Bulk universality for complex eigenvalues of real non-symmetric random matrices with i.i.d. entries

Sofiia Dubova* Kevin Yang*

April 29, 2024

Abstract

We consider an ensemble of non-Hermitian matrices with independent identically distributed real entries that have finite moments. We show that its k-point correlation function in the bulk away from the real line converges to a universal limit.

Contents

1 Introduction **2**
1.1 Notation . . . 5
1.2 Acknowledgements . . . 5

2 Change-of-variables **5**
2.1 Preliminary steps . . . 5
2.2 Re-parameterizing λ_1 and λ_2 . . . 9
2.3 Estimating $\tilde{F}(\mathbf{z}; A^{(2)})$ . . . 11
2.4 Fourier transform computation for $\tilde{K}_j(z_j)$ . . . 26

3 Estimating the integral over θ_1, θ_2 **28**
3.1 Preliminary bounds for integrating on $\mathcal{I}_1$ . . . 29
3.2 Proof of Proposition 8 . . . 34

4 Estimates for $\rho_{\text{main}}(z, \mathbf{z}; A)$ **35**
4.1 Ratio of determinants . . . 35
4.2 Estimating $\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]$ . . . 36
4.3 The dP integration over $M_2^{sa}(\mathbb{R})$ . . . 41
4.4 Putting it altogether . . . 45

5 Replacing resolvents in (4.10) by universal local law approximations **46**
5.1 Replacing the constants, e.g. $\eta_{z,t}$ . . . 46
5.2 Replacing $|\det \mathcal{G}_{j,a,b}|$ . . . 47
5.3 Replacing $\operatorname{Tr} H_{\lambda_j}(\eta_{z,t}) \tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})$ . . . 48
5.4 Replacing $\mathbf{Q}_{a_j,b_j,\theta_j}$ . . . 49
5.5 Replacing the vector in the quadratic form . . . 50

6 Proof of Theorem 2 **51**

*Department of Mathematics, Harvard University, One Oxford Street, Cambridge MA 02138, USA

7 Proof of Theorem 1 **52**
7.1 Proof of Lemma 22 53

A Resolvent estimates for $G_{a,b,\theta}(\eta)$ **56**
A.1 Properties of the operator $\mathcal{X}$. 60

B Computation of resolvent quantities at $\theta = \frac{\pi}{4}$. **61**
B.1 Computing $\mathbf{Q}_{a,b,\frac{\pi}{4}}$. 62

C Jacobian calculation **62**

1 Introduction

Universality of local eigenvalue statistics has long been of interest in random matrix theory. In the case of Hermitian Wigner ensembles this is the content of Wigner-Dyson-Mehta conjecture [24], which was resolved in [17] using a *three-step strategy*:

- Local law for the resolvent $G(z) = (H - z)^{-1}$;
- Bulk universality for the perturbed ensemble $H_t = H + \sqrt{t}B$, where B is Gaussian;
- Comparison of local statistics of H with local statistics of the perturbed matrix H_t.

This strategy has since been widely used to show universality of local statistics of other Hermitian ensembles [21, 2]. An established method of carrying out the second step of the strategy for Hermitian ensembles is Dyson Brownian motion (DBM), a system of stochastic differential equations for the eigenvalues of H_t. For non-Hermitian random matrices, the natural analogue of DBM involves the left- and right-eigenvector overlaps of the ensemble, which makes its analysis very complicated; see Appendix A in [6], for example. In [11] Cipolloni, Erdős and Schröder side-stepped the non-Hermitian DBM using Girko's formula [19] to show the universality of local statistics in the edge regime for both complex and real non-Hermitian i.i.d. ensembles. Girko's formula allows access to the eigenvalue statistics of a non-Hermitian matrix A through the resolvent $G_z(\eta)$ of the Hermitization of A:

$$G_z(\eta) = \begin{pmatrix} -i\eta & A - z \\ A^* - \bar{z} & -i\eta \end{pmatrix}^{-1}.$$

This strategy has not yielded the universality of local statistics in the bulk due to difficulty controlling the correlation of G_{z_1} and G_{z_2} in the intermediate regime $|z_1 - z_2| \asymp \frac{1}{\sqrt{N}}$. Current techniques only allow to show independence of G_{z_1} and G_{z_2} in the regime $|z_1 - z_2| > N^{-\frac{1}{2}+\varepsilon}$, see [14].

Recently, Maltsev and Osman proved bulk universality for complex non-Hermitian i.i.d. matrices in [23] using a different approach based on supersymmetry for the second step of the three-step strategy. They derived an exact integral formula for the k-point correlation function of a random complex non-Hermitian matrix A perturbed by a Gaussian $\sqrt{t}B$ and performed asymptotic analysis to show its convergence to the k-point correlation function of the complex Ginibre ensemble with Gaussian entries. This method was also used to show bulk universality for weakly non-Hermitian complex matrices in [25]. For the purposes of bulk universality, it is enough to take $t = N^{-\epsilon_0}$ (see Section 7 in [23]), though the analysis in [23] holds until the time-scale $N^{-1/2+\delta}$ assuming the initial data matrix satisfies technical (third-order) local law estimates. (Here, $\epsilon_0, \delta > 0$ are any small parameters independent of N.)

In this work we establish bulk universality for the real non-Hermitian i.i.d. ensemble away from the real line using the strategy of [23]. This builds on [28], in which the first four moments of the entries of A are assumed to match those of the Gaussian distribution. (We note that [28], which builds on the methods in [27] for Hermitian matrices, applies to both real and complex matrix ensembles and both real and complex eigenvalue statistics.) We discuss the additional complications that arise in the real case compared to the complex case after we state our main results.

To be precise, the random matrix ensemble of interest in this paper is given by $A = (A_{ij})_{i,j=1}^{N}$, where $N \gg 1$ is the matrix size, and A_{ij} are i.i.d. real-valued random variables that satisfy $\mathbb{E}A_{ij} = 0$ and $\mathbb{E}|A_{ij}|^2 = N^{-1}$ and $\mathbb{E}|A_{ij}|^{2p} \lesssim_p N^{-p}$ for all $p \geq 1$. (Throughout this paper, we write $a \lesssim b$ if $|a| \leq C|b|$, with subscripts to indicate what parameters the constant C depends on.)

Let us now introduce the main objects and constants appearing in this paper.

- First, we define $t = N^{-\epsilon_0}$ to be the time-scale of the Gaussian perturbation, where $\epsilon_0 > 0$ is independent of N.
- For any $z \in \mathbb{C}$, we define $H_z(\eta) := [(A-z)(A-z)^* + \eta^2]^{-1}$ and $\tilde{H}_z(\eta) := [(A-z)^*(A-z)+\eta^2]^{-1}$. We also define the Hermitization
$$\mathcal{H}_z := \begin{pmatrix} 0 & A - z \\ A^* - \overline{z} & 0 \end{pmatrix}$$
and recall the resolvent $G_z(\eta) = [\mathcal{H}_z - i\eta]^{-1}$. It turns out (see [23]) that
$$G_z(\eta) = \begin{pmatrix} i\eta H_z(\eta) & (A-z)H_z(\eta) \\ H_z(\eta)(A-z)^* & i\eta \tilde{H}_z(\eta) \end{pmatrix}. \tag{1.1}$$
- Let $\eta_{z,t}$ solve $t\langle H_z(\eta_{z,t})\rangle = 1$, where $\langle H \rangle = \frac{1}{\dim H} \operatorname{Tr} H$ is normalized trace for any square matrix H. See the proof of Theorem 1.1 in [23] for existence and uniqueness of such $\eta_{z,t}$ for $t = N^{-\epsilon_0}$ (and for the estimate $\frac{1}{C}t \leq \eta_{z,t} \leq Ct$, which holds whenever the local law in Lemma 26 holds).
- Set $g_{z,t} = \eta_{z,t}\langle H_z(\eta_{z,t})\rangle$ and $\alpha_{z,t} = \eta_{z,t}^2 \langle H_z(\eta_{z,t})\tilde{H}_z(\eta_{z,t})\rangle$. Set $\beta_{z,t} = \eta_{z,t}\langle H_z(\eta_{z,t})^2(A-z)\rangle$ and $\gamma_{z,t} = \eta_{z,t}^2\langle H_z(\eta_{z,t})^2\rangle$. Define $\sigma_{z,t} = \alpha_{z,t} + \gamma_{z,t}^{-1}|\beta_{z,t}|^2$.

Now, recall that the k-point correlation function (for any integer $k \geq 1$) of a point process $\{\zeta_j\}_j$ on $\mathbb{C}$ is the function $\rho(z_1, \ldots, z_k)$ satisfying
$$\mathbb{E}\left[\sum_{i_1 \neq i_2 \neq \ldots \neq i_k} \varphi(\zeta_{i_1}, \ldots, \zeta_{i_k})\right] = \int_{\mathbb{C}^k} \varphi(z_1, \ldots, z_k)\rho(z_1, \ldots, z_k) dz_1 \ldots dz_k.$$
The goal of this paper is compute the k-point correlation function in the large N limit for eigenvalues of A in the complex plane, i.e. away from the real line. To state this result precisely, let us fix $k \geq 2$ and define
$$\rho_{\mathrm{GinUE}}^{(k)}(z_1, \ldots, z_k) := \det\left[\frac{1}{\pi}\exp\left\{-\frac{|z_i|^2 + |z_j|^2}{2} + z_i \overline{z}_j\right\}\right]_{i,j=1}^{k}$$
This yields essentially the local distribution of k-many eigenvalues for a matrix whose entries are (normalized) *complex* Gaussians. Remarkably, as we further explain shortly, it also gives the local eigenvalue distribution for real, non-symmetric Gaussian ensembles away from the real line.

Theorem 1. *Fix any $k \geq 2$. For any $O \in C_c^\infty(\mathbb{C}^k)$ and $z \in \mathbb{C}$ such that $|z| < 1$ and $\operatorname{Im}(z) \neq 0$, we have that*
$$\mathbb{E}\left[\sum_{i_1 \neq i_2 \neq \ldots \neq i_k} O(N^{\frac{1}{2}}\sigma_{\mathrm{univ},z}^{\frac{1}{2}}[z - \lambda_{i_1}], \ldots, N^{\frac{1}{2}}\sigma_{\mathrm{univ},z}^{\frac{1}{2}}[z - \lambda_{i_k}])\right]$$
$$- \int_{\mathbb{C}^k} O(z_1, \ldots, z_k)\rho_{\mathrm{univ},z}^{(k)}(z_1, \ldots, z_k) dz_1 \ldots dz_k$$

vanishes as $N \to \infty$, where $\sigma_{\mathrm{univ},z} > 0$ is independent of the distribution of the entries of A, where $\{\lambda_j\}_j$ are eigenvalues of A, where

$$\rho^{(k)}_{\mathrm{univ},z}(z_1, \ldots, z_k) = \Phi_z(z_1, \ldots, z_k)\rho^{(k)}_{\mathrm{GinUE}}(z_1, \ldots, z_k),$$

and where $\Phi_z(z_1, \ldots, z_k)$ is independent of the distribution of the entries of A.

By Theorem 11 in [5], we know $\Phi_z(z_1, \ldots, z_k) \equiv 1$, i.e. that local bulk eigenvalue statistics away from the real line agree with those in the complex case. We only stated Theorem 1 in this way to parallel it with Theorem 2. We note a similar analysis of local laws as in Section 7 of [23] shows $\sigma_{\mathrm{univ},z} = 1$, but we omit this extra computation.

The main step in proving Theorem 1 is to first prove the same statement but for matrices with a Gaussian perturbation. In what follows, B is a real Ginibre matrix of size $N \times N$, i.e. its entries are independent real Gaussians with mean zero and variance N^{-1}.

Theorem 2. *Fix any $k \geq 2$. If we set $t = N^{-\epsilon_0}$, then for $\epsilon_0 > 0$ sufficiently small and for any $O \in C_c^\infty(\mathbb{C}^k)$ and $z \in \mathbb{C}$ such that $|z| < 1$ and $\mathrm{Im}(z) \neq 0$, we have that*

$$\mathbb{E}\left[\sum_{i_1 \neq i_2 \neq \ldots \neq i_k} O(N^{\frac{1}{2}}\sigma_{z,t}^{\frac{1}{2}}[z - \lambda_{i_1}(t)], \ldots, N^{\frac{1}{2}}\sigma_{z,t}^{\frac{1}{2}}[z - \lambda_{i_k}(t)])\right]$$
$$- \int_{\mathbb{C}^k} O(z_1, \ldots, z_k)\rho^{(k)}_{\mathrm{univ},z,t}(z_1, \ldots, z_k)dz_1 \ldots dz_k$$

vanishes as $N \to \infty$, where $\{\lambda_j(t)\}_j$ are eigenvalues of $A + \sqrt{t}B$, where

$$\rho^{(k)}_{\mathrm{univ},z,t}(z_1, \ldots, z_k) = \Phi_{z,t}(z_1, \ldots, z_k)\rho^{(k)}_{\mathrm{GinUE}}(z_1, \ldots, z_k),$$

and where $\Phi_{z,t}(z_1, \ldots, z_k)$ is independent of the distribution of the entries of A.

We give proofs of Theorems 1 and 2 in the case $k = 2$ for (notational) convenience. For general $k \geq 2$, the same argument with minor adjustments works; see [23] for details.

Now we discuss additional difficulties that arise in case of real non-symmetric i.i.d. ensembles in comparison with the complex case. In [23] the authors obtain the integral formula for the k-point correlation function through a change of variables in the space of complex non-Hermitian $N \times N$ matrices $M_N(\mathbb{C})$ obtained by consecutive Householder transformations with respect to the eigenvectors. In our case, since real matrices still have complex eigenvalues, we have to adapt each step of the change of variable to isolate conjugate pairs of eigenvalues. This leads to another integration parameter θ, where 2θ is the angle between the corresponding right eigenvectors. It turns out θ concentrates around the value $\pi/4$. Showing this is one of the key steps, and the main challenge is that local law estimates do not seem to help; instead, we control the region $|\theta - \pi/4| \gg N^{-1/2+\delta}$ directly (here, $\delta > 0$ is any small parameter). Additionally, to carry out the asymptotic analysis we need to show that certain resolvent quantities have universal deterministic approximations. In particular, we derive the local law and the two-resolvent local laws at scales proportional to $t = N^{-\epsilon_0}$ for the Hermitization of $I_2 \otimes A - \Lambda \otimes I_N$, where Λ is a deterministic 2×2 matrix. This is a 2×2 analogue of the 1×1 results in [7, 3, 4, 9]. We prove this using the cumulant expansion argument commonly used for the proofs of two resolvent local laws, see [12, 8, 13, 1]. Our final step is to remove the Gaussian component $\sqrt{t}B$, which amounts to a standard "three-and-a-half" moment comparison approach and the Girko formula as in [18, 10], respectively.

Finally, let us mention that approximately a week before posting to the arXiv, [26] was posted to the arXiv, in which the same result was proven by the same general strategy, but with different approaches to some of the technical challenges alluded to in the previous paragraph. (These technical differences potentially affect the smallest time-scale t achievable.) It is possible that combining the methods of our paper with those of [26] may prove Theorem 2 for the optimal time-scale $t = N^{-1/2+\delta}$. Osman [26] also proves bulk universality on the real line; for this, another method based on "spin variables" is employed to handle additional difficulties occuring therein. See [26] for details.

1.1 Notation

We use big-O notation, i.e. $a = O(b)$ means $|a| \le C|b|$ for some constant $C > 0$. Any subscripts in the big-O notation indicate what parameters the constant C depends on. We also write $a \lesssim b$ to mean $a = O(b)$ and $a \gtrsim b$ to mean $b = O(a)$, with the same disclaimer about subscripts. We use the notation $[[a, b]] = [a, b] \cap \mathbb{Z}$ for an integer range.

1.2 Acknowledgements

We thank Benjamin McKenna, Horng-Tzer Yau, and Jun Yin for very helpful discussions (and in particular Jun Yin for helping with details in the moment comparison step). We would like to thank Mohammed Osman for helpful comments on an earlier draft (and pointing out a mistake). K.Y. was supported in part by the NSF under Grant No. DMS-2203075.

2 Change-of-variables

The goal of this section is to reproduce ideas in [23] but for real non-symmetric matrices.

2.1 Preliminary steps

Define the manifolds

$$\begin{aligned}
\Omega &:= \Omega_1 \times \Omega_2 \times M_{(N-4)\times(N-4)}(\mathbb{R}) \\
\Omega_1 &:= \mathbb{R} \times \mathbb{R}_+ \times [0, \frac{\pi}{2}) \times V^2(\mathbb{R}^N) \times M_{(N-2)\times 2}(\mathbb{R}), \\
\Omega_2 &:= \mathbb{R} \times \mathbb{R}_+ \times [0, \frac{\pi}{2}) \times V^2(\mathbb{R}^{N-2}) \times M_{(N-4)\times 2}(\mathbb{R}).
\end{aligned}$$

Above, $V^2(\mathbb{R}^d)$ is the Stiefel manifold

$$V^2(\mathbb{R}^d) := \mathbf{O}(d)/\mathbf{O}(d-2) = \{(v_1, v_2) \in \mathbb{S}^{d-1} \times \mathbb{S}^{d-1} : v_1^* v_2 = 0\},$$

where $\mathbb{S}^{d-1}$ is the unit sphere in $\mathbb{R}^d$. Define the map $\Phi : \Omega \to M_{N\times N}(\mathbb{R})$ given by

$$\Phi(a_1, b_1, \theta_1, \mathbf{v}, W_1, a_2, b_2, \theta_2, \mathbf{w}, W_2, M^{(2)}) = R_1(\mathbf{v}) \begin{pmatrix} \Lambda_{a_1,b_1,\theta_1} & W_1^* \\ 0 & M^{(1)} \end{pmatrix} R_1(\mathbf{v})^*,$$

where

$$M^{(1)} = R_2(\mathbf{u}) \begin{pmatrix} \Lambda_{a_2,b_2,\theta_2} & W_2^* \\ 0 & M^{(2)} \end{pmatrix} R_2(\mathbf{u})^*.$$

Above, $R_1 : V^2(\mathbb{R}^N) \to \mathbf{O}(N)$ is a smooth map such that for any $\mathbf{v} = (v_1, v_2) \in V^2(\mathbb{R}^N)$, we have $R_1(\mathbf{v})\mathbf{e}_i = v_i$ for $i = 1, 2$. Similarly, $R_2 : V^2(\mathbb{R}^{N-2}) \to \mathbf{O}(N-2)$ is a smooth map such that for any $\mathbf{u} = (u_1, u_2) \in V^2(\mathbb{R}^{N-2})$, we have $R_2(\mathbf{u})\mathbf{e}_i = u_i$ for $i = 1, 2$. Lastly, the matrix $\Lambda_{a,b,\theta}$ is given by

$$\Lambda_{a,b,\theta} = \begin{pmatrix} a & b\tan\theta \\ -\frac{b}{\tan\theta} & a \end{pmatrix}.$$

Throughout, we will use the notation $\lambda_j = a_j + ib_j$.

Lemma 3. *The Jacobian of the map Φ is given by*

$$\begin{aligned}
J(\Phi) = &\frac{256 b_1^2 b_2^2 |\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|} |\lambda_1 - \lambda_2|^2 |\lambda_1 - \overline{\lambda}_2|^2 \\
&\times \left|\det\left(M^{(2)} - \lambda_1\right)\right|^2 \left|\det\left(M^{(2)} - \lambda_2\right)\right|^2.
\end{aligned}$$

Proof. See Appendix C. $\square$

Define the following probability measure on $M_{N\times N}(\mathbb{R})$, in which A is a deterministic matrix:

$$\rho(M)dM := \left(\frac{N}{2\pi t}\right)^{\frac{N^2}{2}} \exp\left\{-\frac{N}{2t}\mathrm{Tr}\ [(M-A)^*(M-A)]\right\} dM.$$

By the change-of-variables formula and Lemma 3, we have

$$\begin{aligned}\rho(M)dM &= \left(\frac{N}{2\pi t}\right)^{\frac{N^2}{2}} \frac{256 b_1^2 b_2^2 |\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|}|\lambda_1-\lambda_2|^2|\lambda_1-\overline{\lambda}_2|^2 \\ &\times \left|\det\left(M^{(2)}-\lambda_1\right)\right|^2\left|\det\left(M^{(2)}-\lambda_2\right)\right|^2 \exp\left\{-\frac{N}{2t}\mathrm{Tr}\,(\Phi^{-1}M)^*(\Phi^{-1}M)\right\} \\ &\times da_1 db_1 d\theta_1 d\mathbf{v} dW_1 da_2 db_2 d\theta_2 d\mathbf{u} dW_2 dM^{(2)},\end{aligned}$$

where $\Phi^{-1}M$ is formal notation for the following matrix:

$$\begin{aligned}\Phi^{-1}M &= \begin{pmatrix}\Lambda_{a_1,b_1,\theta} & W_1^* \\ 0 & M^{(1)}\end{pmatrix}, \\ M^{(1)} &= \begin{pmatrix}\Lambda_{a_2,b_2,\theta_2} & W_2^* \\ 0 & M^{(2)}\end{pmatrix}.\end{aligned}$$

We will integrate out all variables except a_1, b_1, a_2, b_2. In this subsection, we focus on integrating out just W_1, W_2, and $M^{(2)}$.

