), \end{aligned}$$

where Φ^N is given by (4.4) and we introduced its sibling

$$(4.11) \quad \Psi_{i,j}^N = \begin{cases} \psi(i-j), & \text{if } |i-j| \leq (\log N)^6, \\ 0, & \text{if } |i-j| \geq (\log N)^6 + 1. \end{cases}$$

with

$$\psi(j) = \mathbb{E}[f(x)f(T^j x)] = \sum_{k=1}^{\infty} c_k c_{k2^j}, \quad j \in \mathbb{Z}.$$

We note that (4.10) contains all the non-trivial entries of Σ , i.e., all others are zero. The error terms $\mathcal{O}(2^{-(\log N)^6})$ are super-polynomially decaying and can thus be ignored in the following.

Step 2. Reduction to blocks. We introduce the block Green's functions

$$G^1 = G|_{\{1, \dots, N\} \times \{1, \dots, N\}}, \quad G^2 = G|_{\{N+1, \dots, 2N\} \times \{N+1, \dots, 2N\}}$$

and their normalized traces

$$T_\alpha(z) = \frac{1}{N} \text{Tr}[G^\alpha(z)], \quad \alpha = 1, 2.$$

The Schur complement formula and cyclicity readily imply that $T_1(z) = T_2(z)$ and consequently

$$(4.12) \quad m(z) = \frac{1}{2N} \text{Tr}[G(z)] = \frac{T_1(z) + T_2(z)}{2} = T_1(z).$$

We see that it suffices to study the reduced Stieltjes transform $T_1(z)$.

We decompose the self-consistent equation (4.3) into blocks accordingly, taking into account when the correlation matrix vanishes. We first let $i, j \leq N$. Rearranging terms, we obtain

$$(4.13) \quad \begin{aligned} & -\frac{1}{N} \sum_{k,l,m=1}^N G_{ik}^1 \zeta_{k,l+N,j,m+N} G_{lm}^2 - z G_{ij}^1 \\ & = \delta_{ij} + \frac{1}{N} \sum_{k,l,m=1}^N G_{i,k+N} \zeta_{k+N,l,j,m+N} G_{l,m+N} + \mathcal{O}_\sigma\left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}}\right) \\ & = \delta_{ij} + \mathcal{O}_\sigma\left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}}\right). \end{aligned}$$

For the second step, we used the formula from Step 1, Cauchy–Schwarz, the Ward identity, and $\|\psi\|_{\ell^2} < \infty$ (proved below) to estimate

$$\begin{aligned} \frac{1}{N} \left| \sum_{k,l,m=1}^N G_{i,k+N} \zeta_{k+N,l,j,m+N} G_{l,m+N} \right| & \leq \frac{1}{N} \sum_{k,m=1}^N |\overline{\psi(k-m)} G_{i,k+N} G_{j,m+N}| \\ & \leq \frac{\Gamma}{2N\eta} \|\psi\|_{\ell^2} = \mathcal{O}_\sigma\left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}}\right). \end{aligned}$$

To see that $\|\psi\|_{\ell^2} < \infty$, note that ψ is the Fourier transform of

$$\tilde{g}_f(x) = \sum_{\substack{n \geq 0: \\ 2 \nmid n}} \left(\sum_{k=0}^{\infty} c_{n2^k} \exp[2\pi i kx] \right)^2,$$

which is dominated by g_f . Hence, we have

$$\|\psi\|_{\ell^2} \leq \|\tilde{g}_f\|_{L^2([0,1])} \leq \|g_f\|_{L^\infty([0,1])} \leq C\|f\|_{H^1(\mathbb{R})}.$$

Using the formula from Step 1 to rewrite the left-hand side of (4.13), we conclude that

$$(4.14) \quad -\frac{1}{N} G_{ij}^1 \sum_{l,m=1}^N \Phi_{m,l}^N G_{lm}^2 - z G_{ij}^1 = \delta_{ij} + \mathcal{O}_\sigma\left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}}\right).$$

For $i, j > N$ we have to consider the expression

$$-\frac{1}{N} \sum_{k,l,m=1}^N G_{ik}^2 \xi_{k+N,l,j+N,m} G_{lm}^1 - z G_{ij}^2 = \delta_{ij} + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right).$$

Completely analogous reasoning yields the remarkably different formula

$$(4.15) \quad -T_1(z) \sum_{k=1}^N G_{ik}^2 \Phi_{k,j}^N - z G_{ij}^2 = \delta_{ij} + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right).$$

where we used that $\frac{1}{N} \sum_l G_{il}^2 = T_2(z) = T_1(z)$.

Inspecting the system of equations (4.14) and (4.15), we notice that we can reduce G^1 to $T_1(z)$ also from (4.14) by averaging over $i = j$. By contrast, the elements of G^2 are multiplied by the Hermitian Toeplitz matrix Φ^N . We obtain the system

$$(4.16) \quad \begin{aligned} T_1(z) \left(-\frac{1}{N} \text{Tr}[G^2 \Phi^N] - z \right) &= 1 + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right), \\ -T_1(z) G^2 \Phi^N - z G^2 &= \mathbb{1} + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right), \end{aligned}$$

where in the second line the error estimate holds componentwise. We can use the second equation to solve for G^2 as

$$G^2 = \frac{1}{-T_1 \Phi^N - z} \left(1 + \mathcal{O}_\sigma \left(\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}} \right) \right).$$

It remains to understand how the matrix $\frac{1}{-T_1 \Phi^N - z}$ affects the error estimate. This is relegated to the next subsection; see Proposition 4.8. Together with (4.12), this proves Proposition 4.5. \square

4.4 Propagation of error bounds.

We set

$$(4.17) \quad \|M\|_\infty := \sup_{i,j} |M_{i,j}|.$$

Proposition 4.8. *Consider a solution to the system of equations*

$$(4.18) \quad \begin{aligned} T_1(z) \left(-\frac{1}{N} \text{Tr}(G^2(z) \Phi^N) - z \right) &= 1 + e_1 \\ -G^2(z) \Phi T_1(z) - z G^2(z) &= I + E_2, \end{aligned}$$

where e_1 is a number and E_2 is a matrix. Then

$$(4.19) \quad T_1(z) = -\frac{1}{T_1(z)} m_\Phi \left(-\frac{z}{T_1(z)} \right) + \tilde{e}_1,$$

where \tilde{e}_1 satisfies,

$$(4.20) \quad |\tilde{e}_1| \leq \frac{1}{|z|}(|e_1| + (\log N)^6 \|E_2\|_\infty).$$

The proof uses spectral bounds on the finite Toeplitz matrix Φ^N . These are inherited from the infinite Toeplitz matrix Φ as summarized in the following lemma.

Lemma 4.9 (Spectral estimates for finite banded Toeplitz matrix). *There exist N -independent constants $c, C > 0$ such that*

$$(4.21) \quad c \leq \Phi^N \leq C.$$

Proof of Lemma 4.9. Recall that g_f is the Fourier symbol of the infinite Toeplitz matrix Φ and so

$$\inf \text{spec } \Phi = \inf_{x \in [0, 1]} g_f(x) = g_{\min} > 0, \quad \sup \text{spec } \Phi = \sup_{x \in [0, 1]} g_f(x) \leq C \|f\|_{H^1}^2.$$

It is a well-known fact that the finite restriction of an infinite Toeplitz matrix has spectrum lying inside the spectrum determined by the Fourier symbol of the infinite matrix [31, Lemma 6]. In the present case, this implies that if we define the auxiliary Toeplitz matrix

$$(4.22) \quad A_{i,j}^N = \phi(i - j), \quad 1 \leq i, j \leq N$$

then

$$\inf \text{spec } A^N \geq \inf_{x \in [0, 1]} g_f(x) = g_{\min} > 0, \quad \sup \text{spec } A^N \leq \sup_{x \in [0, 1]} g_f(x) \leq C \|f\|_{H^1}^2.$$

Recall that the matrix Φ^N defined by (4.4) is cut off at the band edges $|i - j| = (\log N)^6$. Using Definition (2.8) of ϕ , the fact that $f \in C^2(\mathbb{T})$ and the mean-value theorem, we see that

$$\|A^N - \Phi^N\| = o(1), \quad \text{as } N \rightarrow \infty$$

and so Lemma 4.9 follows from Weyl's norm perturbation theorem. \square

Proof of Proposition 4.8. We abbreviate $T_1 \equiv T_1(z)$. Solving the second identity for G^2 yields $G^2 = -(I + E_2)(\Phi T_1 + z)^{-1}$ and then the first identity gives

$$(4.23) \quad T_1 \left(\frac{1}{N} \text{Tr}(\Phi^N (\Phi^N T_1 + z)^{-1}) - z \right) = 1 + e_1 - \frac{1}{N} \text{Tr}(E_2 (\Phi^N T_1 + z)^{-1} \Phi^N).$$

After some algebraic manipulation, we derive

$$(4.24) \quad T_1 = -\frac{1}{T_1} m_{\Phi^N} \left(-\frac{z}{T_1} \right) + \frac{1}{z} \left(-e_1 + \frac{1}{N} \text{Tr}(E_2 (\Phi^N T_1 + z)^{-1} \Phi^N) \right).$$

Our goal is now to control the second error term $\frac{1}{N}\text{Tr}(E_2(\Phi^N T_1 + z)^{-1} \Phi^N)$. By the cyclic property of the trace, we know that

$$\text{Tr}(E_2(\Phi^N T_1 + z)^{-1} \Phi^N) = \text{Tr}(\Phi^N E_2(\Phi^N T_1 + z)^{-1}).$$

Now, since Φ^N is a band matrix of band size $2(\log N)^6$ and only bounded entries, we see that $\|\Phi^N E_2\|_\infty \leq C(\log N)^6 \|E_2\|_\infty$. To estimate the entries of $(\Phi^N T_1 + z)^{-1}$, we need to use the fact that we are considering the inverse of a banded matrix. We apply the following result on the decay of entries of the inverse of banded matrices from [16].

Lemma 4.10 (Inversion of band matrices [16]). *Let A be an invertible infinite or finite matrix, which is W -banded in the sense that $A(i, j) = 0$ if $|i - j| \geq W$. We have*

$$|(A^{-1})_{i,j}| \leq 2(2W + 1)\kappa(A)\alpha^{(|i-j|-W)_+},$$

where

$$\kappa(A) = \|A\| \cdot \|A^{-1}\|, \quad \alpha = \left(\frac{\kappa(A) - 1}{\kappa(A) + 1}\right)^{2/(2W+1)}.$$

We observe that $\kappa(T_1 \Phi^N + z)$ as defined in Lemma 4.10 is uniformly bounded independently of N . Indeed, by Lemma 4.9, we have

$$\|T_1 \Phi^N + z\| \geq c\omega$$

because

$$\text{Im}\langle v, (T_1 \Phi^N + z)v \rangle = \text{Im}[T_1]\langle v, \Phi^N v \rangle + \text{Im}[z]\langle v, v \rangle \geq (\text{Im}[T_1]c + \text{Im}[z])\langle v, v \rangle.$$

Similar arguments show that $\|T_1 \Phi^N + z\|$ is bounded from above.