Now, we write A in terms of the basis induced by $R_1(\mathbf{v})$ and $R_2(\mathbf{u})$. Precisely, we write

$$\begin{aligned}A &= R_1(\mathbf{v})\begin{pmatrix}A_{11} & B_1^* \\ C_1 & A^{(1)}\end{pmatrix}R_1(\mathbf{v})^* \\ A^{(1)} &= R_2(\mathbf{u})\begin{pmatrix}A_{11}^{(1)} & B_2^* \\ C_2 & A^{(2)}\end{pmatrix}R_2(\mathbf{u}).\end{aligned}$$

We clarify that A_{11} and $A_{11}^{(1)}$ are blocks of size 2×2. Finally, set $B^{(2)} := M^{(2)} - A^{(2)}$. An elementary computation shows that

$$\begin{aligned}\mathrm{Tr}\,(\Phi^{-1}M)^*(\Phi^{-1}M) &= \left\|\left[I_2\otimes A - \Lambda_{a_1,b_1,\theta_1}\otimes I_N\right]\mathbf{v}\right\|^2 \\ &+ \left\|\left[I_2\otimes A^{(1)} - \Lambda_{a_2,b_2,\theta_2}\otimes I_{N-2}\right]\mathbf{u}\right\|^2 \\ &+ \mathrm{Tr}\ [(W_1-B_1)^*(W_1-B_1)] \\ &+ \mathrm{Tr}\ [(W_2-B_2)^*(W_2-B_2)] \\ &+ \mathrm{Tr}\,(B^{(2)})^*B^{(2)}.\end{aligned}$$

From this, we get

$$\begin{aligned}\rho(M)dM &= \tilde{\rho}(a_1,b_1,\theta_1,\mathbf{v},W_1,a_2,b_2,\theta_2,\mathbf{u},W_2,B^{(2)}) \\ &\times da_1 db_1 d\theta_1 d\mathbf{v} dW_1 da_2 db_2 d\theta_2 d\mathbf{u} dW_2 dB^{(2)},\end{aligned}$$

where $\tilde{\rho}(a_1, b_1, \theta_1, \mathbf{v}, W_1, a_2, b_2, \theta_2, \mathbf{u}, W_2, B^{(2)})$ is given by

$$
\begin{aligned}
&\left(\frac{N}{2\pi t}\right)^{\frac{N^2}{2}} \frac{256 b_1^2 b_2^2 |\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|}|\lambda_1-\lambda_2|^2|\lambda_1-\overline{\lambda}_2|^2\\
&\times \left|\det[A^{(2)}+B^{(2)}-\lambda_1]\right|^2 \left|\det[A^{(2)}+B^{(2)}-\lambda_2]\right|^2 e^{-\frac{N}{2t}\mathrm{Tr}\,(B^{(2)})^* B^{(2)}}\\
&\times \exp\left\{-\frac{N}{2t}\left\|[I_2\otimes A-\Lambda_{a_1,b_1,\theta_1}\otimes I_N]\,\mathbf{v}\right\|^2\right\}\\
&\times \exp\left\{-\frac{N}{2t}\left\|\left[I_2\otimes A^{(1)}-\Lambda_{a_2,b_2,\theta_2}\otimes I_{N-2}\right]\mathbf{u}\right\|^2\right\}\\
&\times \exp\left\{-\frac{N}{2t}\mathrm{Tr}\,[(W_1-B_1)^*(W_1-B_1)]\right\}\\
&\times \exp\left\{-\frac{N}{2t}\mathrm{Tr}\,[(W_2-B_2)^*(W_2-B_2)]\right\}.
\end{aligned}
$$

We note that $dM^{(2)}$ has turned into $dB^{(2)}$; the change-of-variable factor here is 1 because the map $M^{(2)} \mapsto B^{(2)}$ is translation by $A^{(2)}$. Also, we clarify that B and B_1 are functions of all parameters except for W_1, W_2, respectively. Thus, by Gaussian integration, we have

$$
\begin{aligned}
\int_{M_{(N-2)\times 2}(\mathbb{R})} \exp\left\{-\frac{N}{2t}\mathrm{Tr}\,[(W_1-B_1)^*(W_1-B_1)]\right\}dW_1 &= \left(\frac{2\pi t}{N}\right)^{N-2}\\
\int_{M_{(N-4)\times 2}(\mathbb{R})} \exp\left\{-\frac{N}{2t}\mathrm{Tr}\,[(W_2-B_2)^*(W_2-B_2)]\right\}dW_2 &= \left(\frac{2\pi t}{N}\right)^{N-4}.
\end{aligned}
$$

So, by integrating out W_1, W_2, which appear only through the Gaussian weights in the last two lines, we get

$$
\begin{aligned}
&\int_{M_{(N-2)\times 2}(\mathbb{R})}\int_{M_{(N-4)\times 2}(\mathbb{R})} \tilde{\rho}(a_1, b_1, \theta_1, \mathbf{v}, W_1, a_2, b_2, \theta_2, \mathbf{u}, W_2, B^{(2)})dW_1 dW_2\\
&= \left(\frac{N}{2\pi t}\right)^{\frac{N^2-4N+12}{2}} \frac{256 b_1^2 b_2^2 |\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|}|\lambda_1-\lambda_2|^2|\lambda_1-\overline{\lambda}_2|^2\\
&\times \left|\det[A^{(2)}+B^{(2)}-\lambda_1]\right|^2 \left|\det[A^{(2)}+B^{(2)}-\lambda_2]\right|^2 e^{-\frac{N}{2t}\mathrm{Tr}\,(B^{(2)})^* B^{(2)}}\\
&\times \exp\left\{-\frac{N}{2t}\left\|[I_2\otimes A-\Lambda_{a_1,b_1,\theta_1}\otimes I_N]\,\mathbf{v}\right\|^2\right\}\\
&\times \exp\left\{-\frac{N}{2t}\left\|\left[I_2\otimes A^{(1)}-\Lambda_{a_2,b_2,\theta_2}\otimes I_{N-2}\right]\mathbf{u}\right\|^2\right\}.
\end{aligned}
$$

We now integrate the previous expression over $B^{(2)}$. First, let $M_k(\mathbb{R})$ be the space of $k\times k$ matrices with real entries.

Lemma 4. *We have the identity*

$$
\begin{aligned}
&\left(\frac{N}{2\pi t}\right)^{\frac{N^2-4N+12}{2}} \int_{M_{N-4}(\mathbb{R})} \prod_{j=1,2} |\det[A^{(2)}+B^{(2)}-\lambda_j]|^2 e^{-\frac{N}{2t}\mathrm{Tr}\,(B^{(2)})^* B^{(2)}} dB^{(2)}\\
&= 2^6\left(\frac{N}{2\pi t}\right)^{2N+4} \int_{M_4^{skew}(\mathbb{C})} e^{-\frac{N}{2t}\mathrm{Tr}\,X^* X}\,\mathrm{Pf}[\mathbf{M}(X)]dX, \qquad (2.1)
\end{aligned}
$$

where $M_4^{skew}(\mathbb{C})$ is the space of 4×4 skew-symmetric complex matrices, and

$$\mathbf{M}(X) := \begin{pmatrix} X \otimes I_{N-4} & A_{\mathbf{w}}^{(2)} \\ -\left(A_{\mathbf{w}}^{(2)}\right)^T & X^* \otimes I_{N-4} \end{pmatrix}, \tag{2.2}$$

$$A_{\mathbf{w}}^{(2)} := \begin{bmatrix} I_2 \otimes A^{(2)} - \mathbf{w} \otimes I_{N-4} & 0 \\ 0 & \left(I_2 \otimes A^{(2)} - \mathbf{w} \otimes I_{N-4}\right)^* \end{bmatrix} \tag{2.3}$$

$$\mathbf{w} := \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{2.4}$$

Proof. This is very similar to the proof of Lemma 5.1 in [23]. We first use vectors $\chi_1, \chi_2, \psi_1, \psi_2$ of anti-commuting Grassmann variables to write

$$|\det[A^{(2)} + B^{(2)} - \lambda_j]|^2 = \int e^{-\chi_j^*(\lambda_j - A^{(2)} - B^{(2)})\chi_j - \psi_j^*(\overline{\lambda}_j - (A^{(2)})^* - (B^{(2)})^*)\psi_j} d\chi_j d\psi_j.$$

From this, we get

$$\begin{aligned} &\left|\det[A^{(2)} + B^{(2)} - \lambda_1]\right|^2 \left|\det[A^{(2)} + B^{(2)} - \lambda_2]\right|^2 e^{-\frac{N}{2t}\mathrm{Tr}\,(B^{(2)})^* B^{(2)}} \\ &= \int \exp\left\{-\frac{N}{2t}\mathrm{Tr}\,(B^{(2)})^* B^{(2)} - \mathrm{Tr}\, B^{(2)} \sum_{j=1,2} \chi_j \chi_j^* - \mathrm{Tr}\,(B^{(2)})^* \sum_{j=1,2} \psi_j \psi_j^*\right\} \\ &\times \exp\left\{-\sum_{j=1,2} \chi_j^*(\lambda_j - A^{(2)})\chi_j - \sum_{j=1,2} \psi_j^*(\overline{\lambda}_j - (A^{(2)})^*)\psi_j\right\} d\chi_1 d\chi_2 d\psi_1 d\psi_2. \end{aligned}$$

We can integrate the first exponential factor over $B^{(2)}$, since it is a Gaussian integral. We have

$$\begin{aligned} &\left(\frac{N}{2\pi t}\right)^{\frac{(N-4)^2}{2}} \int_{M_{N-4}(\mathbb{R})} e^{-\frac{N}{2t}\mathrm{Tr}\,(B^{(2)})^* B^{(2)} - \mathrm{Tr}\, B^{(2)} \sum_{j=1,2} \chi_j \chi_j^* - \mathrm{Tr}\,(B^{(2)})^* \sum_{j=1,2} \psi_j \psi_j^*} dB^{(2)} \\ &= \exp\left\{\frac{t}{2N} \sum_{k,m=1}^{N-4} \left(\sum_{j=1,2} \left(\chi_{j,k}\bar{\chi}_{j,m} + \psi_{j,m}\bar{\psi}_{j,k}\right)\right)^2\right\} \\ &= \exp\left\{-\frac{t}{N} \sum_{j,\ell=1,2} (\chi_j^* \psi_\ell)(\psi_\ell^* \chi_j) - \frac{t}{N}(\chi_1^T \chi_2)(\chi_2^* \bar{\chi}_1) - \frac{t}{N}(\psi_1^T \psi_2)(\psi_2^* \bar{\psi}_1)\right\}, \end{aligned}$$

where the last identity follows by anti-commuting of Grassmann variables. By Gaussian integration, we can again write

$$\exp\left\{-\frac{t}{N} \sum_{j,\ell=1,2} (\chi_j^* \psi_\ell)(\psi_\ell^* \chi_j)\right\} \tag{2.5}$$

$$= \left(\frac{N}{\pi t}\right)^4 \int_{M_2(\mathbb{C})} e^{-\frac{N}{t}\mathrm{Tr}\, Y^* Y - i\sum_{j,\ell=1,2}(Y_{j\ell}\chi_j^*\psi_\ell + \overline{Y}_{j\ell}\psi_\ell^*\chi_j)} dY \tag{2.6}$$

$$= 2^4 \left(\frac{N}{2\pi t}\right)^4 \int_{M_2(\mathbb{C})} e^{-\frac{N}{t}\mathrm{Tr}\, Y^* Y - i\sum_{j,\ell=1,2}(Y_{j\ell}\chi_j^*\psi_\ell + \overline{Y}_{j\ell}\psi_\ell^*\chi_j)} dY. \tag{2.7}$$

and, similarly,

$$\frac{t}{N}(\chi_1^T\chi_2)(\chi_2^*\bar{\chi}_1) = 2\frac{N}{2\pi t}\int_{\mathbb{C}} e^{-\frac{N}{t}|S|^2 - i\bar{S}\chi_1^T\chi_2 - iS\chi_2^*\bar{\chi}_1}dS \tag{2.8}$$

$$\frac{t}{N}(\psi_1^T\psi_2)(\psi_2^*\bar{\psi}_1) = 2\frac{N}{2\pi t}\int_{\mathbb{C}} e^{-\frac{N}{t}|T|^2 - i\bar{T}\psi_1^T\psi_2 - iT\psi_2^*\bar{\psi}_1}dT. \tag{2.9}$$

Now we change the integration variables from Y, S, T to $X \in M_4^{skew}(\mathbb{C})$ defined by

$$X = \begin{pmatrix} \begin{matrix} 0 & iS \\ -iS & 0 \end{matrix} & iY \\ -iY^T & \begin{matrix} 0 & iT \\ -iT & 0 \end{matrix} \end{pmatrix}.$$

The Jacobian of this change of variables is 1. Since all remaining terms involving Grassmann variables are now quadratic in χ, ψ, we can combine them into an exponent of a quadratic form and integrate

$$\int e^{-\frac{1}{2}\phi^T\mathbf{M}(X)\phi}d\phi = \mathrm{Pf}[\mathbf{M}(X)],$$

where $\phi^T = (\chi_1^*, \chi_2^*, \psi_1^T, \psi_2^T, \chi_1^T, \chi_2^T, -\psi_1^*, -\psi_2^*)$ and $\mathbf{M}(X)$ is defined in the statement of the lemma. It remains to collect the Gaussian weights

$$\exp\left\{-\frac{N}{t}\left(\mathrm{Tr}\, Y^*Y + |S|^2 + |T|^2\right)\right\} = \exp\left\{-\frac{N}{2t}\mathrm{Tr}\, X^*X\right\}$$

and the proof is finished. □

2.2 Re-parameterizing λ_1 and λ_2

By Lemma 4, we have

$$\begin{aligned}
&\int_{\substack{M_{(N-2)\times 2}(\mathbb{R}) \\ M_{(N-2)\times 2}(\mathbb{R}) \\ M_{(N-2)\times(N-2)}(\mathbb{R})}} \tilde{\rho}(a_1, b_1, \theta_1, \mathbf{v}, W_1, a_2, b_2, \theta_2, \mathbf{u}, W_2, B^{(2)})dW_1 dW_2 dB^{(2)} \\
&= 2^6\left(\frac{N}{2\pi t}\right)^{2N+4}\frac{256 b_1^2 b_2^2|\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|}|\lambda_1 - \lambda_2|^2|\lambda_1 - \overline{\lambda}_2|^2 \\
&\times \int_{M_4^{skew}(\mathbb{C})} e^{-\frac{N}{2t}\mathrm{Tr}\, X^*X}\,\mathrm{Pf}[\mathbf{M}(X)]dX \\
&\times \exp\left\{-\frac{N}{2t}\left\|[I_2\otimes A - \Lambda_{a_1,b_1,\theta_1}\otimes I_N]\,\mathbf{v}\right\|^2\right\} \\
&\times \exp\left\{-\frac{N}{2t}\left\|\left[I_2\otimes A^{(1)} - \Lambda_{a_2,b_2,\theta_2}\otimes I_{N-2}\right]\mathbf{u}\right\|^2\right\} \\
&=: \tilde{\rho}(\lambda_1, \theta_1, \mathbf{v}, \lambda_2, \theta_2, \mathbf{u}).
\end{aligned}$$

Recall $\lambda_j = a_j + ib_j$. We now re-parameterize λ_1, λ_2 as to be in a neighborhood of a fixed point $z = a + ib \in \mathbb{C}$ of radius $\sqrt{N}$; we emphasize that we always assume $b > 0$. Set $\mathbf{z} = (z_1, z_2)$ and define

$$\rho_t(z, \mathbf{z}; A) := \frac{1}{N^2\sigma_{z,t}^2}\tilde{\rho}\left(z + \frac{z_1}{\sqrt{N\sigma_{z,t}}}, z + \frac{z_2}{\sqrt{N\sigma_{z,t}}}\right), \tag{2.10}$$

$$\tilde{\rho}(\lambda_1, \lambda_2) := \int_{[0,\frac{\pi}{2}]}\int_{[0,\frac{\pi}{2}]}\int_{V^2(\mathbb{R}^N)}\int_{V^2(\mathbb{R}^{N-2})}\tilde{\rho}(\lambda_1, \theta_1, \mathbf{v}, \lambda_2, \theta_2, \mathbf{u})d\theta_1 d\theta_2 d\mathbf{v}d\mathbf{u}. \tag{2.11}$$

Here $d\mathbf{u}$ and $d\mathbf{v}$ are defined as integration with respect to the rotationally invariant volume form on $V^2(\mathbb{R}^N)$ and $V^2(\mathbb{R}^{N-2})$, respectively. In view of the previous definitions, we will also always identify $\lambda_j = z + N^{-1/2}\sigma_{z,t}^{-1/2} z_j$ for $j = 1, 2$. We now follow [23] and define

$$
\begin{aligned}
\tilde{F}(\mathbf{z}; A^{(2)}) &:= \frac{4N}{\pi^4 t^4 \sigma_{z,t}^3} |z_1 - z_2|^2 |z - \overline{z}|^2 \int_{M_4^{skew}(\mathbb{C})} e^{-\frac{N}{2t}\operatorname{Tr} X^* X} \operatorname{Pf}[\mathbf{M}(X)] dX \\
\tilde{K}_1(z_1) &:= \left(\frac{N}{2\pi t}\right)^N \int_{V^2(\mathbb{R}^N)} \exp\left\{-\frac{N}{2t} \left\| \left[I_2 \otimes A - \Lambda_{a_1,b_1,\theta_1} \otimes I_N\right] \mathbf{v} \right\|^2 \right\} d\mathbf{v} \\
\tilde{K}_2(z_2) &:= \left(\frac{N}{2\pi t}\right)^N \int_{V^2(\mathbb{R}^{N-2})} \exp\left\{-\frac{N}{2t} \left\| \left[I_2 \otimes A^{(1)} - \Lambda_{a_1,b_1,\theta_1} \otimes I_{N-2}\right] \mathbf{u} \right\|^2 \right\} d\mathbf{u} \\
d\nu_1(\mathbf{v}) &:= \tilde{K}_1(z_1)^{-1} \exp\left\{-\frac{N}{2t} \left\| \left[I_2 \otimes A - \Lambda_{a_1,b_1,\theta_1} \otimes I_N\right] \mathbf{v} \right\|^2 \right\} d\mathbf{v}, \\
d\nu_2(\mathbf{u}) &:= \tilde{K}_2(z_2)^{-1} \exp\left\{-\frac{N}{2t} \left\| \left[I_2 \otimes A^{(1)} - \Lambda_{a_1,b_1,\theta_1} \otimes I_{N-2}\right] \mathbf{u} \right\|^2 \right\} d\mathbf{u}.
\end{aligned}
$$

The scaling above is chosen because the $V^2(\mathbb{R}^N)$ and $V^2(\mathbb{R}^{N-2})$ integration resemble Gaussian integration with variance N^{-1}; this is why we want N^N factors with $\tilde{K}_j(z_j)$ terms. Finite powers of N, on the other hand, will not be important to keep track of carefully (since all error terms will be either come from multiplicative factors $1 + o(1)$ or be exponentially small in N).

Again, by identifying $\lambda_j = z + N^{-1/2}\sigma_{z,t}^{-1/2} z_j$, in the formula (2.2) for the matrix $\mathbf{M}(X)$, we can write

$$
\mathbf{w} = \begin{pmatrix} z + \frac{z_1}{\sqrt{N\sigma_{z,t}}} & 0 \\ 0 & z + \frac{z_2}{\sqrt{N\sigma_{z,t}}} \end{pmatrix}.
$$

With this notation, we can write

$$
\begin{aligned}
\rho(z; \mathbf{z}, A) := &\int_{[0,\frac{\pi}{2}]^2} \frac{256 b_1^2 b_2^2 |\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1| \sin^2 2\theta_2|} \tilde{F}(\mathbf{z}; A^{(2)}) \tilde{K}_1(z_1) \tilde{K}_2(z_2) d\nu_1(\mathbf{v}) d\nu_2(\mathbf{u}) d\theta_1 d\theta_2 \\
&\times \left[1 + O(N^{-\frac{1}{2}}))\right].
\end{aligned}
$$

We will estimate $\tilde{F}(\mathbf{z}; A^{(2)})$ following the ideas in [23]. We will then give a preliminary Fourier transform computation for $\tilde{K}_j(z_j)$. Ultimately, the main problem for real non-symmetric matrices is the integration over θ, which we explain further and start dealing with in the next section.

2.3 Estimating $\tilde{F}(\mathbf{z}; A^{(2)})$

The goal of this subsection is to prove the following analog of Lemma 4.1 in [23]. Before we state this lemma, we first introduce the following notation (in which $j \in \{1, 2\}$):

$$Z_j := \begin{pmatrix} 0 & z_j \\ \overline{z}_j & 0 \end{pmatrix}, \tag{2.12}$$

$$\delta_{z,t} = \langle [H_z(\eta_{z,t})(A-z)]^2 \rangle, \tag{2.13}$$

$$\tau_{z,t} = \frac{\gamma_{z,t}\delta_{z,t} - \beta_{z,t}^2}{\gamma_{z,t}\sigma_{z,t}}, \tag{2.14}$$

$$\psi_j := \exp\left\{ -\frac{1}{\sqrt{N\sigma_{z,t}}} \mathrm{Tr}\left[G_z(\eta_{z,t})Z_j\right] - \mathrm{Re}(\overline{\tau}_{z,t} z_j^2) + |z_j|^2 \right\} \tag{2.15}$$

$$V_1 := V_1(\mathbf{v}) := \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & 0 & 0 \\ 0 & 0 & \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix}, \tag{2.16}$$

$$V_2 := V_2(\mathbf{u}) := \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & 0 & 0 \\ 0 & 0 & \mathbf{u}_1 & \mathbf{u}_2 \end{pmatrix} \tag{2.17}$$

We clarify that the dimension of V_1 is $2N \times 4$ and the dimension of V_2 is $2(N-2) \times 4$. Also, for the rest of this paper, when we say "locally uniformly in a_j, b_j", we mean for $a_j, b_j = O(1)$, so that the eigenvalues are $\lambda_j = z + O(N^{-1/2})$ uniformly in N. (We will also use "locally uniformly in z_j" to mean the same thing.) Next, let us introduce the constant

$$\begin{aligned} \upsilon_{z,t} :=& \pi \left[Nt^{-1} - N\langle (A-z)^* \tilde{H}_z(\eta_{z,t}) \tilde{H}_{\overline{z}}(\eta_{z,t})(A-z) \rangle \right] \\ &\times \left[Nt^{-1} - N\langle (A-z) H_z(\eta_{z,t}) H_{\overline{z}}(\eta_{z,t})(A-\overline{z}) \rangle \right]. \end{aligned}$$

Lastly, we emphasize that all estimates hold with high probability with respect to randomness of A, i.e. with probability at least $1 - o(1)$. We only need a finite number of estimates, so the end result holds with high probability (i.e. on an event which contributes $o(1)$ in the expectations in Theorems 1 and 2). To avoid being verbose, we will not always explicitly mention this, e.g. when it concerns local law estimates in Lemmas 26 and 29.

Lemma 5. *We have the following locally uniformly in z_1, z_2, where $\kappa > 0$:*

$$\frac{\tilde{F}(\mathbf{z}; A^{(2)})}{\rho^{(2)}_{\mathrm{GinUE}}(z_1, z_2)} = \frac{32\pi^3 |z - \overline{z}|^2}{Nt^4 \gamma_{z,t}^2 \sigma_{z,t}^3 \upsilon_{z,t}^2} e^{-\frac{2N}{t}\eta_{z,t}^2} \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 \tag{2.18}$$

$$\times \prod_{j=1,2} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^2 \psi_j \left[1 + O(N^{-\kappa})\right]. \tag{2.19}$$

Proof. Before we begin, let us first comment on the strategy. The argument amounts to removing the diagonal 2×2 blocks from generic $X \in M_4^{skew}(\mathbb{C})$ in the integration in $\tilde{F}(\mathbf{z}; A^{(2)})$. Indeed, in view of the proof of Lemma 4, these diagonal blocks correspond to extra terms which differentiate the real case from the complex case as in Section 5 of [23]. After this, we are left with a very similar computation as in Section 5 of [23]. (In fact, bulk correlation functions should agree between real and complex matrix ensembles for eigenvalues away from the real line, so we must arrive at the same Harish-Chandra-Itzykson-Zuber integral as in the complex case of [23].)

For any $X \in M_4^{skew}(\mathbb{C})$, set $P = XX^*$ and $Q = X^*X$. For $\delta > 0$ small but fixed, write

$$\begin{aligned}
&\int_{M_4^{skew}(\mathbb{C})} e^{-\frac{N}{2t}\operatorname{Tr} X^*X} \operatorname{Pf}[\mathbf{M}(X)]dX \\
&= \int_{\|P-\eta_{z,t}^2\|_{\max} \le N^{\delta-\frac{1}{2}}} e^{-\frac{N}{2t}\operatorname{Tr} X^*X} \operatorname{Pf}[\mathbf{M}(X)]dX + \Upsilon \\
&=: \Upsilon_0 + \Upsilon,
\end{aligned}$$

where $\|\|_{\max}$ is the max-entry norm, and Υ is an error term that satisfies the bound

$$\begin{aligned}
|\Upsilon| \le &\sum_{j=1,\dots,4} \int_{|P_{jj}-\eta_{z,t}^2| \ge N^{\delta-\frac{1}{2}}} e^{-\frac{N}{2t}\operatorname{Tr} X^*X} |\det[\mathbf{M}(X)]|^{\frac{1}{2}} dX \\
&+ \sum_{j=1,\dots,4} \int_{|Q_{jj}-\eta_{z,t}^2| \ge N^{\delta-\frac{1}{2}}} e^{-\frac{N}{2t}\operatorname{Tr} X^*X} |\det[\mathbf{M}(X)]|^{\frac{1}{2}} dX \\
&+ \sum_{j,\ell,k=1,\dots,4} \int_{\substack{|P_{j\ell}|,|Q_{j\ell}| \ge N^{\delta-\frac{1}{2}} \\ |P_{kk}-\eta_{z,t}^2| \le N^{\delta-\frac{1}{2}} \\ |Q_{kk}-\eta_{z,t}^2| \le N^{\delta-\frac{1}{2}}}} e^{-\frac{N}{2t}\operatorname{Tr} X^*X} |\det[\mathbf{M}(X)]|^{\frac{1}{2}} dX \\
&=: \Upsilon_1 + \Upsilon_2 + \Upsilon_3.
\end{aligned}$$

(We have used the fact that since X is size $O(1)$, we know $\|P - \eta_{z,t}^2\|_{\max} = O(N^{\delta-1/2})$ is equivalent to $\|Q - \eta_{z,t}^2\|_{\max} = O(N^{\delta-1/2})$. We also used $|\operatorname{Pf}[\mathbf{M}]| \le |\det[\mathbf{M}]|^{1/2}$ for a skew symmetric matrix $\mathbf{M}$.) We first bound Υ. We treat the region $|P_{jj} - \eta_{z,t}^2| \ge N^{\delta-1/2}$ for $j = 1, \dots, 4$. We claim

$$|\det[\mathbf{M}(X)]| = |\det[\mathbf{M}(X)\mathbf{M}(X)^*]|^{\frac{1}{2}} \tag{2.20}$$

$$\le \det[P \otimes I_{N-4} + A_{\mathbf{w}}^{(2)}(A_{\mathbf{w}}^{(2)})^*]^{\frac{1}{2}} \det[Q \otimes I_{N-4} + (A_{\mathbf{w}}^{(2)})^* A_{\mathbf{w}}^{(2)}]^{\frac{1}{2}}. \tag{2.21}$$

The second line follows by computing explicitly the diagonal entries of $\mathbf{M}(X)\mathbf{M}(X)^*$ and then using Lemma 2.2 in [23] (Fisher's inequality) to bound $|\det[\mathbf{M}(X)\mathbf{M}(X)^*]|$ by the product of determinants of its diagonal blocks. If we again compute each matrix in the second line and apply Lemma 2.2 in [23], we obtain

$$|\det[\mathbf{M}(X)]|^{\frac{1}{2}} \le \prod_{j=1,\dots,4} |\det[P_{jj} + |A^{(2)} - \lambda_j|^2]|^{\frac{1}{4}} |\det[Q_{jj} + |A^{(2)} - \lambda_j|^2]|^{\frac{1}{4}},$$

where $\lambda_3 = \lambda_1$ and $\lambda_4 = \lambda_2$. Multiply and divide by $\det[(A-z)^*(A-z) + \eta_{z,t}^2]$. This gives

$$\begin{aligned}
|\det[\mathbf{M}(X)]|^{\frac{1}{2}} \le\ & \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 \\
& \times \prod_{j=1,2} \frac{\det[(A^{(2)} - \lambda_j)^*(A^{(2)} - \lambda_j) + \eta_{z,t}^2]}{\det[(A-z)^*(A-z) + \eta_{z,t}^2]} \\
& \times \prod_{j=1,\dots,4} \left\{ \frac{\det[(A^{(2)} - \lambda_j)^*(A^{(2)} - \lambda_j) + P_{jj}]}{\det[(A^{(2)} - \lambda_j)^*(A^{(2)} - \lambda_j) + \eta_{z,t}^2]} \right\}^{\frac{1}{4}} \\
& \times \prod_{j=1,\dots,4} \left\{ \frac{\det[(A^{(2)} - \lambda_j)^*(A^{(2)} - \lambda_j) + Q_{jj}]}{\det[(A^{(2)} - \lambda_j)^*(A^{(2)} - \lambda_j) + \eta_{z,t}^2]} \right\}^{\frac{1}{4}}.
\end{aligned}$$

Let us control the first product, i.e. the ratio of determinants. We first have

$$\frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta_{z,t}^2]}{\det[(A-z)^*(A-z)+\eta_{z,t}^2]} \tag{2.22}$$

$$= \frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta_{z,t}^2]}{\det[(A-\lambda_j)^*(A-\lambda_j)+\eta_{z,t}^2]}\frac{\det[(A-\lambda_j)^*(A-\lambda_j)+\eta_{z,t}^2]}{\det[(A-z)^*(A-z)+\eta_{z,t}^2]} \tag{2.23}$$

$$= \frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta_{z,t}^2]}{\det[(A-\lambda_j)^*(A-\lambda_j)+\eta_{z,t}^2]}\frac{\det\begin{pmatrix} i\eta_{z,t} & A-\lambda_j \\ (A-\lambda_j)^* & i\eta_{z,t}\end{pmatrix}}{\det\begin{pmatrix} i\eta_{z,t} & A-z \\ (A-z)^* & i\eta_{z,t}\end{pmatrix}} \tag{2.24}$$

$$= \frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta_{z,t}^2]}{\det[(A-\lambda_j)^*(A-\lambda_j)+\eta_{z,t}^2]}\det\left[1-\frac{1}{\sqrt{N\sigma_{z,t}}}G_z(\eta_{z,t})Z_j\right]. \tag{2.25}$$

We now rewrite the last determinant as exponential of trace of log, after which we expand the logarithm to get

$$\begin{aligned}\det\left[1-\frac{1}{\sqrt{N\sigma_{z,t}}}G_z(\eta_{z,t})Z_j\right] &= \exp\left\{\mathrm{Tr}\,\log\left[1-\frac{1}{\sqrt{N\sigma_{z,t}}}G_z(\eta_{z,t})Z_j\right]\right\} \\ &= \exp\left\{-\sum_{k=1}^{\infty}\frac{1}{k(N\sigma_{z,t})^{\frac{k}{2}}}\mathrm{Tr}\,\left[(G_z(\eta_{z,t})Z_j)^k\right]\right\}.\end{aligned}$$

We can bound the $k\geq 3$ contribution by using the trivial operator norm bound $\|G_z(\eta_{z,t})\|_{\mathrm{op}}\lesssim \eta_{z,t}^{-1}\lesssim t^{-1}$. Using also uniform boundedness of $\sigma_{z,t}$ away from 0, this gives (for some $C=O(1)$)

$$\left|\sum_{k=3}^{\infty}\frac{1}{k(N\sigma_{z,t})^{\frac{k}{2}}}\mathrm{Tr}\,\left[(G_z(\eta_{z,t})Z_j)^k\right]\right|\lesssim\sum_{k=3}^{\infty}C^kN^{-\frac{k}{2}+1}t^{-\frac{k}{2}}\lesssim N^{-\frac{1}{2}}\eta_{z,t}^{-\frac{3}{2}}$$

assuming $t=N^{-\epsilon_0}$ with $\epsilon_0>0$ small. Let us now separate the $k=1,2$ terms:

$$\exp\left\{-N^{\frac{1}{2}}\sigma_{z,t}^{-\frac{1}{2}}\langle G_z(\eta_{z,t})Z_j\rangle-\frac{1}{2\sigma_{z,t}}\langle G_z(\eta_{z,t})Z_jG_z(\eta_{z,t})Z_j\rangle\right\}.$$

A straightforward computation using (1.1) and $(A-z)^*\tilde{H}_z(\eta)=H_z(\eta)(A-z)$ shows that

$$\begin{aligned}-\frac{1}{2\sigma_{z,t}}\langle G_z(\eta_{z,t})Z_jG_z(\eta_{z,t})Z_j\rangle &= \frac{1}{\sigma_{z,t}}\eta_{z,t}^2|z_j|^2\langle H_z(\eta_{z,t})\tilde{H}_z(\eta_{z,t})\rangle \\ &\quad -\frac{z_j^2\langle[(A-z)^*H_z(\eta_{z,t})]^2\rangle}{2\sigma_{z,t}} \\ &\quad -\frac{\overline{z}_j^2\langle[H_z(\eta_{z,t})(A-z)]^2\rangle}{2\sigma_{z,t}}.\end{aligned}$$

By definition of $\sigma_{z,t}$, we know that the first term on the RHS is equal to $|z_j|^2-\beta_{z,t}^2\sigma_{z,t}^{-1}\gamma_{z,t}^{-1}|z_j|^2$. Next, note $\langle[H_z(\eta_{z,t})(A-z)]^2\rangle=\overline{\langle[(A-z)^*H_z(\eta_{z,t})]^2\rangle}$ since $H_z(\eta_{z,t})$ is self-adjoint. So, the second and third terms on the RHS of the previous identity combine to $-\mathrm{Re}(\overline{\delta_{z,t}}\sigma_{z,t}^{-1}z_j^2)=-\mathrm{Re}(\overline{\tau_{z,t}}\sigma_{z,t}^{-1}z_j^2)+\beta_{z,t}^2\sigma_{z,t}^{-1}\gamma_{z,t}^{-1}\mathrm{Re}(z_j^2)$. Thus, in total, the previous display is $\leq|z_j|^2-\mathrm{Re}(\overline{\tau_{z,t}}\sigma_{z,t}^{-1}z_j^2)$, and

$$\begin{aligned}&\exp\left\{-N^{\frac{1}{2}}\sigma_{z,t}^{-\frac{1}{2}}\langle G_z(\eta_{z,t})Z_j\rangle-\frac{1}{2\sigma_{z,t}}\langle G_z(\eta_{z,t})Z_jG_z(\eta_{z,t})Z_j\rangle\right\} \\ &\leq\exp\left\{-N^{\frac{1}{2}}\sigma_{z,t}^{-\frac{1}{2}}\langle G_z(\eta_{z,t})Z_j\rangle-\mathrm{Re}(\overline{\tau}_{z,t}\sigma_{z,t}^{-1}z_j^2)+|z_j|^2\right\}=\psi_j.\end{aligned}$$

Thus, we have

$$
\begin{aligned}
&\det\left[1-\frac{1}{\sqrt{N\sigma_{z,t}}}G_z(\eta_{z,t})Z_j\right] \\
&=\psi_j\exp\left\{-\sum_{k=3}^{\infty}\frac{1}{k(N\sigma_{z,t})^{\frac{k}{2}}}\mathrm{Tr}\,\left[(G_z(\eta_{z,t})Z_j)^k\right]\right\} \\
&=\psi_j\left[1+O(N^{-\frac{1}{2}}\eta_{z,t}^{-3})\right].
\end{aligned}
$$

We now compute the remaining ratio of determinants using Cramer's rule to replace $A^{(2)}$ by $A^{(1)}$ and then by A. This is the same procedure as in (5.7) of [23], but now the subspaces that we project resolvents onto must be spanned by pairs $\mathbf{v}\in V^2(\mathbb{R}^N)$. Specifically, this gives

$$
\frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta_{z,t}^2]}{\det[(A-\lambda_j)^*(A-\lambda_j)+\eta_{z,t}^2]}=\prod_{\ell=1,2}\left|\det\left[V_\ell^*G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|, \tag{2.26}
$$

where $G_w^{(\ell)}(\eta)$ is the same as $G_w(\eta)$ but replacing A by $A^{(\ell)}$. We now replace $G_{\lambda_j}^{(\ell-1)}$ by $G_z^{(\ell-1)}$. We have

$$
\left|\det\left[V_\ell^*G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|=\left|\det\left[V_\ell^*G_z^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|\frac{\left|\det\left[V_\ell^*G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|}{\left|\det\left[V_\ell^*G_z^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|}
$$

and, by the resolvent identity, we also have the following (in which $G_w=G_w(\eta_{z,t})$ for convenience):

$$
\begin{aligned}
&\frac{\left|\det\left[V_\ell^*G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|}{\left|\det\left[V_\ell^*G_z^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|} \\
&=\left|\det\left[1+(V_\ell^*G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^*(G_{\lambda_j}^{(\ell)}-G_z^{(\ell-1)})V_\ell^*\right]\right| \\
&=\left|\det\left[1-\frac{1}{\sqrt{N\sigma_{z,t}}}(V_\ell^*G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^*G_z^{(\ell-1)}Z_jG_{\lambda_j}^{(\ell-1)}V_\ell\right]\right|.
\end{aligned} \tag{2.27}
$$

To control this determinant we are left with, we will again write it as exponential of a trace of a log and then expand the log. To this end, we need the following estimate, which uses only Lemma 2.1

of [23] and some elementary manipulations:

$$
\begin{aligned}
&N^{-\frac{k}{2}}\left|\mathrm{Tr}\,\left[(V_\ell^* G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^* G_z^{(\ell-1)}Z_j G_{\lambda_j}^{(\ell-1)}V_\ell\right]^k\right| \\
&= N^{-\frac{k}{2}}\left|\mathrm{Tr}\,\left[G_{\lambda_j}^{(\ell-1)}V_\ell(V_\ell^* G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^* G_z^{(\ell-1)}Z_j\right]^k\right| \\
&\lesssim N^{-\frac{k}{2}}\left\|G_{\lambda_j}^{(\ell-1)}V_\ell(V_\ell^* G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^* G_z^{(\ell-1)}Z_j\right\|_{\mathrm{op}}^k \\
&\lesssim C^k N^{-\frac{k}{2}}\left\|G_{\lambda_j}^{(\ell-1)}V_\ell(V_\ell^* G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^* G_z^{(\ell-1)}\right\|_{\mathrm{op}}^k \\
&\leq C^k N^{-\frac{k}{2}}\left\|G_z^{(\ell-1)}V_\ell(V_\ell^* G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^* G_z^{(\ell-1)}\right\|_{\mathrm{op}}^k \|[G_z^{(\ell-1)}]^{-1}G_{\lambda_j}^{(\ell-1)}\|_{\mathrm{op}}^k \\
&\lesssim C^k N^{-\frac{k}{2}}\left\|[\mathrm{Im}(G_z^{(\ell-1)})^{-1}]^{-1}\right\|_{\mathrm{op}}^k \|[G_z^{(\ell-1)}]^{-1}G_{\lambda_j}^{(\ell-1)}\|_{\mathrm{op}}^k \\
&= C^k N^{-\frac{k}{2}}\left\|[\mathrm{Im}(G_z^{(\ell-1)})^{-1}]^{-1}\right\|_{\mathrm{op}}^k \left\|I+N^{-\frac{1}{2}}\sigma_{z,t}^{-\frac{1}{2}}Z_jG_{\lambda_j}^{(\ell-1)}\right\|_{\mathrm{op}}^k \\
&= C^k N^{-\frac{k}{2}}\eta_{z,t}^{-k}(1+O(N^{-\frac{1}{2}}\eta_{z,t}^{-1})).
\end{aligned}
$$

We clarify that the operator norm bound for the trace follows because the matrix in question rank $O(1)$. Since $\eta_{z,t} \gg N^{-1/2}$, the big-O term in the last line is $\ll 1$. Thus, we have

$$
\left|\det\left[1-\frac{1}{\sqrt{N\sigma_{z,t}}}(V_\ell^* G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^* G_z^{(\ell-1)}Z_jG_{\lambda_j}^{(\ell-1)}V_\ell\right]\right| \tag{2.28}
$$

$$
\leq \exp\left\{\sum_{k=1}^{\infty}\frac{1}{k}N^{-\frac{k}{2}}\sigma_{z,t}^{-\frac{k}{2}}\left|\mathrm{Tr}\,\left[(V_\ell^* G_z^{(\ell-1)}V_\ell)^{-1}V_\ell^* G_z^{(\ell-1)}Z_jG_{\lambda_j}^{(\ell-1)}V_\ell\right]^k\right|\right\} \tag{2.29}
$$

$$
= \exp\left\{\sum_{k=1}^{\infty}O(C^kN^{-\frac{k}{2}}\eta_{z,t}^{-k})\right\} = 1+O(N^{-\frac{1}{2}}\eta_{z,t}^{-1}). \tag{2.30}
$$

Gathering all of this, we deduce

$$
\begin{aligned}
|\det[\mathbf{M}(X)]|^{\frac{1}{2}} &\lesssim \det[(A-z)^*(A-z)+\eta_{z,t}^2]^2\psi_1\psi_2\prod_{\ell=1,2}|\det[V_\ell^* G_z^{(\ell-1)}(\eta_{z,t})V_\ell]|^2 \\
&\times \prod_{j=1,\ldots,4}\left\{\frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+P_{jj}]}{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta_{z,t}^2]}\right\}^{\frac{1}{4}} \\
&\times \prod_{j=1,\ldots,4}\left\{\frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+Q_{jj}]}{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta_{z,t}^2]}\right\}^{\frac{1}{4}}.
\end{aligned}
$$

It remains to control the last two products over j; we treat the first, since the second is handled in the same way. By resolvent perturbation and the inequality $\log(1+x) \leq x - \frac{3x^2}{6+4x}$, we have the following, for which we use the notation $\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t}) := A^{(2),*}_{\lambda_j} A^{(2)}_{\lambda_j} + \eta^2_{z,t}$

$$
\begin{aligned}
&\frac{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+P_{jj}]}{\det[(A^{(2)}-\lambda_j)^*(A^{(2)}-\lambda_j)+\eta^2_{z,t}]} \\
&= \det\left[I_{N-4} + (P_{jj}-\eta^2_{z,t})\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})\right] \\
&= \exp\left\{\mathrm{Tr}\,\log\left[I_{N-4} + (P_{jj}-\eta^2_{z,t})\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})\right]\right\} \\
&\leq \exp\left\{(P_{jj}-\eta^2_{z,t})\mathrm{Tr}\,\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t}) - \frac{3(P_{jj}-\eta^2_{z,t})}{6+4\frac{|P_{jj}-\eta_{z,t}|}{\eta^2_{z,t}}}\mathrm{Tr}\,[\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})]^2\right\}.
\end{aligned}
$$

By combining the previous two displays and by $\frac{t}{N}\mathrm{Tr}\,\tilde{H}_z(\eta_{z,t}) = 1$, we have the following parallel to (5.6) in [23]:

$$
\begin{aligned}
\Upsilon_1 \lesssim\; & e^{-\frac{2N\eta^2_{z,t}}{t}} \det[(A-z)^*(A-z)+\eta^2_{z,t}]^2 \psi_1\psi_2 \prod_{\ell=1,2} |\det[V_\ell^* G_z^{(\ell-1)}(\eta_{z,t})V_\ell]|^2 \\
& \times \sum_{k=1,\dots,4} \int_{|P_{kk}-\eta^2_{z,t}|\leq N^{\delta-\frac{1}{2}}} \exp\left\{-\frac{Nt}{4}\sum_{j=1,\dots,4} h_j(\eta^{-2}_{z,t}|p_j|) - \frac{Nt}{4}\sum_{j=1,\dots,4} h_j(\eta^{-2}_{z,t}|q_j|)\right\} dX,
\end{aligned}
$$

where $p_j = P_{jj} - \eta^2_{z,t}$ and $q_j = Q_{jj} - \eta^2_{z,t}$ and

$$
h_j(x) := \frac{\eta^4_{z,t}}{Nt}\mathrm{Tr}\,[\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})]^2 \frac{3x^2}{6+4x} - \frac{\eta^2_{z,t}}{Nt}\left(\mathrm{Tr}\,[\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})] - \mathrm{Tr}\,[\tilde{H}_z(\eta_{z,t})]\right) x.
$$

By interlacing of eigenvalues and the bound $\|\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})\|_{\mathrm{op}} \leq \eta^{-2}_{z,t}$, we replace $N^{-1}\mathrm{Tr}\,[\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})]^2$ by $N^{-1}\mathrm{Tr}\,[\tilde{H}_{\lambda_j}(\eta_{z,t})]^2 + O(N^{-1}\eta^{-4}_{z,t})$. By the same token, we can also replace $N^{-1}\mathrm{Tr}\,[\tilde{H}^{(2)}_{\lambda_j}(\eta_{z,t})]$ by $N^{-1}\mathrm{Tr}\,[\tilde{H}_{\lambda_j}(\eta_{z,t})] + O(N^{-1}\eta^{-2}_{z,t})$. We can then use a resolvent perturbation to show

$$
\begin{aligned}
\mathrm{Tr}\,[\tilde{H}_{\lambda_j}(\eta_{z,t})] - \mathrm{Tr}\,[\tilde{H}_z(\eta_{z,t})] &= \frac{1}{2N\eta_{z,t}}\mathrm{Tr}\,\mathrm{Im}(G_{\lambda_j}(\eta_{z,t}) - G_z(\eta_{z,t})) \\
&= \frac{1}{2N\eta_{z,t}}\mathrm{Im}\mathrm{Tr}\,\left[N^{-\frac{1}{2}}\sigma_{z,t}^{-\frac{1}{2}} G_{\lambda_j}(\eta_{z,t}) Z_j G_z(\eta_{z,t})\right] \\
&= O(N^{-\frac{1}{2}}\eta^{-3}_{z,t}).
\end{aligned}
$$

Lastly, we have the lower bound $N^{-1}\mathrm{Tr}\,[\tilde{H}_{\lambda j}(\eta_{z,t})]^2 \geq C > 0$, since $\tilde{H}_{\lambda_j}(\eta_{z,t})$ is the resolvent of a bounded operator. So, we obtain the lower bound $h_j(x) \geq C\eta^3_{z,t}(\frac{x^2}{1+x} - \frac{1}{N\eta^3_{z,t}}x) \geq C\eta^3_{z,t}\frac{x^2}{1+x}$ because $\eta = N^{-\epsilon_0}$ with $\epsilon_0 > 0$ small. (The discrepancy in the powers of $\eta_{z,t}$ compared to [23] will not be important; they come from using trivial bounds for resolvents as opposed to local law estimates.) In particular, if we take $\epsilon_0 > 0$ small enough in $\eta_{z,t} = N^{-\epsilon_0}$, then as in [23], the integration over $|P_{kk} - \eta^2_{z,t}| \geq N^{-1/2+\delta}$ is exponentially small, i.e. for some $\kappa > 0$, we have

$$
\Upsilon_1 \lesssim e^{-CN^\kappa} e^{-\frac{2N\eta^2_{z,t}}{t}} \det[(A-z)^*(A-z)+\eta^2_{z,t}]^2\psi_1\psi_2 \prod_{\ell=1,2}|\det[V_\ell^* G_z^{(\ell-1)}(\eta_{z,t})V_\ell]|^2. \tag{2.31}
$$

By reversing the roles of X and X^*, the same argument implies

$$
\Upsilon_2 \lesssim e^{-CN^\kappa} e^{-\frac{2N\eta^2_{z,t}}{t}} \det[(A-z)^*(A-z)+\eta^2_{z,t}]^2\psi_1\psi_2 \prod_{\ell=1,2}|\det[V_\ell^* G_z^{(\ell-1)}(\eta_{z,t})V_\ell]|^2. \tag{2.32}
$$

So, we are left to bound Υ_3; let us treat the case $(j,\ell)=(1,2)$, since the other terms in Υ_3 are treated in the same way. For this, we again start with (2.20). Now, let P_1 denote the upper left 2×2 block of $P=XX^*$. We now use Fisher's inequality (Lemma 2.2 in [23]) again, but now with the top left 2×2 block, the $(3,3)$ entry, and the $(4,4)$ entry as our diagonal blocks. This gives

$$\begin{aligned}&\det[P\otimes I_{N-4}+A^{(2)}_{\mathbf{w}}(A^{(2)}_{\mathbf{w}})^*]\\&\leq|\det[P_1\otimes I_{N-4}+(I_2\otimes A^{(2)}-\mathbf{w}\otimes I_{N-4})(I_2\otimes A^{(2)}-\mathbf{w}\otimes I_{N-4})^*]|\\&\times|\det[P_{33}+(A^{(2)}-\lambda_1)(A^{(2)}-\lambda_1)^*]\det[P_{44}+(A^{(2)}-\lambda_2)(A^{(2)}-\lambda_2)^*].\end{aligned}$$