The band width that we have to consider is $W = (\log N)^6$. Thus, we see that the constant α is given by

$$(4.25) \quad \alpha = \left(\frac{\kappa(T_1 \Phi^N + z) - 1}{\kappa(T_1 \Phi^N + z) + 1}\right)^{2/(2(\log N)^{6+1})} = 1 + \mathcal{O}((\log N)^{-6}).$$

We can then control individual entries of the matrix $\Phi^N E_2(\Phi^N T_1 + z)^{-1}$, namely

$$\begin{aligned} & \|\Phi^N E_2(\Phi^N T_1 + z)^{-1}\|_\infty \\ (4.26) \quad & \leq 2(2(\log N)^6 + 1)\kappa(T_1 \Phi^N + z) \sum_{k=-N}^N \|\Phi^N E_2\|_\infty \alpha^{(|k|-(\log N)^6)_+} \\ & \leq C(\log N)^6 \|\Phi^N E_2\|_\infty \left((\log N)^6 + \frac{1}{1-\alpha}\right) \\ & \leq C(\log N)^{18} \|E_2\|_\infty. \end{aligned}$$

The bound on individual entries implies a bound on the normalized trace. This proves Proposition 4.8. \square

5 Derivation of the matrix self-consistent equation

In this section, we prove Lemma 4.3. The following derivation is modeled after [16]. While we simplify some aspects, it is rather technical. Hence, we invite the reader to instead consider Appendix B for a heuristic derivation of (4.3) based on Gaussian integration by parts and the loop equation.

5.1 Basic removal operations and bounds. The following lemma can be seen as a generalization of the Schur complement formula. The proof is via the resolvent equation and can be found in [16, Lemma 3.3].

Lemma 5.1. *Let T, I, J be three subsets of $\{1, \dots, 2N\}$ satisfying*

$$T \cap I, T \cap J = \emptyset.$$

Then, we have the following resolvent identities:

$$(5.1) \quad G_{I,J} = G_{I,J}^{(T)} + G_{I,T}(G_{T,T})^{-1}G_{T,J},$$

$$(5.2) \quad G_{T,J} = -G_{T,T}H_{T,T^c}G_{T^c,J}^{(T)}.$$

As an immediate corollary, we can estimate the sizes of the Green's function after row removal in terms of the original Green's function (more precisely, in terms of the stochastic control parameters from Definition 4.2). This is used as a deterministic a priori estimate later.

Corollary 5.2. *Let i be an integer in $\{1, \dots, 2N\}$ and let*

$$I \subset [i - 2(\log N)^6, i + 2(\log N)^6] \cap \mathbb{Z}.$$

Then

$$(5.3) \quad |G_{kl}^{(I)}| \leq 8(\log N)^6\Gamma^2\gamma, \quad k, l \in \{1, \dots, 2N\} \setminus I.$$

Proof. Consider the identity of (5.1) with $I = \{k\}$, $J = \{l\}$ and $T = I$. We obtain

$$(5.4) \quad G_{kl}^{(I)} = G_{kl} - G_{k,I}(G_{I,I})^{-1}G_{I,j}.$$

We can treat the second term as an inner product and bound it by treating $G_{k,I}$ and $G_{I,j}$ as $4(\log N)^6$ dimensional vectors. That is

$$|G_{k,I}(G_{I,I})^{-1}G_{I,j}| \leq \|G_{k,I}\|_2\|(G_{I,I})^{-1}\|_{2 \rightarrow 2}\|G_{I,j}\|_2 \leq 4(\log N)^6\Gamma^2\gamma.$$

The first term $|G_{kl}|$ can be bounded by Γ , which is less than $4(\log N)^6\Gamma^2\gamma$ since $\Gamma, \gamma \geq 1$ by definition. \square

5.2 Concentration Identities. A key ingredient in the proof of [16, Lemma 3.10] are concentration bounds for variables with finite range of dependence. Here we extend these estimates to the case of logarithmic range of dependence by grouping arguments similar to those in [16].

Lemma 5.3. *Let a_1, \dots, a_N be a family of dependent random variables that satisfy the property that a_i and a_j are independent whenever $|i - j| \geq \lceil (\log N)^6 \rceil$.*

We fix a parameter $\sigma > 0$, a power $p \geq p_0(\sigma)$ and assume the following moment estimate.

$$(5.5) \quad \mathbb{E}[|a_i|^p] \leq \frac{C^p}{\sqrt{N^p}}.$$

Then, there exists a universal constant $C > 0$ such that for any collection of deterministic A_i and B_{ij} with $1 \leq i, j \leq N$, with probability $1 - N^{-\sigma' p}$ and some $\sigma > \sigma'$, we have that

$$(5.6) \quad \left| \sum_{i=1}^N A_i a_i - \mathbb{E} \left[\sum_{i=1}^N A_i a_i \right] \right| \leq CN^\sigma \left(\frac{\sup_i |A_i|}{\sqrt{N}} + \sqrt{\frac{1}{N} \sum_{i=1}^N |A_i|^2} \right),$$

$$(5.7) \quad \left| \sum_{i,j=1}^N B_{ij} a_i a_j - \mathbb{E} \left[\sum_{i,j=1}^N B_{ij} a_i a_j \right] \right| \leq CN^\sigma \left(\frac{\sup_{i,j} |B_{ij}|}{\sqrt{N}} + \sqrt{\frac{1}{N^2} \sum_{i,j=1}^N |B_{ij}|^2} \right).$$

Proof. All sums are implicitly over $\{1, \dots, N\}$. By dividing the integers modulo $\lceil (\log N)^6 \rceil$, we see that each sum

$$(5.8) \quad \mathcal{A}_k = \sum_{i=k \bmod \lceil (\log N)^6 \rceil} A_i a_i - \mathbb{E} \left[\sum_{i=k \bmod \lceil (\log N)^6 \rceil} A_i a_i \right],$$

for $0 \leq k \leq \lceil (\log N)^6 \rceil - 1$, is a sum of terms involving only independent a_i . We can then appeal to standard concentration estimates, see [29, Theorem 7.7], to conclude that with probability $1 - c_p N^{\sigma' p}$, we have

$$\mathcal{A}_k \leq N^{\sigma'} \left(\frac{\sup_i |A_i|}{\sqrt{N}} + \sqrt{\frac{1}{N} \sum_{i=1}^N |A_i|^2} \right),$$

where $\sigma' < \sigma$. A union bound for all the sums \mathcal{A}_k gives the desired estimate. We can absorb $(\log N)^6$ factors into N^σ both in the bound (5.6).

We come to the quadratic summation, where we have to do some more manipulation. First, we split

$$(5.9) \quad \begin{aligned} & \sum_{i,j} B_{i,j} a_i a_j - \mathbb{E} \left[\sum_{i,j} B_{i,j} a_i a_j \right] \\ &= \sum_{|i-j| > (\log N)^6} B_{i,j} a_i a_j + \sum_{|i-j| \leq (\log N)^6} B_{i,j} a_i a_j - \mathbb{E} \left[\sum_{|i-j| \leq (\log N)^6} B_{i,j} a_i a_j \right]. \end{aligned}$$

Consider the first term on the right-hand side. It can be split further into a total of $(\log N)^{12}$ parts based on the moduli of i and j with $(\log N)^6$; call them

$$(5.10) \quad \mathcal{B}_{\alpha,\beta} = \sum_{\substack{|i-j| > (\log N)^6 \\ i \equiv \alpha \pmod{(\log N)^6} \\ j \equiv \beta \pmod{(\log N)^6}}}^N B_{i,j} a_i a_j.$$

Each $\mathcal{B}_{\alpha,\beta}$ is a centered sum of independent random variables and so we can appeal to standard concentration estimates and a union bound with the logarithmic factor $(\log N)^{12}$ absorbed by the power function.

It remains to estimate the second and third terms on the right-hand side of (5.9). To this end, we split the sum into $4(\log N)^6$ parts as follows. Let α, β be integers with

$$1 \leq \alpha \leq \lceil 2(\log N)^6 \rceil \quad \text{and} \quad -(\log N)^6 \leq \beta \leq (\log N)^6$$

and set

$$(5.11) \quad \mathcal{C}_{\alpha,\beta} = \sum_{\substack{i-j=\beta \\ i \equiv \alpha \pmod{\lceil 2(\log N)^6 \rceil}}} B_{i,j} a_i a_j - \mathbb{E} \left[\sum_{\substack{i-j=\beta \\ i \equiv \alpha \pmod{\lceil 2(\log N)^6 \rceil}}} B_{i,j} a_i a_j \right].$$

Each $\mathcal{C}_{\alpha,\beta}$ is a centered sum of independent random variables and there are $4(\log N)^{12}$ choices of α and β , so the result follows from standard concentration estimates and a union bound. \square

We can use the previous lemma, the Ward Identity, and our stochastic control parameters to write out an explicit bound for the sums we consider upon applying resolvent identities.

Corollary 5.4. *Let a_i be random variables satisfying the same conditions as in Lemma 5.3. In addition, assume that all A_i, B_{ij} belong to the set $\{G_{ij}^{(T)}\}_{i,j}$ for some subset T of*

$$[k - 2(\log N)^6, k + 2(\log N)^6] \cap \mathbb{Z}, \quad \text{with } k \in \{1, \dots, 2N\}.$$

Then, for any $\sigma > 0$ and $p \geq p_0(\sigma) \geq 2$, we have with probability at least $1 - N^{-\sigma'p}$, and some $\sigma > \sigma'$, that

$$(5.12) \quad \begin{aligned} \left| \sum_i A_i a_i - \mathbb{E} \left[\sum_i A_i a_i \right] \right| &\leq C \frac{N^\sigma \Gamma^2 \gamma}{\sqrt{N\eta}}, \\ \left| \sum_i B_{ij} a_i a_j - \mathbb{E} \left[\sum_{i,j} B_{i,j} a_i a_j \right] \right| &\leq C \frac{N^\sigma \Gamma^2 \gamma}{\sqrt{N\eta}}. \end{aligned}$$

Proof. We focus on the first estimate in (5.12). By Lemma 5.3, we have with probability at least $1 - N^{-\sigma'p}$ that

$$(5.13) \quad \left| \sum_i A_i a_i - \mathbb{E} \left[\sum_i A_i a_i \right] \right| \leq N^{\sigma'} \left(\frac{\sup_i |G_{ki}^{(T)}|}{\sqrt{N}} + \sqrt{\frac{1}{N} \sum_i |G_{ki}^{(T)}|^2} \right).$$

By applying the Ward Identity, Lemma 4.7, we can bound the sum $\frac{1}{N} \sum_i |G_{ki}^{(T)}|^2$ by $\frac{\text{Im}[G_{kk}^{(T)}]}{N\eta}$. Now we apply Corollary 5.2 to bound the Green function terms by $8(\log N)^6 \Gamma^2 \gamma$. The $(\log N)$ factors can be absorbed into N^σ . This gives us the desired first inequality of (5.12).

The second inequality in (5.12) follows from similar arguments; we omit the details. \square

5.3 Preliminary self-consistent equation. As in the case of independent variables, the derivation of the self-consistent equation with dependence relies on extracting appropriate matrix components to which the concentration estimates can then be applied in a second step.

By definition, the Green's function satisfies

$$(5.14) \quad \sum_k G_{ik} H_{kj} - z G_{ij} = \delta_{ij}.$$

We let T be the set of entries that are correlated with j , i.e.,

$$T = [j - (\log N)^6, j + (\log N)^6] \cap \mathbb{Z}.$$

Definition 5.5. Fix a parameter σ and an integer p such that $p \geq \frac{100}{\sigma}$. For two sequences of random variables $a = a^{(N)}$ and $b = b^{(N)}$, we say that $a = \mathcal{O}_{\sigma,p}(b)$ if there exists universal constants c and N_0 such that for $N > N_0$, we know that $a^{(N)} \leq b^{(N)}$ with probability at least $1 - (\log N)^c N^{-\sigma p}$. We will drop the superscript (N) when the context is clear.

The following lemma specifies a preliminary self-consistent equation.