We now use the Schur complement formula to get the first line below, after which we factor out $P_{22}+|A^{(2)}-\lambda_2|^2$ from the second determinant:

$$\begin{aligned}&\det[P_1\otimes I_{N-4}+(I_2\otimes A^{(2)}-\mathbf{w}\otimes I_{N-4})(I_2\otimes A^{(2)}-\mathbf{w}\otimes I_{N-4})^*]\\&=\det[P_{11}+|A^{(2)}-\lambda_1|^2]\det[P_{22}+|A^{(2)}-\lambda_2|^2-P_{12}(P_{11}+|A^{(2)}-\lambda_1|^2)^{-1}P_{12}]\\&=\det\left[1-|P_{12}|^2H^{(2)}_{\lambda_j}(P_{11})H^{(2)}_{\lambda_j}(P_{22})\right]\det[P_{11}+|A^{(2)}-\lambda_1|^2]\det[P_{22}+|A^{(2)}-\lambda_2|^2].\end{aligned}$$

(Recall $H^{(2)}_w(\eta)=[(A-w)(A-w)^*+\eta^2]^{-1}$.) Now, we use the inequality $\det(I+B)\leq e^{\operatorname{Tr}B}$ and the lower bounds $H^{(2)}_{\lambda_j}(P_{11}),H^{(2)}_{\lambda_j}(P_{22})\geq C>0$ (since they are resolvents of covariance matrices with bounded operator norm) to get

$$\begin{aligned}&\det\left[1-|P_{12}|^2H^{(2)}_{\lambda_j}(P_{11})H^{(2)}_{\lambda_j}(P_{22})\right]\\&\leq\exp\left\{-|P_{12}|^2\operatorname{Tr}H^{(2)}_{\lambda_j}(P_{11})H^{(2)}_{\lambda_j}(P_{22})\right\}\\&=\exp\left\{-|P_{12}|^2\operatorname{Tr}[H^{(2)}_{\lambda_j}(P_{22})]^{\frac12}H^{(2)}_{\lambda_j}(P_{11})[H^{(2)}_{\lambda_j}(P_{22})]^{\frac12}\right\}\\&\leq\exp\{-CN|P_{12}|^2\}.\end{aligned}\tag{2.33}$$

A similar bound holds for Q in place of P. Let us clarify that the effect of using Fisher's inequality with one 2×2 matrix is to produce the determinant factor in (2.33), which produces an a priori exponential bound on the off-diagonal entry of said 2×2 matrix. We shortly apply this principle again. We can now follow our computations for the bounds on Υ_1 and Υ_2; the only difference is the presence of $\exp\{-CN|P_{12}|^2\}$ and $\exp\{-CN|Q_{12}|^2\}$. But we restrict to the set $|P_{12}|,|Q_{12}|\geq N^{\delta-1/2}$. Ultimately, we deduce

$$\Upsilon_3\lesssim e^{-CN^{\kappa}}e^{-\frac{2N\eta_{z,t}^2}{t}}\det[(A-z)^*(A-z)+\eta_{z,t}^2]^2\psi_1\psi_2\prod_{\ell=1,2}|\det[V_\ell^*G_z^{(\ell-1)}(\eta_{z,t})V_\ell]|^2.\tag{2.34}$$

We are left with Υ_0. We must compute $\mathbf{M}(X)\mathbf{M}(X)^*$ more precisely as follows:

$$\begin{aligned}\mathbf{M}(X)\mathbf{M}(X)^*&=\begin{pmatrix}X\otimes I_{N-4}&A^{(2)}_{\mathbf{w}}\\(-A^{(2)}_{\mathbf{w}})^T&X^*\otimes I_{N-4}\end{pmatrix}\begin{pmatrix}X^*\otimes I_{N-4}&-A^{(2)}_{\overline{\mathbf{w}}}\\(A^{(2)}_{\mathbf{w}})^*&X\otimes I_{N-4}\end{pmatrix}\\&=\begin{pmatrix}XX^*\otimes I_{N-4}+A^{(2)}_{\mathbf{w}}[A^{(2)}_{\mathbf{w}}]^*&-X\otimes I_{N-4}A^{(2)}_{\overline{\mathbf{w}}}+A^{(2)}_{\mathbf{w}}X\otimes I_{N-4}\\-[A^{(2)}_{\overline{\mathbf{w}}}]^*X^*\otimes I_{N-4}+X^*\otimes I_{N-4}[A^{(2)}_{\mathbf{w}}]^*&X^*X\otimes I_{N-4}+[A^{(2)}_{\overline{\mathbf{w}}}]^*A^{(2)}_{\overline{\mathbf{w}}}\end{pmatrix}.\end{aligned}\tag{2.35}$$

In particular, we must compute the off-diagonal blocks. Note that the bottom left block is the adjoint of the upper right block (indeed, $\mathbf{M}(X)\mathbf{M}(X)^*$ must be Hermitian). So, we compute the top right block. First, since $X\in M_4^{skew}(\mathbb{C})$, we can write

$$X=\begin{pmatrix}X^1&Y\\-Y^T&X^2\end{pmatrix}=\begin{pmatrix}X^1_{11}&X^1_{12}&Y_{11}&Y_{12}\\-X^1_{12}&X^1_{22}&Y_{21}&Y_{22}\\-Y_{11}&-Y_{21}&X^2_{11}&X^2_{12}\\-Y_{12}&-Y^{22}&-X^2_{12}&X^2_{22}\end{pmatrix}$$

where $X^1, X^2 \in M_2^{skew}(\mathbb{C})$ and $Y \in M_{2\times 2}(\mathbb{C})$. By definition, $A_{\mathbf{w}}^{(2)} X \otimes I_{N-4}$ is equal to

$$\begin{pmatrix} A^{(2)}-\lambda_1 & 0 & 0 & 0 \\ 0 & A^{(2)}-\lambda_2 & 0 & 0 \\ 0 & 0 & A^{(2)}-\overline{\lambda}_1 & 0 \\ 0 & 0 & 0 & A^{(2)}-\overline{\lambda}_2 \end{pmatrix} \begin{pmatrix} X_{11}^1 & X_{12}^1 & Y_{11} & Y_{12} \\ -X_{12}^1 & X_{22}^1 & Y_{21} & Y_{22} \\ -Y_{11} & -Y_{21} & X_{11}^2 & X_{12}^2 \\ -Y_{12} & -Y^{22} & -X_{12}^2 & X_{22}^2 \end{pmatrix}.$$

The top left 2×2 block of this 4×4 matrix is given by

$$\begin{pmatrix} X_{11}^1[A^{(2)}-\lambda_1] & X_{12}^1[A^{(2)}-\lambda_1] \\ -X_{12}^1[A^{(2)}-\lambda_2] & X_{22}^1[A^{(2)}-\lambda_2] \end{pmatrix}.$$

We can compute $X \otimes I_{N-4} A_{\overline{\mathbf{w}}}^{(2)}$ in the exact same way. Its top left 2×2 block is equal to

$$\begin{pmatrix} X_{11}^1[A^{(2)}-\overline{\lambda}_1] & X_{12}^1[A^{(2)}-\overline{\lambda}_2] \\ -X_{12}^1[A^{(2)}-\overline{\lambda}_1] & X_{22}^1[A^{(2)}-\overline{\lambda}_2] \end{pmatrix}.$$

Thus, the top left 2×2 block of $-X\otimes I_{N-4}A_{\overline{\mathbf{w}}}^{(2)} + A_{\mathbf{w}}^{(2)} X \otimes I_{N-4}$ is equal to

$$\begin{pmatrix} X_{11}^1[\overline{\lambda}_1-\lambda_1] & X_{12}^1[\overline{\lambda}_2-\lambda_1] \\ -X_{12}^1[\overline{\lambda}_1-\lambda_2] & X_{22}^1[\overline{\lambda}_2-\lambda_2] \end{pmatrix}. \tag{2.36}$$

The top left 2×2 block of $-[A_{\overline{\mathbf{w}}}^{(2)}]^* X^* \otimes I_{N-4} + X^* \otimes I_{N-4}[A_{\mathbf{w}}^{(2)}]^*$ is just the adjoint of the previous display, as we mentioned earlier. Now, return to (2.35). We apply Fisher's inequality (Lemma 2.2 in [23]) once again with the following choices for blocks. First, we choose the 2×2 block made up of the $(1,1)$ entry of $XX^*\otimes I_{N-4} + A_{\mathbf{w}}^{(2)}[A_{\mathbf{w}}^{(2)}]^*$, the $(2,2)$ entry of $X^*X\otimes I_{N-4} + [A_{\overline{\mathbf{w}}}^{(2)}]^* A_{\overline{\mathbf{w}}}^{(2)}$, the $(1,2)$ entry of (2.36), and the $(2,1)$ entry of its adjoint (in the bottom left block of the second line of (2.35)). Precisely, this block is

$$\begin{pmatrix} P_{11} + (A^{(2)}-\lambda_1)(A^{(2)}-\lambda_1)^* & X_{12}^1[\overline{\lambda}_2-\lambda_1] \\ \overline{X}_{12}[\lambda_2-\overline{\lambda}_1] & Q_{22} + (A^{(2)}-\lambda_1)^*(A^{(2)}-\lambda_1) \end{pmatrix}.$$

This picks out a 2×2 diagonal block in (2.35). Note that its determinant is given by

$$\begin{aligned} &\det \begin{pmatrix} P_{11} + (A^{(2)}-\lambda_1)(A^{(2)}-\lambda_1)^* & X_{12}^1[\overline{\lambda}_2-\lambda_1] \\ \overline{X}_{12}[\lambda_2-\overline{\lambda}_1] & Q_{22} + (A^{(2)}-\lambda_1)^*(A^{(2)}-\lambda_1) \end{pmatrix} \\ &= \det[P_{11} + |A^{(2)}-\lambda_1|^2] \det[Q_{22} + |A^{(2)}-\lambda_2|^2] \det[1 - |X_{12}^1|^2 |\overline{\lambda}_2-\lambda_1|^2 H_{\lambda_j}^{(2)}(\sqrt{P_{11}}) \tilde{H}_{\lambda_j}^{(2)}(\sqrt{Q_{22}})]. \end{aligned}$$

For applying Fisher's inequality, we now take the 1×1 blocks given by the remaining diagonal entries of the second line in (2.35). We get the following (in which $|L|^2 = LL^*$):

$$\begin{aligned} \det[\mathbf{M}(X)\mathbf{M}(X)^*] \leq \prod_{j=1,\ldots,4} &|\det[P_{jj} + |A^{(2)}-\lambda_j|^2]||\det[Q_{jj} + |A^{(2)}-\lambda_j|^2]| \\ &\times \det[1 - |X_{12}^1|^2 |\overline{\lambda}_2-\lambda_1|^2 H_{\lambda_j}^{(2)}(\sqrt{P_{11}}) \tilde{H}_{\lambda_j}^{(2)}(\sqrt{Q_{22}})]. \end{aligned}$$

Now, because $P_{11}, Q_{11} = O(1)$ by our cutoff in the integration domain of Υ_0, as in (2.33), we know that the second line of the previous display is $\leq \exp[-CN|X_{12}^1|^2|\overline{\lambda}_2-\lambda_1|^2]$; this follows by the same argument as in (2.33). But $\lambda_j = z + O(N^{-1/2})$ locally uniformly in a_j, b_j, and $z \notin \mathbb{R}$ is fixed, so $|\overline{\lambda}_2-\lambda_1|^2 \gtrsim 1$ locally uniformly in a_j, b_j, and thus the second line of the previous display is $\leq \exp[-CN|X_{12}^1|^2]$ locally uniformly in a_j, b_j. We can now proceed as in our bound on Υ_3 to restrict further the integration domain in Υ_0 to the set where $|X_{12}| \lesssim N^{-1/2+\delta}$ for any small $\delta > 0$. We can do the same for every other entry of X^1 and every entry of X^2 as well. Ultimately, we have

$$\Upsilon_0 = \Upsilon_{\text{main}} + \Upsilon_4,$$

where for some $\kappa > 0$ small but fixed, we have

$$\Upsilon_{\text{main}} := \int_{\substack{\|P-\eta_{z,t}^2\|_{\max} \le N^{\delta-\frac{1}{2}} \\ \|X^1\|_{\max}, \|X^2\|_{\max} \le N^{\delta-\frac{1}{2}}}} e^{-\frac{N}{2t}\operatorname{Tr} X^* X} \operatorname{Pf}[\mathbf{M}(X)] dX, \tag{2.37}$$
$$|\Upsilon_4| \lesssim e^{-CN^\kappa} e^{-\frac{2N\eta_{z,t}^2}{t}} \det[(A-z)^*(A-z)+\eta_{z,t}^2]^2 \psi_1\psi_2 \prod_{\ell=1,2} |\det[V_\ell^* G_z^{(\ell-1)}(\eta_{z,t}) V_\ell]|^2.$$

In view of the cutoff of Υ_{main}, we now decompose $\mathbf{M}(X)$ as

$$\mathbf{M}(X) = \tilde{\mathbf{M}}(Y) + \tilde{\mathbf{M}}(X^1, X^2),$$

where $\tilde{\mathbf{M}}(X^1, X^2)$ is a diagonal 4×4 block matrix with diagonal entries $X^1 \otimes I_{N-4}, X^2 \otimes I_{N-4}, X^{1,*} \otimes I_{N-4}, X^{2,*} \otimes I_{N-4}$, and where $\tilde{\mathbf{M}}(Y)$ is the following block matrix:

$$\begin{pmatrix} 0 & Y \otimes I_{N-4} & I_2 \otimes A^{(2)} - \mathbf{w} \otimes I_{N-4} & 0 \\ -Y^T \otimes I_{N-4} & 0 & 0 & I_2 \otimes A^{(2)} - \overline{\mathbf{w}} \otimes I_{N-4} \\ -I_2 \otimes A^{(2),*} + \mathbf{w} \otimes I_{N-4} & 0 & 0 & -\overline{Y} \otimes I_{N-4} \\ 0 & -I_2 \otimes A^{(2),*} + \overline{\mathbf{w}} \otimes I_{N-4} & Y^* \otimes I_{N-4} & 0 \end{pmatrix}$$

Because we take Pfaffian, We can change basis by sending $\mathbf{e}_1 \mapsto \mathbf{e}_1, \mathbf{e}_3 \mapsto \mathbf{e}_2, \mathbf{e}_4 \mapsto \mathbf{e}_3, \mathbf{e}_2 \mapsto \mathbf{e}_4$. In this new basis, we then have

$$|\operatorname{Pf}[\mathbf{M}(X)]| = |\operatorname{Pf}[\mathbf{M}(Y) + \mathbf{M}(X^1, X^2)]| = |\det[\mathbf{M}(Y)]|^{\frac{1}{2}} |\det[1 + \mathbf{M}(Y)^{-1}\mathbf{M}(X^1, X^2)]|^{\frac{1}{2}}, \tag{2.38}$$

where

$$\mathbf{M}(X^1, X^2) = \begin{pmatrix} X^1 \otimes I_{N-4} & 0 & 0 & 0 \\ 0 & X^{2,*} \otimes I_{N-4} & 0 & \\ 0 & 0 & X^2 \otimes I_{N-4} & 0 \\ 0 & 0 & 0 & X^{1,*} \otimes I_{N-4} \end{pmatrix},$$
$$\mathbf{M}(Y) = \begin{pmatrix} 0 & M_0(Y) \\ -M_0(Y)^T & 0 \end{pmatrix},$$
$$M_0(Y) = \begin{pmatrix} Y \otimes I_{N-4} & I_2 \otimes A^{(2)} - \mathbf{w} \otimes I_{N-4} \\ -I_2 \otimes A^{(2),*} + \overline{\mathbf{w}} \otimes I_{N-4} & Y^* \otimes I_{N-4} \end{pmatrix}.$$

We now expand

$$\begin{aligned} &|\det[1 + \mathbf{M}(Y)^{-1}\mathbf{M}(X^1, X^2)]|^{\frac{1}{2}} \\ &= \exp\left\{\frac{1}{2}\operatorname{Tr} \mathbf{M}(Y)^{-1}\mathbf{M}(X^1, X^2) - \frac{1}{4}\operatorname{Tr}[\mathbf{M}(Y)^{-1}\mathbf{M}(X^1, X^2)]^2 + O\left(N^{-\frac{1}{2}+3\delta+C\epsilon_0}\right)\right\} \\ &= \exp\left\{-\frac{1}{4}\operatorname{Tr}[\mathbf{M}(Y)^{-1}\mathbf{M}(X^1, X^2)]^2 + O\left(N^{-\frac{1}{2}+3\delta+C\epsilon_0}\right)\right\}, \end{aligned} \tag{2.39}$$

where the bound on the third and higher order terms follows by $\|X^1\|_{\max}, \|X^2\|_{\max} \lesssim N^{-1/2+\delta}$ and $\|\mathbf{M}(Y)^{-1}\|_{\text{op}} \lesssim N^{C\epsilon_0}$, which we verify immediately below. The second identity follows because

$\mathbf{M}(Y)^{-1}\mathbf{M}(X^1, X^2)$ has zero diagonal; this can be checked directly. We now collect a few properties of $\mathbf{M}(Y)^{-1}$. First, we define the following modification of $\mathbf{M}(Y)$ obtained by replacing $\mathbf{w}$ by $\mathbf{z} = zI_2$:

$$\mathbf{M}_z(Y) := \begin{pmatrix} 0 & M_{0,z}(Y) \\ -M_{0,z}(Y)^T & 0 \end{pmatrix},$$
$$M_{0,z}(Y) := \begin{pmatrix} Y \otimes I_{N-4} & I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4} \\ -I_2 \otimes A^{(2),*} + \overline{\mathbf{z}} \otimes I_{N-4} & Y^* \otimes I_{N-4} \end{pmatrix}.$$

An elementary computation shows

$$\mathbf{M}_z(Y)^{-1} = \begin{pmatrix} 0 & -M_{0,z}(Y)^{-1,T} \\ M_{0,z}(Y)^{-1} & 0 \end{pmatrix}$$
$$M_{0,z}(Y)^{-1} = \begin{pmatrix} Y^* \otimes I_{N-4} H^{(2)}_{\mathbf{z}}(Y) & -[I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4}]\tilde{H}^{(2)}_{\mathbf{z}}(Y) \\ [I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4}]^* H^{(2)}_{\mathbf{z}}(Y) & Y \otimes I_{N-4}\tilde{H}^{(2)}_{\mathbf{z}}(Y) \end{pmatrix},$$
$$H^{(2)}_{\mathbf{z}}(Y) = [YY^* \otimes I_{N-4} + (I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})(I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})^*]^{-1}$$
$$\tilde{H}^{(2)}_{\mathbf{z}}(Y) = [Y^*Y \otimes I_{N-4} + (I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})^*(I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})]^{-1}.$$

Because we restrict to $\|XX^* - \eta_{z,t}^2\|_{\max}, \|X^*X - \eta_{z,t}^2\|_{\max}, \|X^1\|_{\max}, \|X^2\|_{\max} \lesssim N^{-1/2+\delta}$ for $\delta > 0$ small, and because $\eta_{z,t}^2 \gtrsim N^{-2\epsilon_0}$ with ϵ_0 small, we know that $\|YY^* - \eta_{z,t}^2\|_{\max}, \|Y^*Y - \eta_{z,t}^2\|_{\max} \lesssim N^{-1/2+\delta}$ and thus $\tilde{H}^{(2)}_{\mathbf{z}}(Y), H^{(2)}_{\mathbf{z}}(Y) \lesssim \eta_{z,t}^{-2} \lesssim N^{2\epsilon_0}$. This implies $\|\mathbf{M}(Y)^{-1}\|_{\mathrm{op}} \lesssim N^{C\epsilon_0}$ for some $C = O(1)$ with high probability (note that $A^{(2)}$ is bounded in operator norm with high probability; indeed, use interlacing to remove the superscript (2) and by the local law in Lemma 26, for example). On the other hand, we know $\|\mathbf{M}_z(Y) - \mathbf{M}(Y)\|_{\mathrm{op}} \lesssim N^{-1/2}$ locally uniformly in z_1, z_2. Thus, resolvent perturbation shows $\mathbf{M}(Y)^{-1} = \mathbf{M}_z(Y)^{-1}[1 + O(N^{-1/2+C\epsilon_0})]$, where the big-O is in operator norm. Finally, we define the matrices below obtained by evaluating resolvents at $\eta_{z,t}$ instead of Y:

$$\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1} = \begin{pmatrix} 0 & -M_{0,z,\eta_{z,t}}(Y)^{-1,T} \\ M_{0,z,\eta_{z,t}}(Y)^{-1} & 0 \end{pmatrix}$$
$$M_{0,z,\eta_{z,t}}(Y)^{-1} = \begin{pmatrix} Y^* \otimes I_{N-4} H^{(2)}_{\mathbf{z}}(\eta_{z,t}) & -[I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4}]\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t}) \\ [I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4}]^* H^{(2)}_{\mathbf{z}}(\eta_{z,t}) & Y \otimes I_{N-4}\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t}) \end{pmatrix},$$
$$H^{(2)}_{\mathbf{z}}(\eta_{z,t}) = [\eta_{z,t}^2 + (I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})(I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})^*]^{-1}$$
$$\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t}) = [\eta_{z,t}^2 + (I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})^*(I_2 \otimes A^{(2)} - \mathbf{z} \otimes I_{N-4})]^{-1}.$$

Before we proceed, we clarify that $\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})$ is block diagonal with entries $\tilde{H}^{(2)}_{z}(\eta_{z,t})$ and $\tilde{H}^{(2)}_{\overline{z}}(\eta_{z,t})$; a similar statement holds for $H^{(2)}_{\mathbf{z}}(\eta_{z,t})$. Thus, its products with $I_2 \otimes A^2 - \mathbf{z} \otimes I_{N-4}$ can be analyzed via the local laws in Lemmas 26 and 29. We will use this shortly.

The bounds $\|YY^* - \eta_{z,t}^2\|_{\max}, \|Y^*Y - \eta_{z,t}^2\|_{\max} \lesssim N^{-1/2+\delta}$ and resolvent perturbation as before show $\|\mathbf{M}_z(Y) - \mathbf{M}_{z,\eta_{z,t}}(Y)\|_{\mathrm{op}} \lesssim N^{-1/2+\delta}$ and $\|\mathbf{M}_z(Y)^{-1} - \mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\|_{\mathrm{op}} \lesssim N^{-1/2+\delta+C\epsilon_0}$. Combining this with $\mathbf{M}(Y)^{-1} = \mathbf{M}_z(Y)^{-1}[1+O(N^{-1/2+C\epsilon_0})]$, we deduce $\|\mathbf{M}(Y)^{-1} - \mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\|_{\mathrm{op}} \lesssim N^{-1/2+\delta+C\epsilon_0}$. Also, by the a priori bounds $\|X_1\|_{\max}, \|X_2\|_{\max} \lesssim N^{-1/2+\delta}$, we also have the estimate

$\|\mathbf{M}(X^1,X^2)\|_{\mathrm{op}} \lesssim N^{-1/2+\delta}$. Thus, we have

$$\begin{aligned}
&\exp\left\{-\frac{1}{4}\left(\operatorname{Tr}[\mathbf{M}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2 - \operatorname{Tr}[\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2\right)\right\}\\
&= \exp\left\{-\frac{1}{4}\operatorname{Tr}\mathbf{M}(Y)^{-1}\mathbf{M}(X^1,X^2)[\mathbf{M}(Y)^{-1}-\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}]\mathbf{M}(X^1,X^2)\right\}\\
&\times \exp\left\{-\frac{1}{4}\operatorname{Tr}[\mathbf{M}(Y)^{-1}-\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}]\mathbf{M}(X^1,X^2)\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\mathbf{M}(X^1,X^2)\right\}\\
&= \exp\left\{O(N^{-\frac{1}{2}+2\delta+C\epsilon_0})\right\},
\end{aligned}$$

where $\delta,\epsilon_0>0$ are small. In particular, we deduce

$$\begin{aligned}
&\exp\left\{-\frac{1}{4}\operatorname{Tr}[\mathbf{M}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2\right\}\\
&= \exp\left\{-\frac{1}{4}\operatorname{Tr}[\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2 + O\left(N^{-\frac{1}{2}+2\delta+C\epsilon_0}\right)\right\}.
\end{aligned} \tag{2.40}$$

An elementary computation shows the following, in which $A^{(2)}(\mathbf{z}) := I_2\otimes A^{(2)} - \mathbf{z}\otimes I_{N-4}$, and in which Y,X^1,X^2 are identified with $Y\otimes I_{N-4}, X^1\otimes I_{N-4}, X^2\otimes I_{N-4}$, respectively, for convenience:

$$\begin{aligned}
-\frac{1}{4}\operatorname{Tr}[\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2 &= -\frac{1}{2}\operatorname{Tr}Y^*H^{(2)}_{\mathbf{z}}(\eta_{z,t})X^1H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})\overline{Y}X^2\\
&- \frac{1}{2}\operatorname{Tr}A^{(2)}(\mathbf{z})\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})X^{2,*}\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}})^TX^2\\
&- \frac{1}{2}\operatorname{Tr}A^{(2)}(\mathbf{z})^*H^{(2)}_{\mathbf{z}}(\eta_{z,t})X^1H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}})X^{1,*}\\
&- \frac{1}{2}\operatorname{Tr}Y\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})X^{2,*}\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})Y^TX^{1,*}.
\end{aligned}$$

Now, observe that $H^{(2)}_{\mathbf{z}}(\eta_{z,t}), \tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t}), H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t}), \tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t}), A^{(2)}(\mathbf{z}), A^{(2)}(\overline{\mathbf{z}}), A^{(2)}(\mathbf{z})^*$ are all of the form $I_2\otimes K$ for some $K\in M_{N-4}(\mathbb{C})$. Thus, we can factorize the traces on the RHS to get

$$\begin{aligned}
-\frac{1}{4}\operatorname{Tr}[\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2 &= -\frac{1}{2}\operatorname{Tr}H^{(2)}_{\mathbf{z}}(\eta_{z,t})H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})\operatorname{Tr}Y^*X^1\overline{Y}X^2\\
&- \frac{1}{2}\operatorname{Tr}A^{(2)}(\mathbf{z})\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}})\operatorname{Tr}X^{2,*}X^2\\
&- \frac{1}{2}\operatorname{Tr}A^{(2)}(\mathbf{z})^*H^{(2)}_{\mathbf{z}}(\eta_{z,t})H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}})\operatorname{Tr}X^1X^{1,*}\\
&- \frac{1}{2}\operatorname{Tr}\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})\operatorname{Tr}YX^{2,*}Y^TX^{1,*}.
\end{aligned}$$

Next, we compute

$$\exp\left\{-\frac{N}{2t}\operatorname{Tr}XX^*\right\} = \exp\left\{-\frac{N}{2t}\operatorname{Tr}X^1X^{1,*}\right\}\exp\left\{-\frac{N}{2t}\operatorname{Tr}X^2X^{2,*}\right\}\exp\left\{-\frac{N}{t}\operatorname{Tr}YY^*\right\}.$$

Therefore, we have

$$
\begin{aligned}
&\exp\left\{-\frac{1}{4}\mathrm{Tr}\,[\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2\right\}\exp\left\{-\frac{N}{2t}\mathrm{Tr}\,XX^*\right\}\\
&=\exp\left\{-\left(\frac{N}{t}+\mathrm{Tr}\,A^{(2)}(\mathbf{z})\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}})\right)\frac{1}{2}\mathrm{Tr}\,X^{2,*}X^2\right\}\\
&\times\exp\left\{-\left(\frac{N}{t}+\mathrm{Tr}\,A^{(2)}(\mathbf{z})^*H^{(2)}_{\mathbf{z}}(\eta_{z,t})H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}})\right)\frac{1}{2}\mathrm{Tr}\,X^{1,*}X^1\right\}\\
&\times\exp\left\{-\frac{1}{2}\mathrm{Tr}\,H^{(2)}_{\mathbf{z}}(\eta_{z,t})H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})\mathrm{Tr}\,Y^*X^1\overline{Y}X^2-\frac{1}{2}\mathrm{Tr}\,\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})\mathrm{Tr}\,YX^{2,*}Y^TX^{1,*}\right\}\\
&\times\exp\left\{-\frac{N}{t}\mathrm{Tr}\,YY^*\right\}.
\end{aligned}
$$

Because X^1,X^2 are 2×2 skew-symmetric, we know that $\mathrm{Tr}\,X^{2,*}X^2=2|x_2|^2$ and $\mathrm{Tr}\,X^{1,*}X^1=2|x_1|^2$ for $x_1,x_2\in\mathbb{C}$. Moreover, by the a priori estimate $\|Y\|_{\mathrm{op}}\lesssim\eta_{z,t}\lesssim t$ that we restrict to in Υ_{main}, we also know that $\mathrm{Tr}\,Y^*X^1\overline{Y}X^2=c_1x_1x_2$ and $\mathrm{Tr}\,YX^{2,*}Y^TX^{1,*}=c_2\overline{x}_1\overline{x}_2$ for constants $c_1,c_2=O(\eta_{z,t}^2)=t^2$. Next, we use Lemma 29 to bound from above both $\mathrm{Tr}\,\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})$ and $\mathrm{Tr}\,H^{(2)}_{\mathbf{z}}(\eta_{z,t})H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})$ by $O(N\eta_{z,t}^{-2})$ with high probability. In particular, we deduce

$$
\begin{aligned}
&\exp\left\{-\frac{1}{2}\mathrm{Tr}\,H^{(2)}_{\mathbf{z}}(\eta_{z,t})H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})\mathrm{Tr}\,Y^*X^1\overline{Y}X^2-\frac{1}{2}\mathrm{Tr}\,\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})\mathrm{Tr}\,YX^{2,*}Y^TX^{1,*}\right\}\\
&=\exp\left\{c_3x_1x_2+c_4\overline{x}_1\overline{x}_2\right\},
\end{aligned}
$$

where $c_3,c_4=O(N)$. By combining the previous two displays, we obtain

$$
\begin{aligned}
&\exp\left\{-\frac{1}{4}\mathrm{Tr}\,[\mathbf{M}_{z,\eta_{z,t}}(Y)^{-1}\mathbf{M}(X^1,X^2)]^2\right\}\exp\left\{-\frac{N}{2t}\mathrm{Tr}\,XX^*\right\}\\
&=\exp\left\{-\begin{pmatrix}x_1\\x_2\end{pmatrix}^*\mathbf{T}_{N,t,\mathbf{z}}(\eta_{z,t})\begin{pmatrix}x_1\\x_2\end{pmatrix}\right\}
\end{aligned}\tag{2.41}
$$

where $\mathbf{T}_{N,t,\mathbf{z}}(\eta_{z,t})$ is the 2×2 matrix

$$
\begin{aligned}
\mathbf{T}_{N,t,\mathbf{z}}&=\begin{pmatrix}\mathbf{T}_{N,t,\mathbf{z},11}&\mathbf{T}_{N,t,\mathbf{z},12}\\\mathbf{T}_{N,t,\mathbf{z},21}&\mathbf{T}_{N,t,\mathbf{z},22}\end{pmatrix},\\
\mathbf{T}_{N,t,\mathbf{z},11}&=\frac{N}{t}+\mathrm{Tr}\,A^{(2)}(\mathbf{z})^*H^{(2)}_{\mathbf{z}}(\eta_{z,t})H^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}}),\\
\mathbf{T}_{N,t,\mathbf{z},22}&=\frac{N}{t}+\mathrm{Tr}\,A^{(2)}(\mathbf{z})\tilde{H}^{(2)}_{\mathbf{z}}(\eta_{z,t})\tilde{H}^{(2)}_{\overline{\mathbf{z}}}(\eta_{z,t})A^{(2)}(\overline{\mathbf{z}}),\\
|\mathbf{T}_{N,t,\mathbf{z},12}|&+|\mathbf{T}_{N,t,\mathbf{z},21}|\lesssim N.
\end{aligned}
$$

We now combine (2.37), (2.38), (2.39), (2.40), and (2.41) to deduce

$$
\begin{aligned}
\Upsilon_{\mathrm{main}}:=&\int_{\substack{\|P-\eta_{z,t}^2\|_{\max}\leq N^{\delta-\frac{1}{2}}\\ \|X^1\|_{\max},\|X^2\|_{\max}\leq N^{\delta-\frac{1}{2}}}}dY e^{-\frac{N}{t}\mathrm{Tr}\,YY^*}|\det[\mathbf{M}(Y)]|^{\frac{1}{2}}\\
&\times\exp\left\{-\begin{pmatrix}x_1\\x_2\end{pmatrix}^*\mathbf{T}_{N,t,\mathbf{z}}(\eta_{z,t})\begin{pmatrix}x_1\\x_2\end{pmatrix}\right\}dx_1dx_2\left\{1+O\left(N^{-\frac{1}{2}+3\delta+C\epsilon_0}\right)\right\}.
\end{aligned}\tag{2.42}
$$

For this, we note the change of variables dX to $dYdx_1dx_2$ has Jacobian 1. Because of the constraints $\|X^1\|_{\max},\|X^2\|_{\max}\leq N^{\delta-1/2}$ above, we can replace $\|P-\eta_{z,t}^2\|_{\max}\leq N^{\delta-1/2}$ by $\|YY^*-\eta_{z,t}^2\|_{\max}\leq$

$N^{\delta-1/2}$ in (2.42). Next, by Lemma 29, the on-diagonal terms of $\mathbf{T}_{N,t,\mathbf{z}}$ are $\gtrsim Nt^{-1}$. Thus, $\mathbf{T}_{N,t,\mathbf{z}} \gtrsim Nt^{-1}$. Since $t = N^{-\epsilon_0}$, we can remove the constraint $\|X^1\|_{\max}, \|X^2\|_{\max} \leq N^{\delta-1/2}$ from (2.42) by using the Gaussian weight in the second line. In particular, we can perform the Gaussian integration over x_1, x_2 and obtain

$$\Upsilon_{\text{main}} = \int_{\|YY^* - \eta_{z,t}^2\|_{\max} \leq N^{\delta-1/2}} e^{-\frac{N}{t}\operatorname{Tr} YY^*} |\det[\mathbf{M}(Y)]|^{\frac{1}{2}} |\det \mathbf{T}_{N,t,z}|^{-2} dY \left[1 + O(N^{-\kappa})\right]$$

for some $\kappa > 0$. As mentioned earlier, the on-diagonal entries of $\mathbf{T}_{N,t,z}$ are $\gtrsim N^{-1}$, whereas the off-diagonal entries are $O(N)$. Thus, for some $\kappa > 0$, we have $\det \mathbf{T}_{N,t,z} = \mathbf{T}_{N,t,\mathbf{z},11}\mathbf{T}_{N,t,\mathbf{z},22}[1 + O(N^{-\kappa})] = \upsilon_{z,t}^{-2}[1 + O(N^{-\kappa})]$ (recall $\upsilon_{z,t}$ from before the statement of this lemma). Therefore, we have

$$\Upsilon_{\text{main}} = \upsilon_{z,t}^{-2} \int_{\|YY^* - \eta_{z,t}^2\|_{\max} \leq N^{\delta-1/2}} e^{-\frac{N}{t}\operatorname{Tr} YY^*} |\det[\mathbf{M}(Y)]|^{\frac{1}{2}} dY \left[1 + O(N^{-\kappa})\right]. \tag{2.43}$$

Now, observe $|\det[\mathbf{M}(Y)]|^{1/2} = \det M_0(Y)$. At this point, we proceed as in Section 5 of [23] to compute the integral in (2.43). Ultimately, after changing variables $Y \mapsto U_1 S U_2^*$ via singular value decomposition with $U_1, U_2 \in \mathbf{U}(2)$ and $S \geq 0$ positive semi-definite, we deduce the following (which is the $\tilde{F}_0$ formula in Section 5 of [23]):

$$\begin{aligned}
\Upsilon_{\text{main}} &= \upsilon_{z,t}^{-2} \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 e^{-\frac{2N}{t}\eta_{z,t}^2} \\
&\times \prod_{j=1,2} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^2 \exp\left\{-\sqrt{\frac{N}{\sigma_{z,t}}}\langle G_z(\eta_{z,t}) Z_j\rangle\right\} \\
&\times |\mathrm{Vol}(\mathbf{U}(2))|^2 \int_{\mathbf{U}(2)\times\mathbf{U}(2)} d\mu(U_1, U_2) e^{-\frac{1}{2\sigma_{z,t}}\sum_{j,k=1,2}\langle G_z(\eta_{z,t}) Z_{jk} G_z(\eta_{z,t}) Z_{kj}\rangle} \\
&\times \int_{\mathbb{R}^2} \prod_{j=1,2} e^{-2N\gamma_{z,t}(s_j - \eta_{z,t})^2} e^{-i\frac{1}{\sqrt{N\sigma_{z,t}}}(s_j - \eta_{z,t})\operatorname{Tr}[G_z(\eta_{z,t})^2 Z_{jj}]} |s_1 - s_2|^2 ds_1 ds_2 \\
&\times \left[1 + O(N^{-\frac{1}{2}+\kappa})\right],
\end{aligned}$$

where Z has 2×2 blocks Z_{jk} indexed by $j, k = 1, 2$, which are defined as

$$Z_{jk} = \begin{pmatrix} 0 & (U_1^* \mathbf{z} U_1)_{jk} \\ (U_2^* \overline{\mathbf{z}} U_2)_{jk} & 0 \end{pmatrix}.$$

We omit the details because they follow exactly as in the complex case (the entries of A essentially have nothing to do with this argument, as long as $\|A\|_{\text{op}} = O(1)$). Now, we change variables $\sqrt{4N\gamma_{z,t}}(s_j - \eta_{z,t}) \mapsto s_j$. This introduces a Jacobian of $\frac{1}{4N\gamma_{z,t}}$, and the Vandermonde determinant $|s_1 - s_2|^2$ gives another factor of $\frac{1}{4N\gamma_{z,t}}$. Hence, we have

$$\begin{aligned}
\Upsilon_{\text{main}} &= \frac{1}{16N^2\gamma_{z,t}^2\upsilon_{z,t}^2} \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 e^{-\frac{2N}{t}\eta_{z,t}^2} \\
&\times \prod_{j=1,2} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^2 \exp\left\{-\sqrt{\frac{N}{\sigma_{z,t}}}\langle G_z(\eta_{z,t} Z_j\rangle\right\} \\
&\times |\mathrm{Vol}(\mathbf{U}(2))|^2 \int_{\mathbf{U}(2)\times\mathbf{U}(2)} d\mu(U_1, U_2) e^{-\frac{1}{2\sigma_{z,t}}\sum_{j,k=1,2}\langle G_z(\eta_{z,t}) Z_{jk} G_z(\eta_{z,t}) Z_{kj}\rangle} \\
&\times \int_{\mathbb{R}^2} \prod_{j=1,2} e^{-\frac{1}{2}s_j^2} e^{-\frac{i}{2\sqrt{\gamma_{z,t}\sigma_{z,t}}} s_j \langle G_z(\eta_{z,t})^2 Z_{jj}\rangle} \left[1 + O(N^{-\frac{1}{2}+C\epsilon_0+\kappa})\right] |s_1 - s_2|^2 ds_1 ds_2.
\end{aligned}$$

Let us first compute $\langle G_z(\eta_{z,t})Z_{jk}G_z(\eta_{z,t})Z_{kj}\rangle$. To this end, we compute

$$\begin{aligned}
G_z(\eta_{z,t})Z_{jk} &= \begin{pmatrix} i\eta_{z,t}H_z(\eta_{z,t}) & H_z(\eta_{z,t})(A-z) \\ (A-z)^*H_z(\eta_{z,t}) & i\eta_{z,t}\tilde{H}_z(\eta_{z,t}) \end{pmatrix} \begin{pmatrix} 0 & (U_1^*\mathbf{z}U_1)_{jk} \\ (U_2^*\overline{\mathbf{z}}U_2)_{jk} & 0 \end{pmatrix} \\
&= \begin{pmatrix} (U_2^*\overline{\mathbf{z}}U_2)_{jk}H_z(\eta_{z,t})(A-z) & i\eta_{z,t}(U_1^*\mathbf{z}U_1)_{jk}H_z(\eta_{z,t}) \\ i\eta_{z,t}(U_2^*\overline{\mathbf{z}}U_2)_{jk}\tilde{H}_z(\eta_{z,t}) & (U_1^*\mathbf{z}U_1)_{jk}(A-z)^*H_z(\eta_{z,t}) \end{pmatrix}.
\end{aligned}$$

The same formula holds for $G_z(\eta_{z,t})Z_{kj}$ if we swap j,k. Hence, we have

$$\begin{aligned}
&\operatorname{Tr}[G_z(\eta_{z,t})Z_{jk}G_z(\eta_{z,t})Z_{kj}] \\
&= -\eta_{z,t}^2(U_1^*\mathbf{z}U_1)_{jk}(U_2^*\overline{\mathbf{z}}U_2)_{kj}\operatorname{Tr}H_z(\eta_{z,t})\tilde{H}_z(\eta_{z,t}) \\
&+ (U_2^*\overline{\mathbf{z}}U_2)_{jk}(U_2^*\overline{\mathbf{z}}U_2)_{kj}\operatorname{Tr}[(H_z(\eta_{z,t})(A-z))^2] \\
&- \eta_{z,t}^2(U_2^*\overline{\mathbf{z}}U_2)_{jk}(U_1^*\mathbf{z}U_1)_{kj}\eta_{z,t}^2\operatorname{Tr}H_z(\eta_{z,t})\tilde{H}_z(\eta_{z,t}) \\
&+ (U_1^*\mathbf{z}U_1)_{jk}(U_1^*\mathbf{z}U_1)_{kj}\operatorname{Tr}[((A-z)^*H_z(\eta_{z,t}))^2] \\
&= -2N\alpha_{z,t}[(U_2^*\overline{\mathbf{z}}U_2)_{jk}(U_1^*\mathbf{z}U_1)_{kj} + (U_1^*\mathbf{z}U_1)_{jk}(U_2^*\overline{\mathbf{z}}U_2)_{kj}] \\
&+ (U_2^*\overline{\mathbf{z}}U_2)_{jk}(U_2^*\overline{\mathbf{z}}U_2)_{kj}\operatorname{Tr}[(H_z(\eta_{z,t})(A-z))^2] \\
&+ (U_1^*\mathbf{z}U_1)_{jk}(U_1^*\mathbf{z}U_1)_{kj}\operatorname{Tr}[((A-z)^*H_z(\eta_{z,t}))^2].
\end{aligned}$$

When we sum over j,k and divide by N to take normalized trace, we get

$$\begin{aligned}
&-\frac{1}{2\sigma_{z,t}}\sum_{j,k=1,2}\langle G_z(\eta_{z,t})Z_{jk}G_z(\eta_{z,t})Z_{jk}\rangle \\
&= \frac{\alpha_{z,t}}{\sigma_{z,t}}\operatorname{Tr}U_2^*\overline{\mathbf{z}}U_2U_1^*\mathbf{z}U_1 - \frac{1}{2\sigma_{z,t}}\langle[H_z(\eta_{z,t})(A-z)]^2\rangle(\overline{z}_1^2+\overline{z}_2^2) \\
&- \frac{1}{2\sigma_{z,t}}\langle[(A-z)^*H_z(\eta_{z,t})]^2\rangle(z_1^2+z_2^2) \\
&= \frac{\alpha_{z,t}}{\sigma_{z,t}}\operatorname{Tr}U_2^*\overline{\mathbf{z}}U_2U_1^*\mathbf{z}U_1 - \sum_{j=1,2}\operatorname{Re}\left(\frac{\overline{\delta}_{z,t}z_j^2}{\sigma_{z,t}}\right).
\end{aligned}$$

Next, we compute $\langle G_z(\eta_{z,t})^2Z_{jj}\rangle$. To this end, we first note that

$$\langle G_z(\eta_{z,t})^2Z_{jj}\rangle = (U_1^*\mathbf{z}U_1)_{jj}\langle[G_z(\eta_{z,t})^2]_{21}\rangle + (U_2^*\overline{\mathbf{z}}U_2)_{jj}\langle[G_z(\eta_{z,t})^2]_{12}\rangle,$$

so only the off-diagonal blocks of $G_z(\eta_{z,t})^2$ matter. These off-diagonal blocks are

$$\begin{aligned}
[G_z(\eta_{z,t})^2]_{12} &= i\eta_{z,t}H_z(\eta_{z,t})^2(A-z) + i\eta_{z,t}H_z(\eta_{z,t})(A-z)\tilde{H}_z(\eta_{z,t}) \\
&= 2i\eta_{z,t}H_z(\eta_{z,t})^2(A-z), \\
[G_z(\eta_{z,t})^2]_{21} &= i\eta_{z,t}(A-z)^*H_z(\eta_{z,t})^2 + i\eta_{z,t}\tilde{H}_z(\eta_{z,t})(A-z)^*H_z(\eta_{z,t}) \\
&= 2i\eta_{z,t}(A-z)^*H_z(\eta_{z,t})^2.
\end{aligned}$$

Note that $\langle[G_z(\eta_{z,t})^2]_{12}\rangle = \overline{\langle[G_z(\eta_{z,t})^2]_{21}\rangle} = 2i\beta_{z,t}$ by invariance of trace under transpose since H_z is self-adjoint. We deduce

$$\langle G_z(\eta_{z,t})^2Z_{jj}\rangle = 2i\beta_{z,t}(U_2^*\overline{\mathbf{z}}U_2)_{jj} + 2i\overline{\beta}_{z,t}(U_1^*\mathbf{z}U_1)_{jj}.$$

We can now rewrite the $\mathbb{R}^2$ integration in Υ_{main} as follows (where $\mathbf{s} = (s_1,s_2)$):

$$\begin{aligned}
&\int_{\mathbb{R}^2}\prod_{j=1,2}e^{-\frac{1}{2}s_j^2}e^{\frac{1}{\sqrt{\gamma_{z,t}\sigma_{z,t}}}s_j[\beta_{z,t}(U_2^*\mathbf{z}U_2)_{jj}+\overline{\beta}_{z,t}(U_1^*\mathbf{z}U_1)_{jj}]}|s_1-s_2|^2ds_1ds_2 \\
&= \int_{\mathbb{R}^2}e^{-\frac{1}{2}\mathbf{s}^2}e^{\frac{1}{\sqrt{\gamma_{z,t}\sigma_{z,t}}}\operatorname{Tr}\mathbf{s}[\beta_{z,t}(U_2^*\overline{\mathbf{z}}U_2)+\overline{\beta}_{z,t}(U_1^*\mathbf{z}U_1)]}|s_1-s_2|^2ds_1ds_2.
\end{aligned}$$