Lemma 5.6. *Let σ and $p \geq p_0(\sigma) \geq 2$. Then*

$$(5.15) \quad -G_{iT}H_{TT^c}G_{T^cT^c}^{(T)}H_{T^c j} - zG_{ij} = \delta_{ij} + \mathcal{O}_{\sigma,p}\left(\frac{N^\sigma \Gamma^2 \gamma}{\sqrt{N\eta}}\right).$$

Proof. We recall the simple bounds $|H_{kj}| \leq \frac{C}{\sqrt{N}}$ and $|G_{ik}| \leq \Gamma$.

Consider (5.14). In the sum $\sum_k G_{ik}H_{kj}$, we bound the terms with $k \in T$ by $\frac{\Gamma}{\sqrt{N}}$ and obtain

$$G_{iT^c}H_{T^c j} - zG_{ij} = \delta_{ij} + \mathcal{O}\left(\frac{(\log N)^6 \Gamma}{\sqrt{N}}\right).$$

We distinguish cases based on whether $i \in T$ or $i \notin T$.

If $i \in T$, we apply equation (5.2) and find

$$-G_{iT}H_{TT^c}G_{T^cT^c}^{(T)}H_{T^c j} - zG_{ij} = \delta_{ij} + \mathcal{O}_{\sigma,p}\left(\frac{(\log N)^6 \Gamma}{\sqrt{N}}\right).$$

Otherwise, if $i \notin T$, we first apply equation (5.1) to obtain

$$\sum_{k \in T^c} G_{ik}^{(T)}H_{kj} + \sum_{k \in T^c} G_{iT}(G_{T,T})^{-1}G_{T,k}H_{kj} - zG_{ij} = \delta_{ij} + \mathcal{O}\left(\frac{(\log N)^6 \Gamma}{\sqrt{N}}\right).$$

We can apply Corollary 5.4 to the first term and (5.2) to the second term to obtain

$$G_{iT}H_{TT^c}G_{T^cT^c}^{(T)}H_{T^c j} - zG_{ij} = \delta_{ij} + \mathcal{O}_{\sigma,p}\left(\frac{N^\sigma \Gamma^2 \gamma}{\sqrt{N\eta}}\right)$$

with probability at least $1 - N^{-\sigma p}$ as desired. \square

5.4 Conclusion. Our goal at this point is to use our concentration estimates to replace $H_{TT^c}G_{T^cT^c}^{(T)}H_{T^c j}$ with its expectation. At this point, we have ensured that the term $H_{T^c j}$ is completely independent of the term $G_{T^cT^c}^{(T)}$. However, we now need to guarantee that the term coming from H_{TT^c} is independent of $G_{T^cT^c}^{(T)}$. Based on the first index $k \in T$ of H_{TT^c} , we need to apply a further resolvent expansion to remove terms in $G_{T^cT^c}^{(T)}$ that are correlated with H_{TT^c} .

Proof of Lemma 4.3. Our starting point is Lemma 5.6, i.e.,

$$-G_{iT}H_{TT^c}G_{T^cT^c}^{(T)}H_{T^c j} - zG_{ij} = \delta_{ij} + \mathcal{O}_{\sigma,p}\left(\frac{N^\sigma \Gamma^2 \gamma}{\sqrt{N\eta}}\right).$$

We let $S \subset T^c$ be the set of entries correlated with T and let $U = T \cup S$. We first split the summation over $T^c = U^c \cup S$ as

$$H_{TT^c}G_{T^cT^c}^{(T)}H_{T^c j} = H_{TU^c}G_{U^cU^c}^{(T)}H_{U^c j} + H_{TU^c}G_{U^cS}^{(T)}H_{Sj} + H_{TS}G_{ST^c}^{(T)}H_{T^c j}.$$

Step 1. We identify the leading term by applying the resolvent identity to replace the first $G_{U^c U^c}^{(T)}$ with $G_{U^c U^c}^{(U)}$ plus what we will show are error terms. Observe from (5.1) that

$$(5.16) \quad G_{U^c U^c}^{(T)} = G_{U^c U^c}^{(U)} + G_{U^c S}^{(T)} (G_{SS}^{(T)})^{-1} G_{S U^c}^{(T)}.$$

In particular, we have

$$(5.17) \quad H_{TT^c} G_{T^c T^c}^{(T)} H_{T^c j} = H_{TU^c} G_{U^c U^c}^{(U)} H_{U^c j} + e_{Tj},$$

where an individual entry e_{kj} of the error vector e_{Tj} can be written as

$$(5.18) \quad e_{kj} = H_{kU^c} G_{U^c S}^{(U)} (G_{SS}^{(T)})^{-1} G_{S U^c}^{(T)} H_{U^c j} + H_{kU^c} G_{U^c S}^{(T)} H_{Sj} + H_{kS} G_{S T^c}^{(T)} H_{T^c j}.$$

Since, by construction, the term $G_{U^c U^c}^{(U)}$ is independent from the terms appearing in H_{TU^c} and $H_{U^c j}$, we are able to apply our concentration estimates as desired.

Indeed, fix an entry $k \in T$. We apply the second concentration estimate in (5.12) and (4.2) to obtain

$$(5.19) \quad H_{kU^c} G_{U^c U^c}^{(U)} H_{U^c j} = \frac{1}{N} \sum_{l, m \in U^c} G_{lm}^{(U)} \xi_{kljm} + \mathcal{O}_{\sigma, p} \left(\frac{N^\sigma \Gamma^2 \gamma}{\sqrt{N\eta}} \right).$$

We now have to replace $G_{lm}^{(U)}$ with G_{lm} by reversing the resolvent expansion, i.e.,

$$(5.20) \quad G_{lm}^{(U)} = G_{lm} - G_{lU} (G_{U,U})^{-1} G_{Um}.$$

We can apply the Cauchy-Schwarz inequality plus the Ward Identity to bound

$$(5.21) \quad \left| \sum_{l, m \in U^c} G_{lU} (G_{U,U})^{-1} G_{Um} \right| \leq \|G_{U,U}\|_2 \| (G_{U,U})^{-1} \|_{2 \rightarrow 2} \| \|G_{U,U}\|_2 \leq \mathcal{O} \left(\frac{\Gamma \gamma}{N\eta} \right).$$

Combining these estimates, we have shown that

$$(5.22) \quad - \sum_{k, l, m} G_{ik} \xi_{kljm} G_{lm} - z G_{ij} = \delta_{ij} + \sum_{k \in T} G_{ik} e_{kj} + \mathcal{O}_{\sigma, p} \left(\frac{N^\sigma \Gamma^3 \gamma}{\sqrt{N\eta}} \right).$$

Step 2. It remains to estimate the error terms e_{kj} from (5.18), or more precisely $\sum_{k \in T} G_{ik} e_{kj}$.

For any fixed entry $s \in S$, we apply Corollary 5.4 to $G_{S U^c}^{(T)} H_{U^c j}$, $H_{kU^c} G_{U^c S}^{(T)}$, and $G_{S T^c}^{(T)} H_{T^c j}$ to obtain

$$(5.23) \quad G_{S U^c}^{(T)} H_{U^c j}, \quad H_{kU^c} G_{U^c S}^{(T)}, \quad G_{S T^c}^{(T)} H_{T^c j} \leq \mathcal{O}_{\sigma, p} \left(\frac{N^\sigma \Gamma^2 \gamma}{\sqrt{N\eta}} \right).$$

For the remaining terms, e.g., $H_{kU^c}G_{U^cS}^{(U)}$, we can no longer apply the concentration estimate. However, observe from the resolvent expansion (5.2), $T^c = U^c \cup S$ and equation (5.3) that

$$(5.24) \quad H_{TU^c}G_{U^cS}^{(T)} = -(G_{TT})^{-1}G_{TS} + \mathcal{O}\left((\log N)^{12} \frac{\Gamma^2 \gamma}{\sqrt{N}}\right) = \mathcal{O}((\log N)^{12} \Gamma^2 \gamma).$$

We conclude that

$$(5.25) \quad \sum_{k \in T} G_{ik} e_{kj} = \mathcal{O}_{\sigma, p}\left(\frac{N^\sigma \Gamma^3 \gamma^2}{\sqrt{N\eta}}\right),$$

as desired. This completes Step 2 and proves Lemma 4.3. \square

6 Analysis of the self-consistent equation

In this section, we study the scalar self-consistent equation that arises from Proposition 4.5.

6.1 Auxiliary self-consistent equation. It is convenient to focus on the following auxiliary scalar self-consistent equation that arises from Proposition 4.5 by dropping the error term, namely

$$(6.1) \quad -\frac{1}{M_N(z)} m_{\Phi^N}\left(-\frac{z}{M_N(z)}\right) = M_N(z).$$

Note that this differs slightly from the limiting self-consistent equation (2.14) satisfied by $m_\infty(z)$. Indeed, (6.1) still features the $N \times N$ Toeplitz matrix Φ^N while (2.14) contains the infinite Toeplitz matrix Φ instead.

First, we prove existence, uniqueness, and stability of solutions to (6.1). Second, in Lemma 6.5, we show that M_N is close to the solution of the limiting equation m_∞ for large N . Finally, we study in detail the propagation of error estimates on the Green's function into the self-consistent equation and show that we can use solutions of (6.1) to approximate the Green's function.

6.2 Existence and uniqueness of solutions. In this section, we prove existence and uniqueness of solutions to the scalar self-consistent equation (6.1). Note that the results in this section follow from more general discussions of matrix-valued self-consistent equations [33].

Lemma 6.1 (Existence). *There exists a solution $M_N(z) : \mathbb{C}^+ \rightarrow \mathbb{C}^+$ to the equation (6.1).*

We emphasize that $M_N(z)$ denotes a scalar quantity. The proof uses Brouwer's fixed point theorem.

Proof. Fix $z \in \mathbb{C}^+$. We define the function $F(w) = -\frac{1}{w}m_{\Phi^N}(-\frac{z}{w})$ that appears on the left-hand side of (6.1). We can rewrite $F(w)$ as

$$(6.2) \quad F(w) = - \int_{\mathbb{R}} \frac{1}{wx + z} d\rho_{\Phi^N}(x)$$

where $d\rho_{\Phi^N}$ is the empirical spectral distribution of Φ^N .

We consider the compact, convex domain

$$(6.3) \quad \mathcal{D}_z := \left\{ w \in \mathbb{C}^+ : |w| \leq \frac{1}{\text{Im}[z]}, \quad \text{Im}[w] \geq \frac{\text{Im}[z]}{\left(\frac{C}{\text{Im}[z]} + |z|\right)^2} \right\}$$

or an appropriate constant $C > 0$ to be determined. Below we show that $F : \mathcal{D}_z \rightarrow \mathcal{D}_z$. Then Brouwer's fixed point theorem implies that F has a fixed point. This fixed point is then the desired solution $M_N(z) \in \mathcal{D}_z$ which we note has positive imaginary part.

It remains to establish $F : \mathcal{D}_z \rightarrow \mathcal{D}_z$. Let $w \in \mathcal{D}_z$. Since z, w have positive imaginary part and x is non-negative, $|wx + z| \geq \text{Im}[z]$ and so

$$|F(w)| \leq \frac{1}{\text{Im}[z]} \rho_{\Phi^N}(\mathbb{R}) = \frac{1}{\text{Im}[z]}.$$

Next, we consider the behavior of the imaginary part,

$$(6.4) \quad \text{Im}[F(w)] = \int_{\mathbb{R}} \frac{\text{Im}[w]x + \text{Im}[z]}{|wx + z|^2} d\rho_{\Phi^N}(x).$$

From (4.21), we conclude that $d\rho_{\Phi^N}$ is supported on $[g_{\min}, C] \subset \mathbb{R}_+$. Using that $|w| \leq \frac{1}{\text{Im}[z]}$, we have on the support of $d\rho_{\Phi^N}$ that $\frac{1}{|wx + z|^2} \geq \frac{1}{\left(\frac{C}{\text{Im}[z]} + |z|\right)^2}$. Hence

$$(6.5) \quad \text{Im}[F(w)] \geq \frac{\text{Im}[z]}{\left(\frac{C}{\text{Im}[z]} + |z|\right)^2}$$

as desired. \square

Since Brouwer's fixed point theorem does not imply uniqueness, we have to prove it by hand using the structure of the self-consistent equation.