Note that the U_1, U_2 dependence in the integrand of this $\mathbb{R}^2$ integral is invariant under right translation by unitary matrices. Thus, we can replace U_1 by $U_1 V$ and U_2 by $U_2 V$, and average over $V \in \mathbf{U}(2)$ according to Haar measure. Then, with a change-of-variables factor of $\frac{|\mathrm{Vol}(\mathbf{U}(2))|}{8\pi^2}|s_1 - s_2|^2$, we can replace integration on $\mathbb{R}^2 \times \mathbf{U}(2)$ by integration on the space $M_2^{sa}(\mathbb{C})$ of Hermitian 2×2 matrices. Ultimately, we deduce that the integral in the previous display is equal to

$$
\begin{aligned}
&\frac{8\pi^2}{|\mathrm{Vol}(\mathbf{U}(2))|} \int_{M_2^{sa}(\mathbb{C})} e^{-\frac{1}{2}\mathrm{Tr}\, Q^2} e^{\frac{1}{\sqrt{\gamma_{z,t}\sigma_{z,t}}} \mathrm{Tr}\,[\beta_{z,t}(U_2^* \overline{\mathbf{z}} U_2) + \overline{\beta}_{z,t}(U_1^* \mathbf{z} U_1)]Q} dQ \\
&= \frac{16\pi^4}{|\mathrm{Vol}(\mathbf{U}(2))|} \exp\left\{ \frac{|\beta_{z,t}|^2}{\gamma_{z,t}\sigma_{z,t}} \mathrm{Tr}\,[U_2^* \overline{\mathbf{z}} U_2 U_1^* \mathbf{z} U_1] \right\} \\
&\times \exp\left\{ \frac{\beta_{z,t}^2}{2\gamma_{z,t}\sigma_{z,t}} \mathrm{Tr}\,[U_2^* \overline{\mathbf{z}} U_2 U_2^* \overline{\mathbf{z}} U_2] + \frac{\overline{\beta}_{z,t}^2}{2\gamma_{z,t}\sigma_{z,t}} \mathrm{Tr}\,[U_1^* \mathbf{z} U_1 U_1^* \mathbf{z} U_1] \right\} \\
&= \frac{16\pi^4}{|\mathrm{Vol}(\mathbf{U}(2))|} \exp\left\{ \left(\frac{\alpha_{z,t}}{\sigma_{z,t}} + \frac{|\beta_{z,t}|^2}{\gamma_{z,t}\sigma_{z,t}} \right) \mathrm{Tr}\,[U_2^* \overline{\mathbf{z}} U_2 U_1^* \mathbf{z} U_1] + \sum_{j=1,2} \mathrm{Re}\left(\frac{\overline{\beta}_{z,t}^2 z_j^2}{\gamma_{z,t}\sigma_{z,t}} \right) \right\}
\end{aligned}
$$

where the integral is computed by Gaussian integration. Putting it altogether to compute Υ_{main}, we deduce

$$
\begin{aligned}
\Upsilon_{\mathrm{main}} &= \frac{1 + O(N^{-\kappa})}{16 N^2 \gamma_{z,t}^2 \upsilon_{z,t}^2} \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 e^{-\frac{2N}{t}\eta_{z,t}^2} \\
&\times \prod_{j=1,2} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^2 \exp\left\{ -\sqrt{\frac{N}{\sigma_{z,t}}} \langle G_z(\eta_{z,t} Z_j \rangle \right\} \\
&\times 16\pi^4 |\mathrm{Vol}(\mathbf{U}(2))| e^{\frac{\alpha_{z,t}}{\sigma_{z,t}} - \sum_{j=1,2} \mathrm{Re}\left(\frac{\overline{\delta}_{z,t} z_j^2}{\sigma_{z,t}} \right)} e^{\sum_{j=1,2} \mathrm{Re}\left(\frac{\overline{\beta}_{z,t}^2 z_j^2}{\gamma_{z,t}\sigma_{z,t}} \right)} \\
&\times \int_{\mathbf{U}(2)\times\mathbf{U}(2)} d\mu(U_1, U_2) \exp\left\{ \left(\frac{\alpha_{z,t}}{\sigma_{z,t}} + \frac{|\beta_{z,t}|^2}{\gamma_{z,t}\sigma_{z,t}} \right) \mathrm{Tr}\,[U_2^* \overline{\mathbf{z}} U_2 U_1^* \mathbf{z} U_1] \right\}.
\end{aligned}
$$

Now, recall that $\alpha_{z,t} + |\beta_{z,t}|^2 \gamma_{z,t}^{-1} = \sigma_{z,t}$, so in the last line, the factor in front of the trace is 1. So, we can combine terms and obtain the following, in which the last two identities are proven by the Harish-Chandra-Itzykson-Zuber formula as in [23], definition of $\rho_{\mathrm{GinUE}}^{(2)}$, and $|\mathrm{Vol}(\mathbf{U}(2))| = (2\pi)^3$:

$$
\begin{aligned}
\Upsilon_{\mathrm{main}} &= \frac{\pi^4[1 + O(N^{-\kappa})]}{N^2 \gamma_{z,t}^2 \upsilon_{z,t}^2} \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 e^{-\frac{2N}{t}\eta_{z,t}^2} \\
&\times \prod_{j=1,2} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^2 \psi_j e^{-z_j^2} \\
&\times |\mathrm{Vol}(\mathbf{U}(2))| \int_{\mathbf{U}(2)\times\mathbf{U}(2)} d\mu(U_1, U_2) \exp\{\mathrm{Tr}\,[U_2^* \overline{\mathbf{z}} U_2 U_1^* \mathbf{z} U_1]\} \\
&= \frac{\pi^4[1 + O(N^{-\kappa})]}{N^2 \gamma_{z,t}^2 \upsilon_{z,t}^2} \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 e^{-\frac{2N}{t}\eta_{z,t}^2} |\mathrm{Vol}(\mathbf{U}(2))| \frac{\det[e^{z_i \overline{z}_j}]}{|\Delta(\mathbf{z})|^2} \\
&\times \prod_{j=1,2} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^2 \psi_j e^{-z_j^2} \\
&= \frac{8\pi^7}{N^2 \gamma_{z,t}^2 \upsilon_{z,t}^2} \det[(A-z)^*(A-z) + \eta_{z,t}^2]^2 e^{-\frac{2N}{t}\eta_{z,t}^2} \frac{\rho_{\mathrm{GinUE}}^{(2)}(z_1, z_2)}{\Delta(\mathbf{z})^2} \\
&\times \prod_{j=1,2} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^2 \psi_j.
\end{aligned}
$$

Now, by using our error estimates for $\Upsilon_1, \Upsilon_2, \Upsilon_3, \Upsilon_4$ and trivial bounds $\gamma_{z,t} \lesssim t$ and $\upsilon_{z,t} \lesssim N^D$ for some $D = O(1)$, we have

$$\Delta(\mathbf{z})^2 \int_{M_{2\times 2}(\mathbb{C})} e^{-\frac{N}{t}\operatorname{Tr} XX^*} \det[iM(X)]dX = \Delta(\mathbf{z})^2 \Upsilon_{\text{main}}(1 + O(N^{-\kappa})) \tag{2.44}$$

locally uniformly in z_1, z_2, at which point we multiply by $\frac{4N|z-\overline{z}|^2}{t^4\pi^4\sigma_{z,t}^3}$ to finish the proof. □

2.4 Fourier transform computation for $\tilde{K}_j(z_j)$

We require the following analog of Lemma 2.3 in [23] for real matrices and $V^2(\mathbb{R}^N)$.

Lemma 6. *Suppose $f \in L^1(M_{2\times N}(\mathbb{R}))$is continuous in a neighborhood of $V^2(\mathbb{R}^N) = \mathbf{O}(N)/\mathbf{O}(N - 2)$. Let $M_2^{sa}(\mathbb{R})$ be the space of symmetric 2×2 matrices, and define*

$$\hat{f}(P) := \sqrt{\frac{1}{2\pi^7}} \int_{M_{2\times N}(\mathbb{R})} e^{-i\operatorname{Tr}[PXX^*]} f(X)dX, \quad P \in M_2^{sa}(\mathbb{R}). \tag{2.45}$$

Then we have

$$\int_{V^2(\mathbb{R}^N)} f(\mathbf{v})d\mathbf{v} = \int_{M_2^{sa}(\mathbb{R})} e^{i\operatorname{Tr} P} \hat{f}(P)dP. \tag{2.46}$$

We recall that $d\mathbf{v}$ is integration with respect to the rotationally invariant volume form on $V^2(\mathbb{R}^N)$.

Proof. We first define

$$I_\epsilon(f) := \frac{1}{2\pi^2\epsilon^2} \int_{M_{2\times N}(\mathbb{R})} e^{-\frac{1}{2\epsilon}\operatorname{Tr}[(XX^*-I_2)^2]} f(X)dX.$$

We use change of variables $X = OP^{\frac{1}{2}}$, where $O \in V^2(\mathbb{R}^N)$ is a matrix of size $2\times N$ and P is a positive definite matrix of size $N \times N$. This change of variables has Jacobian $dX = \frac{1}{4}\det(P)^{\frac{N-3}{2}} dP d\mu_{N,2}(O)$, where $\mu_{N,2}$ is the rotationally invariant volume form on $V^2(\mathbb{R}^N)$; see Proposition 4 in [15]. Thus,

$$\begin{aligned} I_\epsilon(f) &= \frac{1}{2\pi^2\epsilon^2} \int_{P\geq 0} e^{-\frac{1}{2\epsilon}\operatorname{Tr}[(P-I_2)^2]} g(P)dP, \\ g(P) &:= \frac{1}{4}\det(P)^{\frac{N-3}{2}} \int_{V^2(\mathbb{R}^N)} f(OP^{\frac{1}{2}})d\mu_{N,2}(O). \end{aligned}$$

We take the dP integral to be over 2×2 positive definite matrices since all but two eigenvalues of P are 0. After doing so, $(2\pi^2\epsilon^2)^{-1}\exp\{-\frac{1}{2\epsilon}\operatorname{Tr}[(P - I_2)^2]\}$ is an approximation to the delta function at $P = I_2$. Indeed, first, shift $\tilde{P} := P - I_2$; the $\pi^{-2}\epsilon^{-2}$ comes from scaling the matrix entries of $\tilde{P}$ to be Gaussians which integrate to 1. Recall g is continuous at $P = I_2$. Taking $\epsilon \to 0$, we get

$$I_\epsilon(f) \to_{\epsilon\to 0} \frac{1}{4}\det(P)^{\frac{N-3}{2}} \int_{V^2(\mathbb{R}^N)} f(O)d\mu_{N,2}(O),$$

which we can formally rewrite by replacing O by $\mathbf{v}$ and $d\mu_{N,2}(O)$ by $d\mathbf{v}$. We now compute $I_\epsilon(f)$ differently; by a Hubbard-Stratonovich transform, we have

$$\begin{aligned} I_\epsilon(f) &= \frac{1}{2\pi^2\epsilon^2} \int_{M_{2\times N}(\mathbb{R})} \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{M_2^{sa}(\mathbb{R})} e^{-\frac{1}{2}\operatorname{Tr} P^2 + \frac{i}{\sqrt{\epsilon}}\operatorname{Tr}[P(I_2-XX^*)]} f(X)dXdP \\ &= \frac{1}{2^{\frac{5}{2}}\pi^{\frac{7}{2}}} \int_{M_2^{sa}(\mathbb{R})} e^{-\frac{\epsilon}{2}\operatorname{Tr} P^2 + i\operatorname{Tr} P} \int_{M_{2\times N}(\mathbb{R})} e^{-i\operatorname{Tr}[PXX^*]} f(X)dXdP, \end{aligned}$$

at which point we conclude by definition of $\hat{f}(P)$. □

With this in hand, we can now start to compute $\tilde{K}_j(z_j)$. First, we introduce some notation:

$$\begin{aligned}
A^{(j-1)}_{a_j,b_j,\theta_j} &:= I_2 \otimes A^{(j-1)} - \Lambda_{a_j,b_j,\theta_j} \otimes I_{N-2j+2},\\
C^{(j-1)}_{a_j,b_j,\theta_j} &:= (A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j},\\
\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(\eta) &:= [(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j} + \eta^2]^{-1},\\
\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j} &:= \tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(\eta_{z,t}).
\end{aligned}$$

Lemma 7. *We have*

$$\begin{aligned}
\tilde{K}_j(z_j) = &\frac{N^{2j+1}t^{-2j-1}e^{\frac{N}{t}\eta^2_{z,t}}}{2^{2j+\frac{1}{2}}\pi^{2j+\frac{3}{2}}} \det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j} + \eta^2_{z,t}]^{-\frac{1}{2}}\\
&\times \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I_{N-2j+2})[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} dP.
\end{aligned}$$

Proof. By definition, we have

$$\begin{aligned}
\tilde{K}_j(z_j) &= \left(\frac{N}{2\pi t}\right)^N \int_{V^2(\mathbb{R}^{N-2j+2})} \exp\left\{-\frac{N}{2t}\mathbf{v}^* C^{(j-1)}_{a_j,b_j,\theta_j}\mathbf{v}\right\} d\mathbf{v}\\
&= \left(\frac{N}{2\pi t}\right)^N e^{\frac{N\eta^2_{z,t}}{t}} \int_{V^2(\mathbb{R}^{N-2j+2})} \exp\left\{-\frac{N}{2t}\mathbf{v}^*[C^{(j-1)}_{a_j,b_j,\theta_j} + \eta^2_{z,t}]\mathbf{v}\right\} d\mathbf{v}.
\end{aligned}$$

By Lemma 6 and a change-of-variables $P \mapsto \frac{N}{2t}P$ in the dP integral therein, we know

$$\begin{aligned}
\tilde{K}_j(z_j) = &\left(\frac{N}{2\pi t}\right)^N \left(\frac{N}{2t}\right)^3 \sqrt{\frac{1}{2\pi^7}} e^{\frac{N}{t}\eta^2_{z,t}}\\
&\times \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \int_{\mathbb{R}^{2N-4j+4}} e^{-\frac{N}{2t}\mathbf{x}^*[C^{(j-1)}_{a_j,b_j,\theta_j} + \eta^2_{z,t} + iP\otimes I_{N-2j+2}]\mathbf{x}} d\mathbf{x}dP.
\end{aligned}$$

We now perform the $d\mathbf{x}$ integral explicitly since it is a Gaussian. The previous display thus equals

$$\begin{aligned}
&\left(\frac{2\pi t}{N}\right)^{N-2j+2} \left(\frac{N}{2\pi t}\right)^N \left(\frac{N}{2t}\right)^3 \sqrt{\frac{1}{2\pi^7}} e^{\frac{N}{t}\eta^2_{z,t}}\\
&\times \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j} + \eta^2_{z,t} + iP \otimes I_{N-2j+2}]^{-\frac{1}{2}} dP\\
&= 2^{-2j-\frac{3}{2}}\pi^{-2j-\frac{3}{2}} N^{2j+1} t^{-2j-1} e^{\frac{N}{t}\eta^2_{z,t}}\\
&\times \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j} + \eta^2_{z,t} + iP \otimes I_{N-2j+2}]^{-\frac{1}{2}} dP.
\end{aligned}$$

By factoring out $\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}$ from the last determinant, the proof follows. ☐

Combining our computations thus far for $\tilde{F}(\mathbf{z}; A^{(2)})$ and $\tilde{K}_j(z_j)$ yields the following in which

$\kappa > 0$ is independent of N:

$$\frac{\rho_t(z,\mathbf{z};A)}{\rho^{(2)}_{\mathrm{GinUE}}(z_1,z_2)} = \frac{32\pi^3 e^{-\frac{2N}{t}\eta_{z,t}^2}[1+O(N^{-\kappa})]}{Nt^4\gamma_{z,t}^2\sigma_{z,t}^3\upsilon_{z,t}^2}\int_{[0,\frac{\pi}{2}]^2} d\theta_1 d\theta_2 \frac{256 b_1^2 b_2^2 |\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|} \tag{2.47}$$

$$\times \prod_{j=1,2}\int_{V^2(\mathbb{R}^{N-2j+2})} d\mu_j(\mathbf{v}_j)|\det[V_j^* G_z^{(j-1)}(\eta_{z,t})V_j]|^2$$

$$\times \prod_{j=1,2}\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2]\det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j}+\eta_{z,t}^2]^{-\frac{1}{2}}$$

$$\times \prod_{j=1,2}\frac{N^{2j+1}t^{-2j-1}e^{\frac{N}{t}\eta_{z,t}^2}}{2^{2j+\frac{3}{2}}\pi^{2j+\frac{3}{2}}}\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr}P}\det[I+i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}dP$$

$$= \frac{N^7 b^4[1+O(N^{-\kappa})]}{16\pi^4 t^{12}\gamma_{z,t}^2\sigma_{z,t}^3\upsilon_{z,t}^2}\int_{[0,\frac{\pi}{2}]^2} d\theta_1 d\theta_2 \frac{256|\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|}$$

$$\times \prod_{j=1,2}\int_{V^2(\mathbb{R}^{N-2j+2})} d\mu_j(\mathbf{v}_j)|\det[V_j^* G_z^{(j-1)}(\eta_{z,t})V_j]|^2$$

$$\times \prod_{j=1,2}\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2]\det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j}+\eta_{z,t}^2]^{-\frac{1}{2}}$$

$$\times \prod_{j=1,2}\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr}P}\det[I+i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}dP,$$

where the second identity follows by combining factors of $N,t,\pi,2$, viewing V_j as a function of $\mathbf{v}_j$ in the last line above, and noting that $b_1,b_2 = b+O(N^{-1/2})$ (recall $b>0$ is positive).

3 Estimating the integral over θ_1,θ_2

Recall $A^{(j-1)}_{a_j,b_j,\theta_j}$ from before Lemma 7. Let Θ denote the $d\theta_1 d\theta_2$ integration of interest:

$$\Theta := \int_{[0,\frac{\pi}{2}]^2} d\theta_1 d\theta_2 \frac{256|\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|}$$

$$\times \prod_{j=1,2}\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2]\det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j}+\eta_{z,t}^2]^{-\frac{1}{2}}$$

$$\times \prod_{j=1,2}\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr}P}\det[I+i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}dP.$$

Technically, Θ depends on $\mathbf{v}_1,\mathbf{v}_2$ (the integration variables in $V^2(\mathbb{R}^{N-2j+2})$ for $j=1,2$), though this dependence will not play a role in this section. Now, for $\tau>0$ small and fixed (depending only on ϵ_0), define

$$\mathcal{I}_0 = \left\{\theta : |\theta-\pi/4| \le N^{-1/2+\tau}\right\}$$
$$\mathcal{I}_1 = \left[0,\frac{\pi}{2}\right]\setminus\mathcal{I}_0.$$

We write

$$\Theta = \sum_{j,k=0,1}\Theta_{jk},$$

where Θ_{jk} is the same as Θ, but the integration in $d\theta_1 d\theta_2$ is over $\theta_1\in\mathcal{I}_j$ and $\theta_2\in\mathcal{I}_k$. Our goal in this section is the following.

Proposition 8. *There exists $\kappa > 0$ independent of N such that if $(j,k) \neq (0,0)$, then $\Theta_{jk} = O(e^{-N^\kappa})$ locally uniformly in a_j, b_j.*

As an immediate consequence of this result, we have the following estimate.

Corollary 9. *First, define*

$$\rho_{\mathrm{main}}(z,\mathbf{z};A) := \frac{N^7 b^4[1+O(N^{-\kappa})]}{16\pi^4 t^{12}\gamma_{z,t}^2\sigma_{z,t}^3\upsilon_{z,t}^2}\rho_{\mathrm{GinUE}}^{(2)}(z_1,z_2)\int_{\mathcal{I}_0^2} d\theta_1 d\theta_2 \frac{256|\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|}$$
$$\times \prod_{j=1,2}\int_{V^2(\mathbb{R}^{N-2j+2})} d\mu_j(\mathbf{v}_j)|\det[V_j^* G_z^{(j-1)}(\eta_{z,t})V_j]|^2$$
$$\times \prod_{j=1,2}\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2]\det[(A_{a_j,b_j,\theta_j}^{(j-1)})^* A_{a_j,b_j,\theta_j}^{(j-1)}+\eta_{z,t}^2]^{-\frac{1}{2}}$$
$$\times \prod_{j=1,2}\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr}P}\det[I+i[\tilde{H}_{a_j,b_j,\theta_j}^{(j-1)}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}_{a_j,b_j,\theta_j}^{(j-1)}]^{\frac{1}{2}}]^{-\frac{1}{2}}dP.$$

There exists $\kappa > 0$ independent of N such that

$$\rho_t(z,\mathbf{z};A) = \rho_{\mathrm{main}}(z,\mathbf{z};A) + O(e^{-N^\kappa}). \tag{3.1}$$

3.1 Preliminary bounds for integrating on $\mathcal{I}_1$

The crux of this subsection is control integrating $\theta_j \in \mathcal{I}_1$; in this region, we do not have local laws for $A_{a_j,b_j,\theta_j}^{(j-1)}$ and its related resolvents since the operator norm of $A_{a_j,b_j,\theta_j}^{(j-1)}$ blows up as $\theta \to 0, \frac{\pi}{2}$. The first step we take is the following bound on the product of determinants.

Lemma 10. *There exists a constant $C > 0$ such that for any $j \geq 1$, we have the following locally uniformly in a_j, b_j:*

$$\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2]\det[(A_{a_j,b_j,\theta_j}^{(j-1)})^* A_{a_j,b_j,\theta_j}^{(j-1)}+\eta_{z,t}^2]^{-\frac{1}{2}}$$
$$\lesssim N^{C\epsilon_0}\exp\left\{-Cb_j^2 N\eta_{z,t}^2[\tan\theta_j - \tan^{-1}\theta_j]^2\right\}.$$

Proof. As shown in the proof of Lemma 5, we know that

$$\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2] = \det[(A^{(j-1)}-\lambda_j)^*(A^{(j-1)}-\lambda_j)+\eta_{z,t}^2]$$
$$\times \prod_{\ell=1}^{j-1}\left|\det\left[V_\ell^* G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]\right|^{-1}\left[1+O(N^{-\kappa})\right]$$

for some $\kappa > 0$ locally uniformly in a_j, b_j. We now bound the inverse-determinants in the second line. We first claim that the following holds:

$$\left\|\left[V_\ell^* G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]^{-1}\right\|_{\mathrm{op}} = \left\|V_\ell\left[V_\ell^* G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]^{-1}V_\ell^*\right\|_{\mathrm{op}}$$
$$\lesssim \left\|G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\left[V_\ell^* G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell\right]^{-1}V_\ell^* G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})\right\|_{\mathrm{op}}$$
$$\lesssim \|[\mathrm{Im}\mathcal{H}_{\lambda_j}^{(\ell-1)}(\eta_{z,t})]^{-1}\|_{\mathrm{op}} \lesssim \eta_{z,t}^{-1}.$$

The first line follows because V_ℓ is projection onto its image, and $[V_\ell^* G_{\lambda_j}^{(\ell-1)}(\eta_{z,t})V_\ell]^{-1}$is a map from said image to itself. The last line follows by Lemma 2.1 in [23]. Thus, since $\eta_{z,t} = N^{-\epsilon_0}$, it suffices to show that

$$\begin{aligned}&\det[(A^{(j-1)}-\lambda_j)^*(A^{(j-1)}-\lambda_j)+\eta_{z,t}^2]\det[(A^{(j-1)}_{a_j,b_j,\theta_j})^*A^{(j-1)}_{a_j,b_j,\theta_j}+\eta_{z,t}^2]^{-\frac{1}{2}}\\&\lesssim \exp\left\{-Cb_j^2N^2\eta_{z,t}^2[\tan\theta_j-\tan^{-1}\theta_j]^2\right\}.\end{aligned}$$

We claim the following holds:

$$\begin{aligned}&\det[(A^{(j-1)}-\lambda_j)^*(A^{(j-1)}-\lambda_j)+\eta_{z,t}^2]\det[(A^{(j-1)}_{a_j,b_j,\theta_j})^*A^{(j-1)}_{a_j,b_j,\theta_j}+\eta_{z,t}^2]^{-\frac{1}{2}}\\&=\left\{\det\left[I+b_j^2(\tan\theta_j-\tan^{-1}\theta_j)^2\eta_{z,t}^2H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\tilde{H}^{(j-1)}_{\overline{\lambda}_j}(\eta_{z,t})H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\right]\right\}^{-\frac{1}{2}}.\end{aligned}\tag{3.2}$$

Assuming this, the proof follows quickly. Indeed, we know by definition that

$$\begin{aligned}&H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\tilde{H}^{(j-1)}_{\overline{\lambda}_j}(\eta_{z,t})H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\\&=[(A^{(j-1)}-\lambda_j)(A^{(j-1)}-\lambda_j)^*+\eta_{z,t}^2]^{-\frac{1}{2}}\\&\times[(A^{(j-1)}-\overline{\lambda}_j)^*(A^{(j-1)}-\overline{\lambda}_j)+\eta_{z,t}^2]^{-1}\\&\times[(A^{(j-1)}-\lambda_j)(A^{(j-1)}-\lambda_j)^*+\eta_{z,t}^2]^{-\frac{1}{2}}.\end{aligned}$$