Lemma 6.2 (Uniqueness). *Let $z \in \mathbb{C}^+$. Then there is a unique solution $M_N(z)$ to (6.1) with positive imaginary part.*

Proof. Fix $z \in \mathbb{C}^+$. We first show that any solution $M_N(z)$ of (6.1) with positive imaginary part must satisfy

$$(6.6) \quad 1 - \int_{\mathbb{R}} \frac{x}{|M_N(z)x + z|^2} d\rho_{\Phi^N}(x) > 0.$$

Indeed, taking imaginary parts of both sides of (6.1) in the representation (6.4), we obtain

$$(6.7) \quad \begin{aligned} & \operatorname{Im}[M_N(z)] \\ &= \operatorname{Im}[M_N] \int_{\mathbb{R}} \frac{x}{|M_N(z)x + z|^2} d\rho_{\Phi^N}(x) + \operatorname{Im}[z] \int_{\mathbb{R}} \frac{1}{|M_N(z)x + z|^2} d\rho_{\Phi^N}(x) \end{aligned}$$

from which (6.6) follows by rearranging.

Now suppose that we have two solutions M and \tilde{M} to (6.1). Then we rewrite $M - F(M) - (\tilde{M} - \tilde{F}(M)) = 0$ using (6.4) as

$$(6.8) \quad \begin{aligned} & M - \tilde{M} - (M - \tilde{M}) \int_{\mathbb{R}} \frac{x}{(Mx + z)(\tilde{M}x + z)} d\rho_{\Phi^N}(x) = 0, \\ & (M - \tilde{M}) \left[1 - \int_{\mathbb{R}} \frac{x}{(Mx + z)(\tilde{M}x + z)} d\rho_{\Phi^N}(x) \right] = 0. \end{aligned}$$

Now, observe that by (6.6),

$$(6.9) \quad \begin{aligned} & \left| \int_{\mathbb{R}} \frac{x}{(Tx + z)(\tilde{T}x + z)} d\rho_{\Phi^N}(x) \right| \\ & \leq \left(\int_{\mathbb{R}} \frac{x}{|Tx + z|^2} d\rho_{\Phi^N}(x) \right)^{1/2} \left(\int_{\mathbb{R}} \frac{x}{|\tilde{T}x + z|^2} d\rho_{\Phi^N}(x) \right)^{1/2} < 1. \end{aligned}$$

This proves $M = \tilde{M}$ as desired. \square

6.3 Stability estimates. The estimates used in the proof of the last lemma can be refined to show stability of solutions to (6.1). Recall that $M_N(z)$ is the unique solution with positive imaginary part to

$$(6.10) \quad M_N(z) = -\frac{1}{M_N(z)} m_{\Phi^N} \left(-\frac{z}{M_N(z)} \right).$$

Lemma 6.3 (Stability). *Let $\Phi \in \{\Phi^N, \Phi\}$ and let T_0 solve*

$$(6.11) \quad T_0 = -\frac{1}{T_0} m_{\Phi} \left(-\frac{z}{T_0} \right).$$

Moreover, let T be a solution to the approximate self-consistent equation

$$(6.12) \quad T = -\frac{1}{T} m_{\Phi} \left(-\frac{z}{T} \right) + \mathcal{E},$$

for some $\mathcal{E} \in \mathbb{C}$. Fix a parameter $\omega > 0$ and assume we are at a point z such that $\operatorname{Im}[T_0(z)] \geq \omega$. Then, there exist constants $\epsilon_{\omega} > 0$ and C_{ω} depending only on ω such that if $|\mathcal{E}| \leq \epsilon_{\omega}$ and $|T - T_0| \leq \epsilon_{\omega}$, then we have

$$(6.13) \quad |T - T_0| \leq C_{\omega} |\mathcal{E}|.$$

Remark 6.4. The condition that $\text{Im}[T_0] \geq \omega$ means that we consider values of z whose real part corresponds to the bulk of the spectrum.

Proof. Following the manipulations of (6.8) in the previous Lemma 6.2, we can assert that

$$(6.14) \quad (T - T_0) \left(1 - \int_{\mathbb{R}} \frac{x}{(Tx + z)(T_0x + z)} d\rho_{\Phi}(x) \right) = \mathcal{E}.$$

We will write $Tx + z = T_0x + z + (T - T_0)x$ and expand the denominator. By Lemma 4.9, we see that $|T_0x + z| \geq c\text{Im}[T_0] \geq c\omega$ on the support of $d\rho_{\Phi}$. Hence

$$(6.15) \quad \begin{aligned} & \left| \int_{\mathbb{R}} \frac{x}{(Tx + z)(T_0x + z)} d\rho_{\Phi}(x) - \int_{\mathbb{R}} \frac{x}{(T_0x + z)^2} d\rho_{\Phi}(x) \right| \\ & \leq \int_{\mathbb{R}} \frac{x^2 |T - T_0| dx}{|T_0x + z|^2 |T_0x + z| - |(T - T_0)x|} d\rho_{\Phi}(x) \\ & \leq \frac{C^2 \epsilon_{\omega}}{(c\omega)^2 |c\omega - \epsilon_{\omega}|} \end{aligned}$$

where C is an upper bound on the support of $d\rho_{\Phi}$. Writing $K(z)$ for this difference, we have shown that, for sufficiently small ϵ_{ω} ,

$$(6.16) \quad (T - T_0) \left(1 - \int_{\mathbb{R}} \frac{x}{(T_0x + z)^2} d\rho_{\Phi}(x) - K(z) \right) = \mathcal{E}$$

where $|K(z)| \leq \frac{C^2 \epsilon_{\omega}}{c^2 \omega^2 (c\omega - \epsilon_{\omega})}$.

We claim that there exists c_{ω} such that $1 - \int_{\mathbb{R}} \frac{x \rho_{\Phi}(x) dx}{(T_0x + z)^2} > c_{\omega}$. The claim then follows by making ϵ_{ω} sufficiently small, which we point out also makes $|K(z)|$ small.

Recall that from looking at the imaginary parts of the self-consistent equation, we have the relation

$$(6.17) \quad \text{Im}[T_0] \left(1 - \int_{\mathbb{R}} \frac{x}{|T_0x + z|^2} d\rho_{\Phi}(x) \right) = \text{Im}[z] \int_{\mathbb{R}} \frac{1}{|T_0x + z|^2} d\rho_{\Phi}(x),$$

which implies that $\int_{\mathbb{R}} \frac{x}{|T_0x + z|^2} d\rho_{\Phi}(x) < 1$. By using the fact that T_0 has strictly positive imaginary part $> \omega$, we can deduce that there is an even bigger gap between 1 and $\int_{\mathbb{R}} \frac{x}{(T_0x + z)^2} d\rho_{\Phi}(x)$.

Notice that in the support of $d\rho_{\Phi}$ we know that $\text{Im}[T_0x + z] \geq c\omega$ and $|T_0x + z| \leq C(z)$ for some constant $C(z) > 0$. When writing $T_0x + z = |T_0x + z| e^{i\theta(x, z)}$ in polar coordinates, we find that $\sin \theta > \frac{c\omega}{C(z)}$. In particular, there is some θ_{ω} such that $\theta_{\omega} \leq \theta(x, z) \leq \pi - \theta_{\omega}$. We have

$$(6.18) \quad \frac{1}{(T_0x + z)^2} = \frac{e^{-2i\theta(x, z)}}{|T_0x + z|^2}.$$

By integrating this over x , we obtain

$$(6.19) \quad \int_{\mathbb{R}} \frac{x}{(T_0 x + z)^2} d\rho_{\Phi}(x) = \int_{\mathbb{R}} e^{-2i\theta(x,z)} \frac{x}{|T_0 x + z|^2} d\rho_{\Phi}(x).$$

Recalling that $\int_{\mathbb{R}} \frac{x d\rho_{\Phi}(x)}{|T_0 x + z|^2} < 1$ and $2\theta_{\omega} \leq 2\Theta \leq 2\pi - 2\theta_{\omega}$, this relation implies that there is a strictly positive gap between 1 and $\int_{\mathbb{R}} \frac{x d\rho_{\Phi}(x)}{(T_0 x + z)^2}$ that depends only on ω .

Thus, for sufficiently small ϵ_{ω} , we know that there exists some constant C_{ω}^{-1} such that

$$(6.20) \quad \left| 1 - \int_{\mathbb{R}} \frac{x}{(T_0 x + z)^2} d\rho_{\Phi}(x) - K \right| > C_{\omega}^{-1}.$$

Considering (6.16), this implies that $|T - T_0| \leq C_{\omega} |\mathcal{E}|$. \square

6.4 Comparison to the limiting self-consistent equation. Recall that $m_{\infty}(z)$ is the unique solution of the limiting self-consistent equation (2.14). Here we prove that it is close to $M_N(z)$, the solution of the auxiliary self-consistent equation (6.1) that was analyzed in the preceding subsections

We can combine our discussion in the following lemma:

Lemma 6.5. *Assume that $\text{Im}[M_n(z)] \geq \omega$. Then*

$$(6.21) \quad |M_N(z) - m_{\infty}(z)| \leq C \frac{(\log N)^{12}}{N}.$$

The proof idea is that, while the $N \times N$ matrix Φ^N does not have an explicitly computable eigenvalue distribution, a logarithmic-rank perturbation will. Since Φ^N was a band matrix with band size $(\log N)^6$, we can extend the band by wrapping around the other side of the matrix, as if it were a torus (such matrices are called circulant matrices) and then explicitly compute the spectrum.

Proof. We define the circulant matrix

$$(6.22) \quad (\tilde{\Phi}^N)_{i,j} = \phi(i - j), \quad \text{if } |i - j \bmod N| \leq (\log N)^6.$$

Note that $\tilde{\Phi}^N$ is a Hermitian matrix and a rank $2(\log N)^6$ perturbation of Φ^N . By taking the finite Fourier series, we see that $\tilde{\Phi}^N$ has eigenvalues explicitly given by

$$\tilde{e}_k := \sum_{j=-(\log N)^6}^{(\log N)^6} \phi(j) e^{2\pi i \frac{k}{N} j}$$

where k can take any integer value between 0 and $N-1$. We order these eigenvalues as

$$(6.23) \quad \tilde{e}_{\pi(1)} \leq \tilde{e}_{\pi(2)} \leq \tilde{e}_{\pi(3)} \leq \cdots \leq \tilde{e}_{\pi(n)}$$

for an appropriate permutation $\pi \in S_N$.

Observe the following inequality for some constant C :

$$(6.24) \quad |\tilde{e}_k - \tilde{e}_{k+1}| \leq \sum_{j=-(\log N)^6}^{(\log N)^6} |\phi(j)| |e^{2\pi i \frac{j}{N}} - 1| \leq C \sum_{j=-(\log N)^6}^{(\log N)^6} |\phi(j)| \frac{(\log N)^6}{N}.$$

This inequality implies that $|\tilde{e}_{\pi(k)} - \tilde{e}_{\pi(k+1)}| \leq C \frac{(\log N)^6}{N}$ for a potentially different constant C .

If we now denote the eigenvalues of Φ^N by

$$(6.25) \quad e_{\pi(1)} \leq e_{\pi(2)} \leq e_{\pi(3)} \leq \cdots \leq e_{\pi(n)},$$

then we see that by interlacing (since the difference between Φ^N and $\tilde{\Phi}^N$ is a rank $2(\log N)^6$ matrix), we have

$$(6.26) \quad \tilde{e}_{\pi(j-2(\log N)^6)} \leq e_{\pi(j)} \leq \tilde{e}_{\pi(j+2(\log N)^6)}.$$

Moreover, this implies that for $2(\log N)^6 < j < N - 2(\log N)^6$ we have

$$(6.27) \quad |e_{\pi(j)} - \tilde{e}_{\pi(j)}| \leq \frac{C(\log N)^{12}}{N}.$$

We will assume in what follows that we want to show that $M_N(z)$ is an approximate solution to the limiting self-consistent equation (2.14) instead of (6.1), as it was previously. The idea is now to use these estimates to show that $T \equiv M_N(z)$ is an approximate solution to (2.14) and conclude by stability. We start from

$$(6.28) \quad \begin{aligned} & \left| T - \frac{1}{T} m_\Phi \left(-\frac{z}{T} \right) \right| \\ &= \left| \frac{1}{T} m_{\Phi^N} \left(-\frac{z}{T} \right) - \frac{1}{T} m_\Phi \left(-\frac{z}{T} \right) \right| \\ &= \left| \sum_{j=1}^N \left[\frac{1}{N} \frac{1}{Te_{\pi(j)} + z} - \int_{\frac{\pi(j)}{N}}^{\frac{\pi(j)+1}{N}} \frac{dx}{T \sum_{j=-\infty}^{\infty} \phi(j) e^{2\pi i j x} + z} \right] \right| \\ &\leq \sum_{j=1}^N \int_{\frac{\pi(j)}{N}}^{\frac{\pi(j)+1}{N}} \left| -\frac{1}{Te_{\pi(j)} + z} + \frac{1}{T \sum_{j=-\infty}^{\infty} \phi(j) e^{2\pi i j x} + z} \right| dx \\ &\leq C \sum_{j=1}^N \int_{\frac{\pi(j)}{N}}^{\frac{\pi(j)+1}{N}} \left| e_{\pi(j)} - \sum_{j=-\infty}^{\infty} \phi(j) e^{2\pi i j x} \right|. \end{aligned}$$

We estimate the last term by using that, for $(\log N)^6 < j < N - (\log N)^6$ and x as

above,

$$\begin{aligned}
& \left| e_{\pi(j)} - \sum_{j=-\infty}^{\infty} \phi(j) e^{2\pi i j x} \right| \\
& \leq |e_{\pi(j)} - \tilde{e}_{\pi(j)}| + \left| \tilde{e}_{\pi(j)} - \sum_{j=-\infty}^{\infty} \phi(j) e^{2\pi i j x} \right| \\
& \leq |e_{\pi(j)} - \tilde{e}_{\pi(j)}| + \sum_{j=-(\log N)^6}^{(\log N)^6} |\phi(j)| |e^{2\pi i j x} - e^{2\pi i j \frac{\pi(j)}{N}}| \\
(6.29) \quad & + \sum_{j \in \mathbb{Z}, j \notin [-(\log N)^6, (\log N)^6]} |\phi(j)| \\
& \leq C \frac{(\log N)^{12}}{N} + C \frac{(\log N)^6}{N} \sum_{j=-(\log N)^6}^{(\log N)^6} |\phi(j)| \\
& + \sum_{j \in \mathbb{Z}, j \notin [-(\log N)^6, (\log N)^6]} |\phi(j)|.
\end{aligned}$$

The other eigenvalues with $j \leq (\log N)^6$ or $j \geq N - (\log N)^6$ are bounded by uniform a priori bounds. It follows that

$$(6.30) \quad T = -\frac{1}{T} m_{\Phi} \left(-\frac{z}{T} \right) + \mathcal{E},$$

with the error estimate

$$(6.31) \quad |\mathcal{E}| \leq C \frac{(\log N)^{12}}{N} + C \sum_{j \in \mathbb{Z}, j \notin [-(\log N)^6, (\log N)^6]} |\phi(j)|.$$

By applying our stability result, Lemma 6.3, we see that for N sufficiently large,

$$(6.32) \quad |M_N(z) - m_{\infty}(z)| \leq C \frac{(\log N)^6}{N} + C \sum_{j \in \mathbb{Z}, j \notin [-(\log N)^6, (\log N)^6]} |\phi(j)|.$$

The claim now follows by estimating the second term, using Definition (2.8) of ϕ , the fact that $f \in C^2(\mathbb{T})$ and the mean-value theorem. \square

6.5 Application to Green's function estimates. In this subsection, we show that we can recover the Green's function block G^2 from solutions $m_{\infty}(z)$ to the limiting scalar self-consistent equation (2.14). The exact same method will give us bounds on the entries of G^3 , G^4 , and the off-diagonal entries of G^1 ,

Lemma 6.6. *Let (T^1, G^2) be a solution to the system of equations (4.18). Assume in addition to this that $|T^1 - m_\infty(z)| \leq \mathcal{E}$ where $\mathcal{E} = o(1)$ as $N \rightarrow \infty$. Also recall the other matrices G^1, G^3 and G^4 from (4.5). In this proof we will let $T_0(z) \equiv m_\infty(z)$, the solution to the limiting self-consistent equation (2.14). Then*

$$(6.33) \quad \begin{aligned} \|G^1 - T_0(z)\mathbf{1}\| &\leq C(\log N)^6 \|E_1\|_\infty + C(\log N)^{12} \mathcal{E}, \\ \|G^2 + (\Phi^N T_0(z) + z)^{-1}\|_\infty &\leq C(\log N)^6 \|E_2\|_\infty + C(\log N)^{12} \mathcal{E}, \\ \|G^3\|_\infty &\leq C(\log N)^6 \|E_3\|_\infty + C(\log N)^{12} \mathcal{E}, \\ \|G^4\|_\infty &\leq C(\log N)^6 \|E_4\|_\infty + C(\log N)^{12} \mathcal{E}. \end{aligned}$$

Proof. We will drop the argument z from $T_1(z)$ and $T_0(z)$ when the context is obvious. It suffices to prove the result for G^2 . All other blocks are similar.

It follows from (4.18) that

$$(6.34) \quad G^2 = -(\Phi^N T_1 + z)^{-1} - (\Phi^N T_1 + z)^{-1} E_2.$$

We recall $T_0 \equiv m_\infty(z)$. We expand the first resolvent using the resolvent identity and find that

$$(6.35) \quad (\Phi^N T_1 + z)^{-1} = (\Phi^N T_0 + z)^{-1} - (\Phi^N T_0 + z)^{-1} (T_1 - T_0) (\Phi^N T_1 + z)^{-1}.$$

We bound the entries of $(\Phi^N T_0 + z)^{-1}$ and $(\Phi^N T_1 + z)^{-1}$ by Lemma 4.10 since they are inverses of a band matrix. The function κ is bounded as in the proof of Lemma 4.10. Hence, we obtain for every $1 \leq i, j \leq N$

$$(6.36) \quad \begin{aligned} &|((\Phi^N T_0 + z)^{-1} (\Phi^N T_1 + z)^{-1})_{ij}| \\ &\leq \sum_{k=1}^N |[(\Phi^N T_0 + z)^{-1}]_{ik}| |[(\Phi^N T_1 + z)^{-1}]_{kj}| \\ &\leq C \sum_{k=1}^N \alpha_0^{(|i-k|-(\log N)^6)_+} (\alpha^1)^{(|j-k|-(\log N)^6)_+} \\ &\leq C \sum_{k=-\infty}^{\infty} \alpha_0^{(|i-k|-(\log N)^6)_+} \sum_{k=-\infty}^{\infty} (\alpha^1)^{(|j-k|-(\log N)^6)_+} \\ &\leq C \left((\log N)^6 + \frac{1}{1-\alpha_0} \right) \left((\log N)^6 + \frac{1}{1-\alpha^1} \right) (\log N)^0, \end{aligned}$$

Here, α_0 and α^1 are the corresponding constants from the application of Lemma 4.10 to the matrices $(\Phi^N T_0 + z)$ and $(\Phi^N T_1 + z)$, respectively. The assumption that $\text{Im}[T_0] \geq \omega$ clearly shows that $\|\Phi^N T_0 + z\|$ is bounded above and below and, thus, we can assert that $\alpha_0 = 1 + O((\log N)^{-6})$ and the result of infinite summation $(1 - \alpha_0)^{-1} = O((\log N)^6)$.

In addition to this, we have already assumed that $|T_0 - T_1| = \mathcal{E} = o(1)$, so clearly we have the same upper and lower bounds on $\|\Phi^N T_1 + z\|$, and we may still assert that $\alpha^1 = 1 + O((\log N)^6)$. This shows that we have the following estimate:

$$(6.37) \quad \|G^2 + (\Phi^N T_0 + z)^{-1}\|_\infty \leq C(\log N)^6 \|E_2\|_\infty + C(\log N)^{12} \mathcal{E}. \quad \square$$

7 Establishing the local law

Once we have derived the error estimates for our self-consistent equation along with the stability estimates, we can prove a local law via a standard continuity approach.

7.1 The global law. The first goal is to establish the local law at large scales, e.g., a global law. We will establish the following theorem,

Theorem 7.1. *Let \mathcal{M} be the exact solution to the matrix of self-consistent equations (4.6), let $m_\infty(z)$ be the exact solution to the infinite self-consistent equation (2.14), and let G be the Green's function of the matrix Y as in Section 4. Let \mathcal{D} be a compact subset of \mathbb{C}^+ . Then, for $N_{\mathcal{D},v,p}$ sufficiently large depending on \mathcal{D} , v and p , we can establish the following bound for $N \geq N_{\mathcal{D},v,p}$.*

$$(7.1) \quad \begin{aligned} \mathbb{P}\left(\sup_{z \in \mathcal{D}, i,j} |G_{ij} - \mathcal{M}_{ij}| \geq \frac{N^v}{\sqrt{N\eta}}\right) &\leq N^{-vp}, \\ \mathbb{P}\left(\sup_{z \in \mathcal{D}} \left|\frac{1}{2N} \text{Tr}G(z) - m_\infty(z)\right| \geq \frac{N^v}{\sqrt{N\eta}}\right) &\leq N^{-vp}. \end{aligned}$$

Proof. We see from Lemma 4.3 that we know that the matrix G satisfies the self-consistent equation up to an error of order given by $\frac{N^{2\sigma} \Gamma^5 \gamma^3}{\sqrt{N\eta}}$. From the results of Lemmas 6.3, 6.6, and 6.5, we know that we can derive the desired result (7.1) as long as we know that $\Gamma, \gamma \lesssim 1$ for $z \in \mathcal{D}$. (Note that since G and \mathcal{M} are Lipschitz, we can derive the high probability bounds on a discrete grid that is of polynomial size and extend to the entire set \mathcal{D} without too much loss in probability.)

We have the deterministic bound that $\Gamma \leq \frac{1}{\eta}$, which will be bounded by 1 in our compact region \mathcal{D} . It suffices to establish bounds on γ . By using the Schur complement formula, we know that

$$(7.2) \quad (G_{I,I}^{(J)})^{-1} = H_{I,I} - z + H_{I,K^c} G_{K^c, K^c}^{(K)} H_{K^c, I},$$

for $K = I \cup J$, which will be a subset of $[i - 2(\log N)^6, i + 2(\log N)^6]$.