The operator norm of each matrix we take inverse of on the RHS is $O(1)$; there is no θ-dependence. So, we know that for some $C>0$ independent of N,θ, we have

$$H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\tilde{H}^{(j-1)}_{\overline{\lambda}_j}(\eta_{z,t})H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\geq C.$$

This gives

$$\begin{aligned}&\left\{\det\left[I+b_j^2(\tan\theta_j-\tan^{-1}\theta_j)^2\eta_{z,t}^2H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\tilde{H}^{(j-1)}_{\overline{\lambda}_j}(\eta_{z,t})H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\right]\right\}^{-\frac{1}{2}}\\&\leq\left\{\det\left[I+Cb_j^2(\tan\theta_j-\tan^{-1}\theta_j)^2\eta_{z,t}^2\right]\right\}^{-\frac{1}{2}}\\&\leq(1+Cb_j^2(\tan\theta_j-\tan^{-1}\theta_j)^2\eta_{z,t}^2)^{-\frac{N}{2}}\lesssim e^{-NC'b_j^2\eta_{z,t}^2(\tan\theta_j-\tan^{-1}\theta_j)^2}.\end{aligned}$$

At this point, the claim follows, so it suffices to prove (3.2). To ease notation, let us focus on $j=1$. We let matrices R and L consist of right and left eigenvectors of $\Lambda_{a_1,b_1,\theta_1}$. More precisely,

$$L=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -i\tan\theta_1\\ 1 & i\tan\theta_1\end{pmatrix},\qquad R=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -i\tan^{-1}\theta_1\\ 1 & i\tan^{-1}\theta_1\end{pmatrix}.$$

Then it is easy to see that

$$\Lambda_{a_1,b_1,\theta_1}=R^*\begin{pmatrix}\lambda_1 & 0\\ 0 & \overline{\lambda}_1\end{pmatrix}L,\qquad \Lambda^*_{a_1,b_1,\frac{\pi}{2}-\theta_1}=R^*\begin{pmatrix}\lambda & 0\\ 0 & \overline{\lambda}_1\end{pmatrix}^*L.$$

Consider a permutation matrix

$$J=\begin{pmatrix}I_N & 0 & 0 & 0\\ 0 & 0 & I_N & 0\\ 0 & I_N & 0 & 0\\ 0 & 0 & 0 & I_N\end{pmatrix}.$$

Now, we introduce the following $4N\times 4N$ matrix:

$$\Gamma_{\lambda_1,\theta_1}(\eta):=\begin{pmatrix}i\eta & A_{a_1,b_1,\theta_1}\\ A_{a_1,b_1,\frac{\pi}{2}-\theta_1} & i\eta\end{pmatrix}^{-1}.$$

It is straightforward to check that

$$\Gamma_{\lambda_1,\theta_1}(\eta) = \begin{pmatrix} R^* \otimes I_N & 0 \\ 0 & R^* \otimes I_N \end{pmatrix} J \begin{pmatrix} G_{\lambda_1}(\eta) & 0 \\ 0 & G_{\overline{\lambda}_1}(\eta) \end{pmatrix} J \begin{pmatrix} L \otimes I_N & 0 \\ 0 & L \otimes I_N \end{pmatrix}.$$

Thus

$$\begin{aligned} \det \Gamma_{\lambda_1,\theta_1}(\eta) = \det G_{\lambda_1}(\eta) \det G_{\overline{\lambda}_1}(\eta) &= \left\{ \det \left[(A-\lambda_1)(A-\lambda_1)^* + \eta^2 \right] \right\}^{-2} \\ &= \left\{ \det \begin{pmatrix} i\eta & A-\lambda_1 \\ A^T - \overline{\lambda}_1 & i\eta \end{pmatrix} \right\}^{-2}. \end{aligned}$$

We plug this in to get

$$\begin{aligned} &\det[(A-\lambda_1)^*(A-\lambda_1) + \eta_{z,t}^2] \det[A^*_{a_1,b_1,\theta_1} A_{a_1,b_1,\theta_1} + \eta_{z,t}^2]^{-\frac{1}{2}} \\ &= \det \begin{pmatrix} i\eta_{z,t} & A-\lambda_1 \\ A^T - \overline{\lambda}_1 & i\eta_{z,t} \end{pmatrix} \left\{ \det \begin{pmatrix} i\eta_{z,t} & A_{a_1,b_1,\theta_1} \\ A^*_{a_1,b_1,\theta_1} & i\eta_{z,t} \end{pmatrix} \right\}^{-\frac{1}{2}} \\ &= \left\{ \det \begin{pmatrix} i\eta_{z,t} & A_{a_1,b_1,\theta_1} \\ A^*_{a,b,\frac{\pi}{2}-\theta_1} & i\eta_{z,t} \end{pmatrix} \right\}^{\frac{1}{2}} \left\{ \det \begin{pmatrix} i\eta_{z,t} & A_{a_1,b_1,\theta_1} \\ A^*_{a_1,b_1,\theta_1} & i\eta_{z,t} \end{pmatrix} \right\}^{-\frac{1}{2}} \\ &= \left\{ \det \left[I_{4N} + \Gamma_{\lambda_1,\theta_1}(\eta_{z,t}) \begin{pmatrix} 0 & 0 \\ \left(\Lambda_{a_1,b_1,\frac{\pi}{2}-\theta_1} - \Lambda_{a_1,b_1,\theta_1}\right)^* \otimes I_N & 0 \end{pmatrix} \right] \right\}^{-\frac{1}{2}} \\ &= \left\{ \det \left[I_{4N} - \begin{pmatrix} G_{\lambda_1}(\eta_{z,t}) & 0 \\ 0 & G_{\overline{\lambda}_1}(\eta_{z,t}) \end{pmatrix} b_1 \left(\tan\theta_1 - \tan^{-1}\theta_1 \right) Q_{\theta_1} \otimes I_N \right] \right\}^{-\frac{1}{2}}, \end{aligned}$$

where

$$\begin{aligned} Q_{\theta_1} &= J \begin{pmatrix} L & 0 \\ 0 & L \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} R^* & 0 \\ 0 & R^* \end{pmatrix} J \\ &= \frac{i}{2} \begin{pmatrix} -(\tan\theta_1 - \tan^{-1}\theta_1) & -(\tan\theta_1 + \tan^{-1}\theta_1) \\ \tan\theta_1 + \tan^{-1}\theta_1 & \tan\theta_1 - \tan^{-1}\theta_1 \end{pmatrix} \otimes E_{(2)}, \end{aligned}$$

and, for any integer $k \geq 1$, we set

$$E_{(k)} = \begin{pmatrix} 0 & 0 \\ I_k & 0 \end{pmatrix}.$$

Thus, we can compute the LHS of the identity in (3.2) as follows:

$$\begin{aligned} &\det[(A-\lambda_1)^*(A-\lambda_1) + \eta_{z,t}^2] \det[A^*_{a_1,b_1,\theta_1} A_{a_1,b_1,\theta_1} + \eta_{z,t}^2]^{-\frac{1}{2}} \\ &= \left\{ \det \begin{pmatrix} I_{2N} + \frac{ib_1}{2} \left(\tan\theta_1 - \tan^{-1}\theta_1\right)^2 G_{\lambda_1}(\eta_{z,t}) E_{(N)} & \frac{ib}{2} \left(\tan^2\theta_1 - \tan^{-2}\theta_1\right) G_{\lambda_1}(\eta_{z,t}) E_{(N)} \\ -\frac{ib_1}{2} \left(\tan^2\theta_1 - \tan^{-2}\theta_1\right) G_{\overline{\lambda}_1}(\eta_{z,t}) E_{(N)} & I_N - \frac{ib_1}{2} \left(\tan\theta_1 - \tan^{-1}\theta_1\right)^2 G_{\overline{\lambda}_1}(\eta_{z,t}) E_{(N)} \end{pmatrix} \right\}^{-\frac{1}{2}}. \end{aligned}$$

Since the $(1,1)$ and $(1,2)$ blocks of this matrix commute, this determinant is equal to

$$\begin{aligned} &\left\{ \det \left[\left(I_N + \frac{ib_1}{2} \left(\tan\theta_1 - \tan^{-1}\theta_1\right)^2 G_{\lambda_1}(\eta_{z,t}) E_{(N)} \right) \left(I_N - \frac{ib_1}{2} \left(\tan\theta_1 - \tan^{-1}\theta_1\right)^2 G_{\overline{\lambda}_1}(\eta_{z,t}) E_{(N)} \right) \right. \right. \\ &- \left. \left. \frac{b_1^2}{4} \left(\tan^2\theta_1 - \tan^{-2}\theta_1\right)^2 G_{\lambda_1}(\eta_{z,t}) E_{(N)} G_{\overline{\lambda}_1}(\eta_{z,t}) E_{(N)} \right] \right\}^{-\frac{1}{2}} \\ &= \left\{ \det \left[I + \frac{ib_1}{2} \left(\tan\theta_1 - \tan^{-1}\theta_1\right)^2 \left(G_{\lambda_1}(\eta_{z,t}) - G_{\overline{\lambda}_1}(\eta_{z,t}) \right) E_{(N)} \right. \right. \\ &+ \left. \left. b_1^2 \left(2 - \tan^2\theta_1 - \tan^{-2}\theta_1 \right) G_{\lambda_1}(\eta_{z,t}) E_{(N)} G_{\overline{\lambda}_1}(\eta_{z,t}) E_{(N)} \right] \right\}^{-\frac{1}{2}}. \end{aligned}$$

Note that

$$G_{\lambda_1}(\eta_{z,t}) - G_{\overline{\lambda}_1}(\eta_{z,t}) = G_{\lambda_1}(\eta_{z,t}) \begin{pmatrix} 0 & \lambda_1 - \overline{\lambda}_1 \\ \overline{\lambda}_1 - \lambda_1 & 0 \end{pmatrix} G_{\overline{\lambda}_1}(\eta_{z,t})$$
$$= 2b_1 i G_{\lambda_1}(\eta_{z,t})(E^*_{(N)} - E_{(N)})G_{\overline{\lambda}_1}(\eta_{z,t}).$$

Combining the previous three displays gives (3.2). □

We now control the dP integral.

Lemma 11. *Fix $j \geq 1$ and $\theta_j \in [0, \frac{\pi}{4}]$. We have*

$$\left| \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} dP \right|$$
$$\lesssim N^C + C^N N^{-10(N-2j+6)}\theta_j^{-2N} + N^{-5N+10j+12}\theta_j^{-2N}.$$

If $\theta_j \in [\frac{\pi}{4}, \frac{\pi}{2}]$, the same bound holds upon replacing θ_j by $\frac{\pi}{2} - \theta_j$.

Proof. We write the proof in the case where $\theta_j \leq \frac{\pi}{4}$. For the case $\theta_j \geq \frac{\pi}{4}$, it suffices to use the same argument but replace θ_j by $\frac{\pi}{2} - \theta_j$.

For the sake of an upper bound, we can move the absolute value inside the integration and drop the complex exponential. For any $P \in M_2^{sa}(\mathbb{R})$, write $P = UDU^*$, where $U \in \mathbf{U}(2)$ is unitary and D is diagonal. Change-of-variables then shows

$$\int_{M_2^{sa}(\mathbb{R})} |\det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}| dP$$
$$\lesssim \int_{\mathbf{U}(2)} \int_{\mathbb{R}^2} |\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-\frac{1}{2}} |D_{11} - D_{22}| dD_{11} dD_{22} dU,$$

where

$$\mathbf{H} := U^* \otimes I_{N-2j+2} \tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(\eta_{z,t})^{\frac{1}{2}} U \otimes I_{N-2j+2}.$$

Because $\mathbf{U}(2)$ is compact, we only need to control the integral on $\mathbb{R}^2$ uniformly in U. We split the $\mathbb{R}^2$ integration as follows:

$$\int_{\mathbb{R}^2} |\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-\frac{1}{2}} |D_{11} - D_{22}| dD_{11} dD_{22}$$
$$= \int_{|D_{11}|,|D_{22}| \leq N^{10}} |\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-\frac{1}{2}} |D_{11} - D_{22}| dD_{11} dD_{22}$$
$$+ \int_{|D_{11}|,|D_{22}| \geq N^{10}} |\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-\frac{1}{2}} |D_{11} - D_{22}| dD_{11} dD_{22}$$
$$+ \int_{\substack{|D_{11}| \leq N^{10} \\ |D_{22}| \geq N^{10}}} |\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-\frac{1}{2}} |D_{11} - D_{22}| dD_{11} dD_{22}$$
$$+ \int_{\substack{|D_{11}| \geq N^{10} \\ |D_{22}| \leq N^{10}}} |\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-\frac{1}{2}} |D_{11} - D_{22}| dD_{11} dD_{22}$$
$$= \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}.$$

Since $\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}$ is self-adjoint (as $\mathbf{H}$ is self-adjoint), we get that $|\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]| \geq 1$. This implies $|\mathrm{I}| \leq N^C$. Next, we note that

$$|\det[I + i\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-1/2} \leq |\det[\mathbf{H}D \otimes I_{N-2j+2}\mathbf{H}]|^{-1/2}$$
$$= |D_{11}|^{-\frac{N-2j+2}{2}} |D_{22}|^{-\frac{N-2j+2}{2}} |\det \mathbf{H}^2|^{-1/2}$$

for the same reason. But the eigenvalues of $\mathbf{H}^2$ are those of $H^{(j-1)}_{a_j,b_j,\theta_j}(\eta_{z,t})$, which are uniformly $\gtrsim \theta_j^2$ by definition of $H^{(j-1)}_{a_j,b_j,\theta_j}(\eta_{z,t})$ and a bound on the operator norm of $A^{(j-1)}_{a_j,b_j,\theta_j}$ of $\lesssim \theta_j^{-2}$. This gives

$$\begin{aligned}|\mathrm{II}| &\lesssim \int_{|D_{11}|,|D_{22}|\geq N^{10}} |D_{11}|^{-\frac{N-2j+2}{2}}|D_{22}|^{-\frac{N-2j+2}{2}}\theta_j^{-2N}|D_{11}-D_{22}|dD_{11}dD_{22}\\ &\lesssim C^N N^{-10(N-2j+6)}\theta_j^{-2N}.\end{aligned}$$

We are left to bound III; the bound for IV follows by the same argument but swapping D_{11} and D_{22}. For convenience, write

$$\mathbf{H}_1 := I + i\mathbf{H}\begin{pmatrix}0 & 0\\ 0 & D_{22}\end{pmatrix}\otimes I_{N-2j+2}\mathbf{H}.$$

We have

$$\begin{aligned}&|\det[I+i\mathbf{H}D\otimes I_{N-2j+2}\mathbf{H}]|\\ &= |\det[\mathbf{H}_1]|\times\det\left[I+i\mathbf{H}_1^{-\frac{1}{2}}\mathbf{H}\begin{pmatrix}D_{11} & 0\\ 0 & 0\end{pmatrix}I_{N-2j+2}\mathbf{H}\mathbf{H}_1^{-\frac{1}{2}}\right]\\ &\geq \left|\det\left[I+i\mathbf{H}^2\begin{pmatrix}0 & 0\\ 0 & D_{22}\end{pmatrix}\otimes I_{N-2j+2}\right]\right|.\end{aligned}$$

where the last inequality holds because the second factor in the second line has the form $|\det[I+iA]|$ with A self-adjoint. Now, write $\mathbf{H}^2 = \begin{pmatrix}\mathbf{L}_{11} & \mathbf{L}_{12}\\ \mathbf{L}_{12} & \mathbf{L}_{22}\end{pmatrix}$, and compute

$$\mathbf{H}^2\begin{pmatrix}0 & 0\\ 0 & D_{22}\end{pmatrix}\otimes I_{N-2j+2} = D_{22}\begin{pmatrix}0 & \mathbf{L}_{12}\\ 0 & \mathbf{L}_{22}\end{pmatrix}.$$

This gives us

$$\left|\det\left[I+i\mathbf{H}^2\begin{pmatrix}0 & 0\\ 0 & D_{22}\end{pmatrix}\otimes I_{N-2j+2}\right]\right| = \left|\det\begin{pmatrix}1 & iD_{22}\mathbf{L}_{12}\\ 0 & 1+iD_{22}\mathbf{L}_{22}\end{pmatrix}\right| \geq |\det[D_{22}\mathbf{L}_{22}]|.$$

Since $\mathbf{L}_{22}$ is a diagonal block of $\mathbf{H}^2$, and since $\mathbf{H}$ is self-adjoint, we know that $\mathbf{L}_{22}$ is self-adjoint. We also recall $\mathbf{H}^2 \gtrsim \theta_j^2$ from earlier in this proof. This implies $\mathbf{L}_{22} \gtrsim \theta_j^2$ by restricting to vectors in the block corresponding to $\mathbf{L}_{22}$. Thus, the previous display is $\geq C^{-N}D_{22}^{N-2j+2}\theta_j^{-2N}$. Ultimately,

$$\begin{aligned}|\mathrm{III}| &\lesssim \int_{\substack{|D_{11}\leq N^{10}\\ |D_{22}|\geq N^{10}}} |\det[D_{22}\mathbf{L}_{22}]|^{-\frac{1}{2}}|D_{11}-D_{22}|dD_{11}dD_{22}\\ &\lesssim N^{10}\int_{|D_{22}|\geq N^{10}} C^N D_{22}^{-\frac{N-2j+2}{2}}\theta_j^{-2N}dD_{22}\\ &\lesssim C^N N^{-5N+10j+12}\theta_j^{-2N}.\end{aligned}$$

As mentioned earlier, the same bound holds for IV. This completes the proof. □

By the previous lemmas, we can now bound the integral $d\theta_j$ when $\theta_j \in \mathcal{I}_1$.

Lemma 12. *Fix any $j \geq 1$. There exists $\kappa > 0$ such that locally uniformly in a_j, b_j, we have*

$$\begin{aligned}&\int_{\mathcal{I}_1}\frac{|\cos 2\theta_j|}{\sin^2 2\theta_j}\psi_j\det[(A-z)^*(A-z)+\eta_{z,t}^2]\det[(A^{(j-1)}_{a_j,b_j,\theta_j})^*A^{(j-1)}_{a_j,b_j,\theta_j}+\eta_{z,t}^2]^{-\frac{1}{2}}\\ &\times\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr}P}\det[I+i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}dPd\theta_j\\ &\lesssim e^{-N^\kappa}.\end{aligned}$$

Proof. We break up $\mathcal{I}_1 = \mathcal{I}_{11} \cup \mathcal{I}_{12}$, where $\mathcal{I}_{11} = [0, \frac{\pi}{4} - N^{-1/2+\tau}]$ and $\mathcal{I}_{12} = [\frac{\pi}{4} + N^{-1/2+\tau}, \frac{\pi}{2}]$. We will prove the estimate

$$\begin{aligned}
&\int_{\mathcal{I}_{11}} \frac{|\cos 2\theta_j|}{\sin^2 2\theta_j} \psi_j \det[(A-z)^*(A-z) + \eta_{z,t}^2] \det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j} + \eta_{z,t}^2]^{-\frac{1}{2}} \\
&\times \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} dP d\theta_j \\
&\lesssim e^{-N^\kappa}.
\end{aligned}$$

The estimate for $\mathcal{I}_{12}$ instead of $\mathcal{I}_{11}$ follows by the same argument after changing variables $\theta_j \mapsto \frac{\pi}{2} - \theta_j$.

By Lemmas 10 and 11, it suffices to show

$$\begin{aligned}
&\int_{\mathcal{I}_{11}} \frac{|\cos 2\theta_j|}{\sin^2 2\theta_j} N^{C\epsilon_0} \exp\left\{-Cb_j^2 N^2 \eta_{z,t}^2 [\tan\theta_j - \tan^{-1}\theta_j]^2\right\} \\
&\times \left[N^C + C^N N^{-10(N-2j+6)}\theta_j^{-2N} + N^{-5N+10j+12}\theta_j^{-2N}\right] d\theta_j \\
&\lesssim e^{-N^\kappa}.
\end{aligned} \tag{3.3}$$

Now, further decompose $\mathcal{I}_{11} = [0, c] \cup [c, \frac{\pi}{4} - N^{-1/2+\tau}]$, where $c > 0$ is a small, fixed constant.

Assume first that $\theta_j \in [c, \frac{\pi}{4} - N^{-1/2+\tau}]$. We know $|\theta_j - \frac{\pi}{4}| \geq N^{-1/2+\tau}\eta_{z,t}^{-1}$ for $\tau > 0$ small and fixed. Thus, the exponential in the first line in (3.3) is $O(\exp[-CN^{2\tau}])$, and $|\cos 2\theta_j|^{-1} \lesssim N^{1/2}$. So

$$\begin{aligned}
&\int_c^{\frac{\pi}{4} - N^{-1/2+\tau}} \frac{|\cos 2\theta_j|}{\sin^2 2\theta_j} N^{C\epsilon_0} \exp\left\{-Cb_j^2 N^2 \eta_{z,t}^2 [\tan\theta_j - \tan^{-1}\theta_j]^2\right\} \\
&\times \left[N^C + C^N N^{-10(N-2j+6)}\theta_j^{-2N} + N^{-5N+10j+12}\theta_j^{-2N}\right] d\theta_j \\
&\lesssim D^N e^{-CN^{2\tau}} \lesssim e^{-N^\kappa},
\end{aligned}$$

where $D = O(1)$. We now handle the integral on $[0, c]$. In this case, if $c > 0$ is small enough, we have the lower bound $|\tan\theta_j - \tan^{-1}\theta_j| \gtrsim \theta_j^{-1}$. We can also bound $|\cos 2\theta_j| = O(1)$. Thus, we have the following estimate for $D = O(1)$ and $C > 0$:

$$\begin{aligned}
&\int_0^c \frac{|\cos 2\theta_j|}{\sin^2 2\theta_j} N^{C\epsilon_0} \exp\left\{-Cb_j^2 N^2 \eta_{z,t}^2 [\tan\theta_j - \tan^{-1}\theta_j]^2\right\} \\
&\times \left[N^C + C^N N^{-10(N-2j+6)}\theta_j^{-2N} + N^{-5N+10j+12}\theta_j^{-2N}\right] d\theta_j \\
&\lesssim \int_0^c \frac{1}{\sin^2 2\theta_j} D^N \theta_j^{-2N} \exp[-Cb_j^2 N^2 \eta_{z,t}^2 \theta_j^{-2}] d\theta_j.
\end{aligned}$$

But $N^2\eta_{z,t}^2 \gg N$ since $\eta_{z,t} \gtrsim t = N^{-\epsilon_0}$. So, the exponential decays faster than θ_j^{-2N} or $\sin^{-2} 2\theta_j$ blow up as $\theta_j \to 0$. In particular, the last line is $\lesssim \exp[-N^\kappa]$ for some $\kappa > 0$ by elementary calculus. This completes the proof. □

3.2 Proof of Proposition 8

Lemmas 10 and 11 and a straightforward bound $|\det[V_j^* G_z^{(j-1)}(\eta_{z,t})V_j]| \lesssim \|G_z^{(j-1)}(\eta_{z,t})\|_{\text{op}}^4 \lesssim \eta_{z,t}^{-4} \lesssim N^{4\epsilon_0}$ shows that $\Theta_0 \lesssim N^D$ for some $D = O(1)$. Thus, $\Theta_{10}, \Theta_{11} = O(\exp[-N^\kappa])$ by Lemma 12. □

4 Estimates for $\rho_{\mathrm{main}}(z,\mathbf{z};A)$

Let us recall

$$\begin{aligned}
\rho_{\mathrm{main}}(z,\mathbf{z};A) &:= \frac{N^7 b^4[1+O(N^{-\kappa})]}{16\pi^4 t^{12}\gamma_{z,t}^2\sigma_{z,t}^3\upsilon_{z,t}^2}\rho^{(2)}_{\mathrm{GinUE}}(z_1,z_2)\int_{\mathcal{I}_0^2} d\theta_1 d\theta_2 \frac{256|\cos 2\theta_1||\cos 2\theta_2|}{|\sin^2 2\theta_1|\sin^2 2\theta_2|} \\
&\times \prod_{j=1,2}\int_{V^2(\mathbb{R}^{N-2j+2})} d\mu_j(\mathbf{v}_j)|\det[V_j^* G_z^{(j-1)}(\eta_{z,t})V_j]|^2 \\
&\times \prod_{j=1,2}\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2]\det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A^{(j-1)}_{a_j,b_j,\theta_j}+\eta_{z,t}^2]^{-\frac{1}{2}} \\
&\times \prod_{j=1,2}\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr}P}\det[I+i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}dP.
\end{aligned}$$

The goal of this section is to express every term above in terms of traces of resolvents.