We can attempt to estimate the operator norm of the right-hand side. First, observe that we can bound the operator norm of $H_{I,I}$ by the Frobenius norm,

$$(7.3) \quad \|H_{I,I}\| \leq \sqrt{\sum_{a,b \in I} |H_{ab}|^2},$$

which can be bounded by $O(\frac{N^\sigma}{\sqrt{N\eta}})$. This only uses the fact that $|H_{ab}| = \frac{O(1)}{\sqrt{N}}$ and we have at most $O((\log N)^{12})$ terms in the sum.

Finally, we can estimate

$$(7.4) \quad \|H_{I,K^c} G_{K^c K^c}^{(K)} H_{K^c I}\| \leq \|H_{I,K^c}\| \|G_{K^c K^c}^{(K)}\| \|H_{K^c,I}\| \leq \|G_{K^c,K^c}^{(K)}\| \sum_{i \in I, k \in K^c} |H_{ik}|^2.$$

We know that $\|G_{K^c,K^c}^{(K)}\| \leq \eta^{-1}$ and $\sum_{i \in I, k \in K^c} |H_{ik}|^2 \leq O((\log N)^6)$. Combining our previous estimates by the triangle inequality, this shows that $\gamma \leq O(N^\sigma)$ in \mathcal{D} . This completes the proof. \square

7.2 Proving the Local Law. As we have seen in the previous proof, it was necessary to get a bound on γ and Γ . We will establish the local analogue first.

Lemma 7.2. *Assume that*

$$(7.5) \quad \|G - \mathcal{M}\|_\infty \leq N^{-\varepsilon},$$

for some parameter $\varepsilon > 0$. Then, we have that

$$(7.6) \quad \Gamma, \gamma \leq O(1).$$

Proof. The explicit form of the solution of the self-consistent equation ensures that $\|\mathcal{M}\|_\infty = O(1)$. Clearly, the condition (7.5) would imply that $\Gamma = O(1)$ immediately. We need to do more work to establish the same result for γ .

As is usual, we let J and I be non-intersecting subsets of

$$[i - 2(\log N)^6, \dots, i + 2(\log N)^6]$$

for some integer i .

Recall the resolvent identity,

$$(7.7) \quad G_{J,J}^{(I)} = G_{J,J} - G_{J,I}(G_{I,I})^{-1}G_{I,J}.$$

We will compare this to the corresponding quantity in M ,

$$(7.8) \quad \mathcal{M}_{J,J}^{(I)} := \mathcal{M}_{J,J} - \mathcal{M}_{J,I}(\mathcal{M}_{I,I})^{-1}\mathcal{M}_{I,J}.$$

Using the fact that $\|G - \mathcal{M}\|_\infty \leq N^{-\varepsilon}$ and the fact that $|J|, |I| \leq 2(\log N)^6$ will show after some manipulation that $\|\mathcal{M}_{J,J}^{(I)} - G_{J,J}^{(I)}\|_\infty \leq C(\log N)^{12}N^{-\varepsilon}$. With this $\|\cdot\|_\infty$ bound in hand, we can show by an application of the resolvent formula that $\|(\mathcal{M}_{J,J}^{(I)})^{-1}\| - \|(G_{J,J}^{(I)})^{-1}\| = O(1)$. It suffices to understand the operator norm of $\|(\mathcal{M}_{J,J}^{(I)})^{-1}\|$ in order to understand $\|(G_{J,J}^{(I)})^{-1}\|$.

We note that it is an algebraic fact that $\mathcal{M}_{J,J}^{(I)}$ satisfies the equation

$$(7.9) \quad \mathcal{M}_{J,J}^{(I)} = (-z - \Xi^{(I)}(\mathcal{M}^{(I)}))_{J,J}^{-1},$$

where $\Xi^{(I)}$ is an operator like in (4.7), but with the covariance terms ζ that involve any index of (I) to be set to 0. This representation allows us to determine a lower bound on the norm of $\|\mathcal{M}_{J,J}^{(I)}\|$:

$$(7.10) \quad \|\mathcal{M}_{J,J}^{(I)}\| \geq \sup_{\|v\|=1, v \in \mathbb{R}^{|I|}} \text{Im}[v^* \mathcal{M}_{J,J}^{(I)} v] \geq \sup_v \frac{\text{Im}[v^* \Xi^{(I)}(\mathcal{M}_{J,J}^{(I)}) v]}{\|z + \Xi^{(I)}(\mathcal{M})\|^2}.$$

The covariance structure of $\Xi^{(I)}$ and the fact that $\text{Im}(\mathcal{M})$ is positive Hermitian allows one to assert that $\text{Im}[v^* \Xi^{(I)}(\mathcal{M}_{J,J}^{(I)}) v]$ is bounded below. This will show that $\|\mathcal{M}_{J,J}^{(I)}\|$ is bounded below, so $\|(\mathcal{M}_{J,J}^{(I)})^{-1}\|$ is bounded above, as desired. \square

Now, we have the necessary estimate in order to complete our proof of the local law.

Lemma 7.3. *Fix some constant $\omega > 0$ and recall the solution $m_\infty(z)$ to the self-consistent equation from (2.14). Let \mathcal{D} be a subset of \mathbb{C}^+ such that if $z \in \mathcal{D}$ then $\text{Im}[m_\infty(z)] \geq \omega$.*

Fix some parameters $v > 0$, $\sigma > 0$ and $p > \frac{100}{\sigma}$. We consider the subset $\mathcal{D}_v : \mathcal{D} \cap \{z : \text{Im}[z] \geq N^{-1+v}\}$. There is some $N_{v,\sigma,p}$ such that for $N \geq N_{v,\sigma,p}$ we could derive the following probability estimate:

$$(7.11) \quad \begin{aligned} \mathbb{P}\left(\sup_{z \in \mathcal{D}_v} \left| \frac{1}{2N} \text{Tr}G(z) - m_\infty(z) \right| \geq \frac{N^\sigma}{\sqrt{N\eta}}\right) &\leq N^{-\sigma p}, \\ \mathbb{P}\left(\sup_{z \in \mathcal{D}_v, i,j} |G_{ij} - \mathcal{M}_{ij}| \geq \frac{N^\sigma}{\sqrt{N\eta}}\right) &\leq N^{-\sigma p}. \end{aligned}$$

Proof. As mentioned before, we show that it suffices to prove the probability bounds on a sufficiently dense grid on the set \mathcal{D}_v . First observe that the Lipschitz constants of the function G_{ij} and the corresponding values in \mathcal{M} are bounded by $\frac{1}{\eta^2} \leq N^2$. We see that if we prove the high probability bounds on a grid whose grid distance is N^{-4} , we would be able to derive high probability bounds uniformly over \mathcal{D}_v .

Now, we turn to establishing the result over a finite grid \mathcal{G} in \mathcal{D}_v . Let z be a grid point in \mathcal{G} and let z_1, z_2, \dots, z_k be the grid points in \mathcal{G} whose real part matches that of z . The points are ordered so that $\text{Im}[z_1] \geq \text{Im}[z_2] \geq \dots \geq \text{Im}[z_k]$. First assume that we have established the high probability bounds as in (7.11) for some z_i . We will now prove the results for z_{i+1} .

Since we know the local law at z_i , we know that

$$(7.12) \quad \|\mathcal{M}(z_k) - G(z_k)\|_\infty \leq \frac{N^\sigma}{\sqrt{N\eta}},$$

with high probability.

By using the Lipschitz continuity of \mathcal{M} and G in z , this will also establish that

$$(7.13) \quad \|\mathcal{M}(z_{k+1}) - G(z_{k+1})\|_\infty \leq N^{-\varepsilon},$$

for some $\varepsilon > 0$. This is good enough to apply Lemma 7.2 as well as the self-consistent estimates in Lemmas 6.3 and 6.6. This shows that we can get the desired high probability bound at the point z_{k+1} . Taking a union bound over all elements will give us a high probability bound over our grid \mathcal{G} and establish the local law. \square

8 The comparison to the Gaussian

A very powerful tool in proving universality of various random matrix models is the study of the Dyson Brownian motion. The study was initially pioneered in a series of papers by Erdos, Schlein, Yau, and collaborators [22, 21, 27] and culminating in an optimal time proof of universality in [28]. The study of the Dyson Brownian motion has since been used to great effect in many papers, such as [17, 23, 24, 30, 25, 38]. In this section, we apply the Dyson-Brownian motion to prove universality for the ensemble \mathcal{H} .

8.1 Local law estimates under interpolation. We consider the evolution of the Green's function under the modified Ornstein–Uhlenbeck (OU) process given as follows. Recall $H = [Y_{ij}]$, our $2N \times 2N$ Hermitian block matrix with the $N \times N$ diagonal blocks set to 0, and let dB_{ij} be a matrix-valued Brownian motion with correlation structure given by

$$(8.1) \quad \text{Cov}[B_{ab}(t)\bar{B}_{cd}(t)] = t\xi_{abcd}.$$

Under this convention, we see that $B_{ab} = 0, B_{N+a, N+b}$ for $1 \leq a, b \leq N$.

We consider the matrix evolution on H to be given by

$$(8.2) \quad dH_{ab}(t) = \frac{1}{\sqrt{N}}dB_{ab} - \frac{1}{2}H_{ab}(t),$$

with $H(0) = H$ being our initial matrix and $H(t)$ the result after running the Brownian motion for time t .

By our choice of Brownian motion, we see that $H(t)$ has the same covariance and independence structure as the matrix $H(0)$. Thus, we can show that a local law holds for $H(t)$ without much difficulty.

The following integration-by-parts lemma will be useful in understanding the time evolution of functions of the matrix $H(t)$.

Lemma 8.1. *Let (x_1, \dots, x_N) be an array of J correlated random variables with mean 0 (where J is allowed to be a function of N). Assume further that $\mathbb{E}[|x_k|^3]$ is bounded uniformly for all k . Pick some index $i \in [1, \dots, N]$ and let \mathcal{T} be the set of indices that are correlated with i . Then, we have the following relation:*

$$(8.3) \quad \mathbb{E}[f(x_1, \dots, x_N)x_i] = \sum_{j \in \mathcal{T}} \mathbb{E}[\partial_j f] \mathbb{E}[x_j x_i] + O(|K|^2 \|D^2 f\|_\infty).$$

Proof. The proof is an exercise in applying the Taylor expansion.

Let $x^{(\mathcal{T})}$ be the tuple of integers $(x_1 \mathbb{1}(1 \notin \mathcal{T}), \dots, x_N \mathbb{1}(N \notin \mathcal{T}))$. We see that if we expand f in the variables in \mathcal{T} , then we can derive the expression

$$(8.4) \quad \begin{aligned} & f(x_1, \dots, x_N)x_i \\ &= f(x^{(\mathcal{T})})x_i + \sum_{j \in \mathcal{T}} \partial_j f(x^{(\mathcal{T})})x_j x_i \\ &+ \frac{1}{2} \sum_{k,j \in \mathcal{T}} (1-t) \int_0^1 \partial_{k,j} f(x^{(\mathcal{T})} + t(x - x^{(\mathcal{T})}))x_k x_j x_i dt. \end{aligned}$$

We can bound,

$$(8.5) \quad \begin{aligned} & \frac{1}{2} \sum_{k,j \in \mathcal{T}} \int_0^t \mathbb{E}[\partial_{k,j} f(x^{(\mathcal{T})} + t(x - x^{(\mathcal{T})}))x_k x_j x_i] dt \\ & \leq \frac{1}{2} |\mathcal{T}|^2 \|D^2 f\|_\infty \mathbb{E}[|x_k|^3]^{1/3} \mathbb{E}[|x_j|^3]^{1/3} \mathbb{E}[|x_i|^3]^{1/3}. \end{aligned}$$

We now let U_j be the set of integers that are correlated with j . We can again apply the Taylor expansion to compute the expectation of

$$(8.6) \quad \begin{aligned} & \mathbb{E}[\partial_j f(x^{(\mathcal{T})})x_j x_i] \\ &= \mathbb{E}[\partial_j f(x^{(\mathcal{T} \cup U_j)})x_j x_i] \\ &+ \sum_{k \in U_j \setminus \mathcal{T}} \int_0^1 (1-t) \mathbb{E}[\partial_j \partial_k f(x^{(\mathcal{T} \cup U_j)} + t(x^{(\mathcal{T})} - x^{(\mathcal{T} \cup U_j)}))x_j x_i x_k] dt. \end{aligned}$$

The second term on the right hand side above can be estimated as we have done previously.

We see that

$$(8.7) \quad \mathbb{E}[\partial_j f(x^{(\mathcal{T} \cup U_j)}) x_j x_i] = \mathbb{E}[\partial_j f(x^{(\mathcal{T} \cup U_j)})] \mathbb{E}[x_j x_i].$$

We can now reverse the application of the Taylor expansion and write

$$(8.8) \quad \begin{aligned} & \mathbb{E}[\partial_j f(x^{(\mathcal{T} \cup U_j)})] \\ &= \mathbb{E}[\partial_j f(x)] - \sum_{k \in U_j \cup \mathcal{T}} \int_0^1 (1-t) \mathbb{E}[\partial_k \partial_j f(x^{(\mathcal{T} \cup U_j)} + t(x - x^{(\mathcal{T} \cup U_j)}))] dt. \end{aligned}$$

Substituting this expression inside $\mathbb{E}[\partial_j f(x^{(\mathcal{T} \cup U)})] \mathbb{E}[x_j x_i]$ gives us the expression $\mathbb{E}[\partial_j f(x)] \mathbb{E}[x_j x_i]$ plus an error term expression which can be bounded in the same way as we have done previously. This completes the proof of the expression. \square

We will apply the previous lemma when we compute the time evolution of functions of $H(t)$.

Lemma 8.2. *Let f be a function in C^3 from $\mathbb{C}^{2N \times 2N} \rightarrow \mathbb{C}$. Then, we have the following relation,*

$$(8.9) \quad \mathbb{E}[f(H(t))] - \mathbb{E}[f(H(0))] = O(tN^{1/2}(\log N)^{12} \mathbb{E}[\|D^3 f\|_\infty]).$$

Proof. We start with applying Ito's Lemma. We see that

$$(8.10) \quad d\mathbb{E}[f(H(t))] = -\frac{1}{2} \sum_{ab} \mathbb{E}[\partial_{ab} f(H(t)) x_{ab}] + \frac{1}{2N} \sum_{ab,cd} \mathbb{E}[\partial_{ab} \overline{\partial_{cd}} f(H(t))] \xi_{abcd}.$$

To evaluate the first term on the right-hand side of the above equation, we may apply Lemma 8.1. We see that we may derive

$$(8.11) \quad -\frac{1}{2} \sum_{ab} \sum_{cd} \mathbb{E}[\partial_{ab} \overline{\partial_{cd}} f(H(t))] \xi_{abcd} + O((\log N)^{12} \mathbb{E}[\|D^3 f\|_\infty] N^{-3/2}).$$

Here, we used the fact that $\mathbb{E}[|H_{ij}(t)|^3] = O(N^{-3/2})$.

In what follows, it will be useful to state exactly what derivatives we need to control in the expression $\|D^3 f\|_\infty$, rather than apply a supremum bound.

Given a pair (i, j) , we define the set $\mathcal{T}^{i,j}$ as follows:

$$\mathcal{T}^{i,j} = \{(i', j') : |i - i'| \leq 4(\log N)^{12} \text{ or } |j - j'| < 4(\log N)^{12}\}.$$

Essentially, if one lets $\tilde{\mathcal{T}}^{i,j}$ be the set of indices of entries that could be correlated with $H_{ij}(t)$, then $\mathcal{T}^{i,j}$ is the set of indices of entries that could be correlated with entries whose indices are in $\tilde{\mathcal{T}}^{i,j}$.

When we apply the Taylor expansion, we see that we consider expressions of the form

$$(8.12) \quad \partial_{ij} \sum_{ab,cd \in \mathcal{T}^{i,j}} \partial_{ab} \overline{\partial_{cd} f}(H(t)^{(\mathcal{T}^{i,j})} + \theta(H(t) - H(t)^{(\mathcal{T}^{i,j})})).$$

Here, we apply the notation from Lemma 8.1 to let $H(t)^{(\mathcal{T}^{i,j})}$ represent the matrix $H(t)$ with certain entries set to 0.

In the proof of the previous lemma, θ is a constant between 0 and 1. \square

With the above lemma in hand, we can now establish a Green's function comparison theorem.

Lemma 8.3. *Recall the setting of Lemma 3.3, namely, let $n \geq 1$, $\varepsilon > 0$ and let E_1, \dots, E_n satisfy $\rho_\infty(E_i) \geq \varepsilon$, where ρ_∞ is the density associated with m_∞ . Given $\sigma_1, \dots, \sigma_n \leq \sigma$, we set*

$$z_j = E_j + i\eta_j, \quad \eta_j = N^{-1-\sigma_j}.$$

Consider the matrix dynamics $H(t)$ with $H(0)$ coming from our initial matrix distribution as in (3.3). We let G^t be the Green's function of $H(t) - z$ with normalized trace m^t and G^0 be the Green's function of $H(0) - z$ with normalized trace m^0 . Then, there exists $C_\sigma > 0$ depending only on σ such that

$$\left| \prod_{k=1}^n \operatorname{Im} m^t(z_k) - \prod_{k=1}^n \operatorname{Im} m^0(z_k) \right| \leq \frac{C_\sigma}{N}.$$

Proof. We will prove the comparison when $n = 1$. The proof of the general statement follows similar details.

We will try to apply the previous Lemma 8.2. We see that it suffices to derive a bound on the third derivatives,

$$(8.13) \quad \mathbb{E} \left| \partial_{ab} \partial_{cd} \partial_{ef} \frac{1}{2N} \operatorname{Tr}(G^{\theta,s}(z)) \right|.$$

Here, cd and ef are entries in $\mathcal{T}^{a,b}$ and $G^{\theta,s}(z)$ is the Green's function of the matrix $H(s) - z + \theta(H^{(\mathcal{T}^{a,b})}(s) - H(s))$, where θ is a constant between 0 and 1.

First fix a time s and set $\theta = 0$. We will first establish a bound here before discussing the general case.

By direct computation, one can see that

$$\left| \partial^\alpha \frac{1}{2N} \operatorname{Tr}[G^s(z)] \right| \leq \Gamma^4(z),$$

where, recall, $\Gamma(z)$ is a uniform upper bound for the entries of $G^s(z)$ and ∂^α indicates any third-order partial derivative.

Again, by direct computation, one can see that the change of Γ as the imaginary part η of z changes satisfies a useful inequality,

$$\left| \frac{\partial \Gamma}{\partial \eta} \right| \leq \frac{\Gamma}{\eta}.$$

One can integrate this differential equation to see that

$$(8.14) \quad \Gamma(E + i\eta) \leq \Gamma(E + iN^{-1+\epsilon})N^{2\epsilon}$$

whenever $N^{-1-\epsilon} \leq \eta \leq N^{-1}$. Now, when E is in the bulk of the distribution, we can apply our local law to ensure that $\Gamma(E + iN^{-1+\epsilon}) = O(N^\epsilon)$. Thus, with probability $1 - N^{-D}$ for some large D , we could ensure that $|\partial^{|\alpha|} \frac{1}{2N} \text{Tr}[G^s(z)]| \leq CN^{12\epsilon}$ and the trivial bound N^8 otherwise.

By Lipschitz continuity, one can establish these results on a discrete grid of times and extend to the entire interval $[0, t]$. In addition, one can show that matrices of the form $H(t)^{\mathcal{T}^{i,j}} + \theta(H(t) - H(t)^{\mathcal{T}^{i,j}})$ satisfy a similar local law. Again, applying local law results to a discrete grid of θ s and noting the fact that there are no more than N^2 choices of these special $\mathcal{T}^{i,j}$ modifications will allow us to get a uniform probability bound on all choices of θ and i, j . This gives us a desired proof of the bounds on the derivatives we need to apply Lemma 8.2 and complete the proof of the Theorem. \square

The results of the above Green's function comparison theorem can be used to prove the following comparison on correlation functions, as we have seen earlier in the proof of Lemma 3.4.

Theorem 8.4. *Fix a time $t = N^{-1+\epsilon}$ with $\epsilon > 0$. Consider the matrix dynamics $H(t)$ with $H(0)$ coming from our initial matrix distribution as in (3.3). Let $p_N^{(k),t}$ be the correlation functions of $H(t)$. Let ρ be the density corresponding to the limiting spectral distribution of $H(0)$, and let E be a point in the support of ρ . Then, for any compactly supported continuous test function O from $\mathbb{R}^k \rightarrow \mathbb{R}$, we have the following comparison estimate:*

$$(8.15) \quad \int_{\mathbb{R}^k} O(\alpha) \left[p_N^{(k),t} \left(E + \frac{\alpha}{2N} \right) - p_N^{(k),0} \left(E + \frac{\alpha}{2N} \right) \right] d\alpha = O(N^{-c}).$$

8.2 Comparing to the GOE. At this point, we have established that the statistics of $H(0)$ match those of $H(t)$. We will be finished once we show that the statistics of $H(t)$ match those of the GOE.

However, recall from our interpolation that $H(t)$ has a correlated Gaussian component

$$\sqrt{t}N \begin{bmatrix} 0 & C \\ C^* & 0 \end{bmatrix},$$

where G has the correlation structure given by $\mathbb{E}[C_{ab}C_{cd}] = \xi_{abcd}$. However, because our covariance matrix $\xi < c_0$ is positive semidefinite, we can split the matrix C as $C = \tilde{C} + G$ where \tilde{C} and G are independent Gaussians and G is a GUE matrix.

Thus, the matrix $H(t)$ can be represented in the form $\tilde{H} + cGUE$ for $c \geq N^{-1+\epsilon}$ and \tilde{H} independent of the GUE. The results in Section 3 of [17] (see also Theorem 2.2 of [38]) prove that the matrix $H(t)$ will have universal spectral statistics. Theorem 8.4 shows that $H(0)$ will have the same spectral statistics as $H(t)$. This proves Theorem 3.2. Finally, using Lemma 3.4, this will further prove Theorem 2.10.

Appendix A On the limiting objects $m_\infty(z)$ and $\rho_\infty(z)$

In this appendix, we prove Propositions 2.5 and 2.7.

Proof of Proposition 2.5. This was essentially already proved in the main text in the pre-limit case. Existence of m_∞ follows from Brouwer's fixed point theorem as in the proof of Lemma 6.1. Uniqueness is proved by following the argument establishing Lemma 6.2. We omit the details. \square

Proof of Proposition 2.7. In a first step, we prove that $m_\infty(z)$ is the Stieltjes transform of a Borel measure on \mathbb{R} . The analytic implicit function theorem and the condition on the imaginary part imply that $m_\infty(z)$ is a Herglotz function. Hence, the Herglotz representation theorem yields constants $a \in \mathbb{R}$, $b \geq 0$, and a Borel measure $d\rho_\infty$ on \mathbb{R} satisfying $\int_{\mathbb{R}} \frac{1}{1+x^2} d\rho_\infty(x) < \infty$ such that

$$m_\infty(z) = a + bz + \int_{\mathbb{R}} \left(\frac{1}{x-z} - \frac{1}{1+x^2} \right) d\rho_\infty(x)$$

Let G be an $N \times N$ Gaussian matrix with the same correlation structure as the matrix Y from (3.2) and define its Hermitization

$$H_G = \begin{pmatrix} 0 & G \\ G^\dagger & 0 \end{pmatrix}.$$

We can repeat the proof of the local law that was given for \tilde{H} for the matrix H_G because only the correlation structure and range of dependence (which equals the

range of correlation for a Gaussian matrix) is used. The upshot is that $m_\infty(z)$ arises as the limiting spectral density of the Gaussian matrix ensemble H_G . Since H_G are Hermitian matrices, this implies that

$$|m_\infty(z)| \leq \frac{1}{\operatorname{Im} z}, \quad \operatorname{Im} z > 0.$$

By considering the asymptotics of this estimate for $z = iy$ with $y \rightarrow \infty$, the Herglotz representation formula reduces to

$$(A.1) \quad m_\infty(z) = \int_{\mathbb{R}} \frac{1}{x - z} d\rho_\infty(x),$$

for a finite Borel measure $d\rho_\infty(x)$.

In a second step, we use free probability theory to prove that the Borel measure $d\rho_\infty(x)$ in (A.1) has a continuous density. Alternatively, this can be derived from general results about the matrix Dyson equation such as, e.g., [6, Corollary 4.14].

Recall that f is an admissible evaluation function. By Lemma 4.9, this implies that the correlation matrix $\Phi^N \geq \frac{g_{\min}}{2} > 0$ is strictly positive definite. Hence, we can decompose

$$G = \frac{g_{\min}}{2} G_1 + G_2$$

where G_1 is a Ginibre matrix (independent Gaussian entries with variance $\frac{1}{N}$) and G_2 is a Gaussian matrix independent of G_1 with entries of variance $\sim \frac{1}{N}$ and correlation matrix $\geq \frac{g_{\min}}{2}$. This decomposition extends to the Hermitization

$$H_G = \frac{g_{\min}}{2} H_1 + H_2, \quad H_i = \begin{pmatrix} 0 & G_i \\ G_i^\dagger & 0 \end{pmatrix}.$$

Both matrices H_1 and H_2 satisfy a local law for respective limiting densities $d\rho_1(x) \equiv d\rho_{\text{sc}}(x)$ and $d\rho_2(x)$. Thus $d\rho_\infty(x)$ arises as the free convolution of $d\rho_2$ with the Wigner semicircle law. A result of Biane [12, Cor. 2] then says that $d\rho_\infty(x)$ has a continuous density. It is computable from (A.1) via the Stieltjes inversion formula. \square

Appendix B Heuristic derivation of the self-consistent equation

Recall that, by definition, we have the equation $GY - zG = I$. To derive the form of the self-consistent equation, it would be required to consider a matrix \tilde{Y} of Gaussian random variables whose covariance structure matches the covariance structure of Y ; thus, we consider the equation $G\tilde{Y} - zG = I$, take the expectation

of both sides of the equation, and simplify by integrating by parts with respect to the Gaussian variables in \tilde{Y} . This procedure will result in the equation

$$(B.1) \quad \mathbb{E}[(G\tilde{Y} - zG)_{ab}] = \mathbb{E}\left[-\frac{1}{N} \sum_{k,l,m} G_{ik}\xi_{kljm}G_{lm} - zG_{ij}\right] = \delta_{ij}.$$

Here, we wrote $(G\tilde{Y})_{ij} = \sum G_{im}\tilde{Y}_{mj}$. Observe now that $\partial_{\tilde{Y}_{kl}}G_{im} = G_{ik}G_{lm}$. Integrating by parts and using the fact that $\mathbb{E}[\tilde{Y}_{kl}\tilde{Y}_{mj}] = \xi_{kljm}$ will show that there is a prefactor of $\xi_{kljm}G_{ik}G_{lm}$ associated with this quantity. At the last step we can remove the expectation, anticipating that these quantities will be concentrated.

Acknowledgments. The authors thank an anonymous referee for helpful remarks. A.A. would like to thank NSF award number 2102842 and the Harvard GSAS Merit/Graduate Society Term-Time Research Fellowship for support during part of this project.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution and reproduction in any medium, provided the appropriate credit is given to the original authors and the source, and a link is provided to the Creative Commons license, indicating if changes were made (<https://creativecommons.org/licenses/by/4.0/>).

Open Access funding enabled and organized by Projekt DEAL.

REFERENCES

- [1] A. Adhikari and Z. Che, *Edge universality for correlated Gaussians*, Electron. J. Probab. **24** (2019), Article no. 44.
- [2] A. Adhikari and M. Lemm, *Local law for singular values of oscillatory matrices*, Int. Math. Res. Not. **2023** (2023), 3907–3947.
- [3] A. Adhikari, M. Lemm and H.-T. Yau, *Global eigenvalue distribution of matrices defined by the skew-shift*, Anal. PDE **14** (2021), 1153–1198.
- [4] O. Ajanki, L. Erdős and T. Kruger, *Local spectral statistics of Gaussian matrices with correlated entries*, J. Stat. Phys. **163** (2016), 280–302.
- [5] J. Alt, L. Erdős, T. Kruger and D. Schröder, *Correlated random matrices: band rigidity and edge universality*, Ann. Prob. **48** (2020), 963–1001.
- [6] J. Alt and T. Krüger, *Inhomogeneous circular law for correlated matrices*, J. Funct. Anal. **281** (2011), Article no. 109120.
- [7] A. Avila and S. Jitomirskaya, *The ten Martini problem*, Ann. of Math. (2) **170** (2009), 303–342.
- [8] A. Avila, D. Damanik and Z. Zhang, *Schroedinger operators with potentials generated by hyperbolic transformations: I. positivity of the Lyapunov exponent*, arXiv:2011.10146 [math.SP]
- [9] M. Banna, F. Merlevéde and M. Peligrad, *On the limiting spectral distribution for a large class of random matrices with correlated entries*, Stochastic Process. Appl. **125** (2015), 2700–2726.

- [10] R. Bauerschmidt, J. Huang, A. Knowles and H.-T. Yau, *Bulk eigenvalue statistics for random regular graphs*, Ann. Probab. **45** (2017), 3626–3663.
- [11] R. Bauerschmidt, A. Knowles and H.-T. Yau, *Local semicircle law for random regular graphs*, Comm. Pure Appl. Math. **70** (2017), 1898–1960.
- [12] P. Biane, *On the free convolution with a semi-circular distribution*, Indiana Univ. Math. J. **46** (1997), 705–718.
- [13] J. Bourgain, M. Goldstein and W. Schlag, *Anderson localization for Schrödinger operators on \mathbb{Z} with potentials given by skew-shift*, Comm. Math. Phys. **220** (2001), 583–621.
- [14] J. Bourgain and W. Schlag, *Anderson localization for Schrödinger operators on \mathbb{Z} with strongly mixing potentials*, Comm. Math. Phys. **215** (2000), 143–175.
- [15] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Springer, Berlin–New York, 1975.
- [16] Z. Che, *Universality of random matrices with correlated entries*, Electron. J. Probab. **22** (2017), Article no. 30.
- [17] Z. Che and P. Lopatto, *Universality of the least singular value for sparse random matrices*, Electron. J. Probab. **24** (2019), 1–53.
- [18] V. Chulaevsky and T. Spencer, *Positive Lyapunov exponents for a class of deterministic potentials*, Comm. Math. Phys. **168** (1995), 455–466.
- [19] D. Damanik, R. Killip and B. Simon, *Perturbations of orthogonal polynomials with periodic recursion coefficients*, Ann. of Math. (2) **171** (2010), 1931–2010.
- [20] B. de Monvel, A. Khorunzhy and V. Vasilchuk, *Limiting eigenvalue distribution of random matrices with correlated entries*, Markov Process. Related Fields **2** (1996), 607–636.
- [21] L. Erdős, S. Peche, J. Ramirez, B. Schlein and H.-T. Yau, *Bulk universality for Wigner matrices*, Comm. Pure Appl. Math. **63** (2010), 895–925.
- [22] L. Erdős, J. Ramirez, B. Schlein and H.-T. Yau, *Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation*, Electron. J. Probab. **15** (2010), 526–604.
- [23] L. Erdős, B. Schlein, H.-T. Yau and J. Yin, *The local relaxation flow approach to universality of the local statistics for random matrices*, Ann. Inst. Henri Poincaré Probab. Stat. **48** (2012), 1–46.
- [24] L. Erdős, H.-T. Yau and J. Yin, *Bulk universality for generalized Wigner matrices*, Probab. Theory Related Fields **154** (2012), 341–407.
- [25] L. Erdős, H.-T. Yau and J. Yin, *Rigidity of eigenvalues of generalized Wigner matrices*, Adv. Math. **229** (2012), 1435–1515.
- [26] L. Erdős, T. Krüger and D. Schröder, *Random matrices with slow correlation decay*, Forum Math. Sigma **7** (2019), Article no. e8.
- [27] L. Erdős, J. Ramirez, B. Schlein, T. Tao, V. Vu and H.-T. Yau, *Bulk universality for Wigner Hermitian matrices with subexponential decay*, Math. Res. Lett. **17** (2010), 667–674.
- [28] L. Erdős, B. Schlein and H.-T. Yau, *Universality of random matrices and local relaxation flow*, Invent. Math. **185** (2011), 75–119.
- [29] L. Erdős and H.-T. Yau, *A Dynamical Approach to Random Matrix Theory*, American Mathematical Society, Providence, RI, 2017.
- [30] L. Erdős, H.-T. Yau and J. Yin, *Universality for generalized Wigner matrices with Bernoulli distribution*, J. Comb. **2** (2011), 15–81.
- [31] R. Gray, *Toeplitz and circulant matrices: A review*, in *Foundations and Trends in Communications and Information Theory*. Vol. 2, Now Publishers, Hanover, MA, 2006, pp. 155–239.
- [32] R. Han, M. Lemm and W. Schlag, *Effective multiscale approach to the Schrödinger cocycle over a skew-shift base*, Ergodic Theory Dynam. Systems **40** (2020) 2788–2853.
- [33] W. Helton, R. Far, and R. Speicher, *Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints*, Int. Math. Res. Not. IMRN **2007** (2007), Article no. rnm086.

- [34] J. Huang and B. Landon, *Spectral statistics of sparse Erdős–Rényi graph Laplacians*, Ann. Inst. Henri Poincaré Probab. Stat. **56** (2020), 120–154.
- [35] S. Jitomirskaya and W. Liu, *Universal hierarchical structure of quasiperiodic eigenfunctions*, Ann. of Math. (2) **187** (2018), 721–776.
- [36] A. Khorunzhy, *Eigenvalue distribution of large random matrices with correlated entries*, Mat. Fiz. Anal. Geom. **3** (1996), 80–101.
- [37] H. Kruger, *Multiscale analysis for ergodic Schrödinger operators and positivity of Lyapunov exponents*, J. Anal. Math. **115** (2011), 343–387.
- [38] B. Landon, P. Sosoe and H.-T. Yau, *Fixed energy universality of Dyson Brownian motion*, Adv. Math. **346** (2019), 1137–1332.
- [39] N. Minami, *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*, Comm. Math. Phys. **177** (1996), 709–725.
- [40] T. Tao and V. Vu, *Random matrices: Universality of local eigenvalue statistics*, Acta Math. **206** (2011), 127–204.

Arka Adhikari

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CA 94305, USA
email: arkaa@stanford.edu

Marius Lemm

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TÜBINGEN
72076 TÜBINGEN, GERMANY
email: marius.lemm@uni-tuebingen.de

(Received February 14, 2022 and in revised form December 30, 2022)