4.1 Ratio of determinants

As in the proof of Lemma 10, we start with the estimate

$$\begin{aligned}
\psi_j \det[(A-z)^*(A-z)+\eta_{z,t}^2] &= \det[(A^{(j-1)}-\lambda_j)^*(A^{(j-1)}-\lambda_j)+\eta_{z,t}^2] \\
&\times \prod_{\ell=1}^{j-1}\left|\det\left[V_\ell^* G^{(\ell-1)}_{\lambda_j}(\eta_{z,t})V_\ell\right]\right|^{-1}\left[1+O(N^{-\kappa})\right]
\end{aligned}$$

for some $\kappa>0$ locally uniformly in a_j,b_j. The second line will be addressed by Lemma 13 below, so we deal with the first determinant on the RHS. Unlike the proof of Lemma 10, we want to compute (3.2) more precisely in terms of traces of resolvents; the point is that we now have the a priori estimate $|\theta-\frac{\pi}{4}|\lesssim N^{-1/2+\tau}\eta_{z,t}^{-1}$ since we restrict to $\theta_j\in\mathcal{I}_0$. So, take the RHS of (3.2) and expand in terms of the trace. In particular, we have

$$\begin{aligned}
&\det[(A^{(j-1)}-\lambda_j)(A^{(j-1)}-\lambda_j)^*+\eta_{z,t}^2]\det[A^{(j-1)}_{a_j,b_j,\theta_j}(A^{(j-1)}_{a_j,b_j,\theta_j})^*+\eta_{z,t}^2]^{-\frac{1}{2}} \\
&= \exp\left\{-\sum_{k=1}^{\infty}\frac{1}{k}b_j^{2k}(\tan\theta_j-\tan^{-1}\theta_j)^{2k}\eta_{z,t}^{2k}\operatorname{Tr}\left[H^{(j-1)}_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\right]^k\right\}.
\end{aligned}$$

Since $\theta_j\in\mathcal{I}_0$, we know that $|\tan\theta_j-\tan^{-1}\theta_j|\lesssim N^{-1/2+\tau}\eta_{z,t}^{-1}$, where $\tau>0$ is small. Thus, if we trivially bound the trace of the k-th power by $O(N\eta_{z,t}^{-4k})$, since $\eta_{z,t}=N^{-\epsilon_0}$ with $\epsilon_0>0$ small, the contribution of the sum from $k=2$ and on is $O(N^{-\kappa})$. In particular, we have

$$\begin{aligned}
&\exp\left\{-\sum_{k=1}^{\infty}\frac{1}{k}b_j^{2k}(\tan\theta_j-\tan^{-1}\theta_j)^{2k}\eta_{z,t}^{2k}\operatorname{Tr}\left[H^{(j-1)}_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\right]^k\right\} \\
&= \exp\left\{-b_j^2(\tan\theta_j-\tan^{-1}\theta_j)^2\eta_{z,t}^2\operatorname{Tr}H^{(j-1)}_{\lambda_j}(\eta_{z,t})\tilde{H}^{(j-1)}_{\overline{\lambda}_j}(\eta_{z,t})\right\}\left[1+O(N^{-\kappa})\right].
\end{aligned}$$

By Cauchy interlacing and trivial resolvent bounds, we can remove the superscript:

$$\begin{aligned}
\operatorname{Tr}H^{(j-1)}_{\lambda_j}(\eta_{z,t})\tilde{H}^{(j-1)}_{\overline{\lambda}_j}(\eta_{z,t}) &= \operatorname{Tr}H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\tilde{H}^{(j-1)}_{\overline{\lambda}_j}(\eta_{z,t})H^{(j-1)}_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}} \\
&= \operatorname{Tr}H_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})H_{\lambda_j}(\eta_{z,t})^{\frac{1}{2}}+O(N^{-1+2\tau}\eta_{z,t}^{-D}) \\
&= \operatorname{Tr}H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})+O(N^{-1+2\tau}\eta_{z,t}^{-D}).
\end{aligned} \tag{4.1}$$

The error term on the RHS is obtained by our bound on $\theta_j - \frac{\pi}{4}$ and a trivial operator norm bound on resolvents of $\eta_{z,t}^{-D}$. But $\eta_{z,t} = N^{-\epsilon_0}$, so if we choose ϵ_0 small enough, this cost is $\lesssim N^{-\kappa}$. Putting this altogether, we have

$$\psi_j \det[(A-z)^*(A-z) + \eta_{z,t}^2] \det[(A^{(j-1)}_{a_j,b_j,\theta_j})^* A_{a_j,b_j,\theta_j} + \eta_{z,t}^2]^{-\frac{1}{2}}$$
$$\approx \exp\left\{-b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2 \eta_{z,t}^2 \mathrm{Tr}\, H_{\lambda_j}(\eta_{z,t}) \tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\right\} \prod_{\ell=1}^{j-1} \left|\det\left[V_\ell^* G^{(\ell-1)}_{\lambda_j}(\eta_{z,t}) V_\ell\right]\right|^{-1},$$

where $\approx$ means equal to modulo a factor of $1 + O(N^{-\kappa})$.

4.2 Estimating $\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]$

Write $G_z^{(j-1)} = G_z^{(j-1)}(\eta_{z,t})$ and $H_z^{(j-1)} = H_z^{(j-1)}(\eta_{z,t})$. For simplicity, we focus on the case $j = 1$. We comment on the more general case at the end of this subsection. For $\mathbf{v}_1 \in V^2(\mathbb{R}^N)$, we write $\mathbf{v}_1 = (\mathbf{v}_{11}, \mathbf{v}_{12})$. By the block representation for G_z, we have

$$V_1^* G_z V_1 = \begin{pmatrix} i\eta \mathbf{v}_{11}^* H_z \mathbf{v}_{11} & i\eta \mathbf{v}_{11}^* H_z \mathbf{v}_{12} & \mathbf{v}_{11}^* H_z (A-z) \mathbf{v}_{11} & \mathbf{v}_{11}^* H_z (A-z) \mathbf{v}_{12} \\ i\eta \mathbf{v}_{12}^* H_z \mathbf{v}_{11} & i\eta \mathbf{v}_{12}^* H_z \mathbf{v}_{12} & \mathbf{v}_{12}^* H_z (A-z) \mathbf{v}_{11} & \mathbf{v}_{12}^* H_z (A-z) \mathbf{v}_{21} \\ \mathbf{v}_{11}^* (A-z)^* H_z \mathbf{v}_{11} & \mathbf{v}_{11}^* (A-z)^* H_z \mathbf{v}_{12} & i\eta \mathbf{v}_{11}^* \tilde{H}_z \mathbf{v}_{11} & i\eta \mathbf{v}_{11}^* \tilde{H}_z \mathbf{v}_{12} \\ \mathbf{v}_{12}^* (A-z)^* H_z \mathbf{v}_{11} & \mathbf{v}_{12}^* (A-z)^* H_z \mathbf{v}_{12} & i\eta \mathbf{v}_{12}^* \tilde{H}_z \mathbf{v}_{11} & i\eta \mathbf{v}_{12}^* \tilde{H}_z \mathbf{v}_{12} \end{pmatrix}.$$

Before we estimate its determinant, we must introduce notation. For any a, b, θ, define $H_{a,b,\theta}(\eta) = [(I_2 \otimes A - \Lambda_{a,b,\theta} \otimes I_N)(I_2 \otimes A - \Lambda_{a,b,\theta} \otimes I_N)^* + \eta^2]^{-1}$ and $H_{a,b,\theta} = H_{a,b,\theta}(\eta_{z,t})$.

Lemma 13. *For any $p = O(1)$, there exists $\kappa > 0$ such that*

$$\int_{V^2(\mathbb{R}^N)} |\det[V_1^* G_z(\eta_{z,t}) V_1]|^p d\mu_1(\mathbf{v}_1) = |\det \mathcal{G}|^p \left[1 + O(N^{-\kappa})\right],$$

where $\mathcal{G}$ is the following 4×4 matrix:

$$\mathcal{G} := \begin{pmatrix} \mathcal{G}_{11} & \mathcal{G}_{12} \\ \mathcal{G}_{21} & \mathcal{G}_{22} \end{pmatrix}$$
$$\mathcal{G}_{11} := \begin{pmatrix} i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(H_z \otimes E_{11}) & i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(H_z \otimes E_{12}) \\ i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(H_z \otimes E_{21}) & i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(H_z \otimes E_{22}) \end{pmatrix}$$
$$\mathcal{G}_{12} := \begin{pmatrix} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[H_z(A-z) \otimes E_{11}] & \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[H_z(A-z) \otimes E_{12}] \\ \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[H_z(A-z) \otimes E_{21}] & \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[H_z(A-z) \otimes E_{22}] \end{pmatrix}$$
$$\mathcal{G}_{21} := \begin{pmatrix} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[(A-z)^* H_z \otimes E_{11}] & \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[(A-z)^* H_z \otimes E_{12}] \\ \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[(A-z)^* H_z \otimes E_{21}] & \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}[(A-z)^* H_z \otimes E_{22}] \end{pmatrix}$$
$$\mathcal{G}_{22} := \begin{pmatrix} i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(\tilde{H}_z \otimes E_{11}) & i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(\tilde{H}_z \otimes E_{12}) \\ i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(\tilde{H}_z \otimes E_{21}) & i\eta_{z,t} \frac{t}{N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1}(\tilde{H}_z \otimes E_{22}) \end{pmatrix},$$
$$E_{11} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{22} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad E_{12} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad E_{21} := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Proof. Let $A_{a,b,\theta} = I_2 \otimes A - \Lambda_{a,b,\theta} \otimes I_N$, For any Hermitian matrix F, we start by introducing

$$m_F(r) := \frac{e^{-\frac{rt}{2N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1} F}}{K_{a_1,b_1,\theta_1}} \int_{V^2(\mathbb{R}^N)} \exp\left\{-\frac{N}{t} \mathbf{v}^* \left(A^*_{a_1,b_1,\theta_1} A_{a_1,b_1,\theta_1} - \frac{rt}{N} F\right) \mathbf{v}\right\} d\mathbf{v}$$
$$= \frac{e^{-\frac{rt}{2N} \mathrm{Tr}\, \tilde{H}_{a_1,b_1,\theta_1} F} e^{\frac{2N}{t} \eta_{z,t}^2}}{K_{a_1,b_1,\theta_1}} \int_{V^2(\mathbb{R}^N)} \exp\left\{-\frac{N}{t} \mathbf{v}^* \left(A^*_{a_1,b_1,\theta_1} A_{a_1,b_1,\theta_1} + \eta_{z,t}^2 - \frac{rt}{N} F\right) \mathbf{v}\right\} d\mathbf{v},$$

where K_{a_1,b_1,θ_1} is a normalizing constant chosen so that $m_F(0) = 1$. By Lemma 6, as in the proof of Lemma 7, we can compute

$$\begin{aligned}
&\int_{V^2(\mathbb{R}^N)} \exp\left\{-\frac{N}{t}\mathbf{v}^*\left(A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t} - \frac{rt}{N}F\right)\mathbf{v}\right\} d\mathbf{v} \\
&= C_{N,t}\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{t}\operatorname{Tr} P}\left\{\det\left(A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t} - \frac{rt}{N}F + iP\otimes I_N\right)\right\}^{-\frac{1}{2}} dP \\
&= C_{N,t}\left\{\det\left(A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t} - \frac{rt}{N}F\right)\right\}^{-\frac{1}{2}} \\
&\times \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{t}\operatorname{Tr} P}\left\{\det\left(1 + \sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(iP\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)\right\}^{-\frac{1}{2}} dP,
\end{aligned}$$

where $\tilde{H}^{(rF)}_{a_1,b_1,\theta_1} := (A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t} - \frac{rt}{N}F)^{-1}$. Above, and throughout this proof, $C_{N,t}$ is a constant that comes from our application of Lemma 6; its exact value is not important. Next,

$$\begin{aligned}
&\left\{\det\left(A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t} - \frac{rt}{N}F\right)\right\}^{-\frac{1}{2}} \\
&= \left\{\det\left(A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t}\right)\right\}^{-\frac{1}{2}}\left\{\det\left(1 - \tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\right)\right\}^{-\frac{1}{2}} \\
&= \left\{\det\left(A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t}\right)\right\}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\operatorname{Tr}\log\left(1 - \tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\right)\right\},
\end{aligned}$$

where $\tilde{H}_{a_1,b_1,\theta_1} := \tilde{H}^{(0)}_{a_1,b_1,\theta_1}$. Next, we compute

$$\begin{aligned}
&\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{t}\operatorname{Tr} P}\left\{\det\left(1 + \sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(iP\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)\right\}^{-\frac{1}{2}} dP \\
&= \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{t}\operatorname{Tr} P}\exp\left\{-\frac{1}{2}\operatorname{Tr}\log\left(1 + \sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(iP\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)\right\} dP.
\end{aligned}$$

Putting all the previous displays together, we get

$$\begin{aligned}
m(r) &= \frac{C_{N,t}e^{2\frac{N}{t}\eta^2_{z,t}}}{K_{a_1,b_1,\theta_1}}\left\{\det\left(A^*_{a_1,b_1,\theta_1}A_{a_1,b_1,\theta_1} + \eta^2_{z,t}\right)\right\}^{-\frac{1}{2}} \\
&\times \exp\left\{-\frac{rt}{2N}\operatorname{Tr}\tilde{H}_{a_1,b_1,\theta_1}F - \frac{1}{2}\operatorname{Tr}\log\left(1 - \tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\right)\right\} \\
&\times \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{t}\operatorname{Tr} P}\exp\left\{-\frac{1}{2}\operatorname{Tr}\log\left(1 + \sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(iP\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)\right\} dP.
\end{aligned} \tag{4.2}$$

This identity can only hold if $1 - \tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1} > 0$; we will always have this condition for any F we take. To control the second line, we have the following as in Lemma 6.2 of [23]:

$$\begin{aligned}
&\exp\left\{-\frac{rt}{2N}\operatorname{Tr}\tilde{H}_{a_1,b_1,\theta_1}F - \frac{1}{2}\operatorname{Tr}\log\left(1 - \tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\right)\right\} \\
&\lesssim \exp\left\{\frac{1}{\|1 - \tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\|_{\mathrm{op}}}\operatorname{Tr}\left(\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\right)^2\right\}.
\end{aligned}$$

Let us now control the dP integral in (4.2). Next, by Taylor expansion as in Section 6 of [23],

$$\begin{aligned}
&\exp\left\{-\frac{1}{2}\mathrm{Tr}\,\log\left(1+\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(iP\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)\right\}\\
&=\exp\left\{-\frac{i}{2}\mathrm{Tr}\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}-\frac{1}{4}\mathrm{Tr}\left(\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)^2+\mathcal{E}\right\},
\end{aligned}$$

where $\mathcal{E}$ satisfies both $|e^{\mathcal{E}}|=O(1)$ and $|\mathcal{E}|\lesssim N\eta_{z,t}^{-6}\|P\|_{\mathrm{op}}^3$. Note that

$$\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}=\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\left(1-\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\frac{rt}{N}F\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\right)^{-1}\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}.$$

Now, assume that $\|\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\|_{\mathrm{op}}\lesssim\frac{N^{\epsilon}}{N\eta_{z,t}^3}$ uniformly in N. We will also always have this constraint, and it clearly implies $1-\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}>0$ if $\eta_{z,t}\gg N^{-\frac{1}{3}+\epsilon}$. Next, we compute

$$\begin{aligned}
&\mathrm{Tr}\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}=\mathrm{Tr}\,\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}(P\otimes I_N)\\
&=\mathrm{Tr}\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\left(1-\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\frac{rt}{N}F\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\right)^{-1}\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}(P\otimes I_N)\\
&=\mathrm{Tr}\,\tilde{H}_{a_1,b_1,\theta_1}(P\otimes I_N)+O\left(\frac{N^{\epsilon}}{\eta_{z,t}^5}\|P\|_{\mathrm{op}}\right),
\end{aligned}\tag{4.3}$$

where the last line also uses $\|\tilde{H}_{a_1,b_1,\theta_1}\|_{\mathrm{op}}\leq\eta_{z,t}^{-2}$. By the same token, for some constant $C>0$, we have

$$\begin{aligned}
&\left(1-\frac{CN^{\epsilon}}{\eta_{z,t}^{10}}\|P\|_{\mathrm{op}}^2\right)\mathrm{Tr}\left(\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\right)^2\\
&\leq\mathrm{Tr}\left(\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)^2\\
&\leq\left(1-\frac{CN^{\epsilon}}{\eta_{z,t}^{10}}\|P\|_{\mathrm{op}}^2\right)\mathrm{Tr}\left(\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\right)^2.
\end{aligned}\tag{4.4}$$

This has two consequences. First, by the lower bound above and the lower bound $\tilde{H}_{a_1,b_1,\theta_1}\gtrsim1$,

$$\mathrm{Tr}\left(\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)^2\gtrsim N|P_{11}|^2+N|P_{22}|^2+N|P_{12}|^2.\tag{4.5}$$

Thus, we can restrict the P integration to $|P_{11}|+|P_{12}|+|P_{22}|\lesssim N^{-1/2}\log N$ at the cost of something exponentially small in N. In particular, we have $\mathcal{E}=O(N^{-1/2+\kappa})$ for some $\kappa>0$ small. This bound on entries of P combined with (4.3) also implies that

$$\begin{aligned}
&\mathrm{Tr}\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}=\mathrm{Tr}\,\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}(P\otimes I_N)\\
&=\mathrm{Tr}\,\tilde{H}_{a_1,b_1,\theta_1}(P\otimes I_N)+O\left(\frac{N^{2\kappa}}{\sqrt{N}\eta_{z,t}^5}\right).
\end{aligned}$$

Moreover, in the region $|P_{11}|+|P_{12}|+|P_{22}|\lesssim N^{-1/2}\log N$, we have

$$\mathrm{Tr}\left(\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\right)^2\lesssim\eta_{z,t}^{-C}\log^2N.$$

By combining this with (4.4), in this region, we have the following for $\kappa > 0$ small:

$$\operatorname{Tr}\left(\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)^2 = \operatorname{Tr}\left(\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\right)^2 + O\left(\frac{N^{\kappa}}{N\eta_{z,t}^{10}}\right).$$

Ultimately, for possibly different but still small $\kappa > 0$, we get

$$\begin{aligned}
&\exp\left\{-\frac{i}{2}\operatorname{Tr}\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}} - \frac{1}{4}\operatorname{Tr}\left(\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}^{(rF)}_{a_1,b_1,\theta_1}}\right)^2 + \mathcal{E}\right\}\\
&= \exp\left\{-\frac{i}{2}\operatorname{Tr}\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}_{a_1,b_1,\theta_1}} - \frac{1}{4}\operatorname{Tr}\left(\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}(P\otimes I_N)\sqrt{\tilde{H}_{a_1,b_1,\theta_1}}\right)^2\right\}\\
&\times \exp\left\{iO\left(\frac{N^{\kappa}}{\sqrt{N}\eta_{z,t}^{5}}\right) + O\left(\frac{N^{\kappa}}{N\eta_{z,t}^{10}}\right)\right\}.
\end{aligned}$$

Note that the second line above, which contains the main term on the RHS of this identity, is independent of r. Thus, by the previous display, we have $m_F(r) \lesssim m_F(0)$ as long as F is Hermitian and $\|\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}F\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\|_{\mathrm{op}} \lesssim \frac{N^{\epsilon}}{N\eta_{z,t}^{3}} \ll 1$. Recalling that $m_F(0) = 1$, under this condition, we have $m_F(r) = O(1)$. Concentration of matrix entries of $V_1^* G_z V_1$ now follows as Lemma 6.2 in [23]. In particular, we will apply this inequality for $r = N^{-\kappa}$ with $\kappa > 0$ small and for the following choices of F:

$$\begin{aligned}
F &= \eta_{z,t}\mathrm{Re}\left[E\otimes H_z\right],\\
F &= \eta_{z,t}\mathrm{Im}\left[E\otimes H_z\right],\\
F &= \eta_{z,t}\mathrm{Re}\left[E\otimes \tilde{H}_z\right],\\
F &= \eta_{z,t}\mathrm{Im}\left[E\otimes \tilde{H}_z\right],\\
F &= \mathrm{Re}\left[E\otimes H_z(A-z)\right],\\
F &= \mathrm{Im}\left[E\otimes H_z(A-z)\right],
\end{aligned}$$

where E is any of the four matrices

$$\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

and $\mathrm{Re}\,X = \frac{1}{2}(X^*+X)$, $\mathrm{Im}\,X = \frac{i}{2}(X^*-X)$. (We note that our choice of r is larger than the choices made in the proof of Lemma 6.2 in [23], giving weaker concentration estimates, but this is fine.) Let us illustrate one example. Upon treating μ_j as a probability measure over $\mathbf{v}$, by the Markov inequality, we have

$$\mu_j\left(\mathbf{v}^* F\mathbf{v} - \frac{t}{N}\operatorname{Tr}\tilde{H}_{a_1,b_1,\theta_1}F \geq r\right) \leq e^{-Kr^2}m_{KF}(r).$$

Choose $F = \eta_{z,t}\mathrm{Re}\left[E\otimes H_z\right]$ and $r = N^{-\kappa}$ and $K = N^{3\kappa}$ for $\kappa > 0$ small. Then we have the estimate $\|\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\frac{rt}{N}KF\tilde{H}^{\frac{1}{2}}_{a_1,b_1,\theta_1}\|_{\mathrm{op}} \lesssim N^{-1+2\kappa}\eta_{z,t}^{-3} \lesssim 1$, so $m_{KF}(r) = O(1)$. This shows that the event on the LHS is exponentially small in μ_j-measure. By the same token, we know the same is true if we replace $\geq r$ on the LHS by $\leq r$, and we can replace $\mathrm{Re}\,[E\otimes H_z]$ by $\mathrm{Im}\,[E\otimes H_z]$. For example, this implies $v_1^* H_z v_1 = \frac{\eta_{z,t}t}{N}\operatorname{Tr}\tilde{H}_{a_1,b_1,\theta_1}H_z + O(N^{-\kappa})$ outside an event of μ_j-probability that is exponentially small. Doing the same procedure for each entry of $V_1^* G_z V_1$ shows that each said entry matches the corresponding entry of $\mathcal{G}$ up to $O(N^{-\kappa})$ outside an event of μ_j-probability that is exponentially small. Since the determinant is smooth in its entries, and since the entries of $V_1^* G_z V_1$ are $O(\eta_{z,t}^{-1}) = O(N^{-\epsilon_0})$ deterministically with $\epsilon_0 > 0$ small, we deduce

$$\int_{V^2(\mathbb{R}^N)} |\det[V_1^* G_z(\eta_{z,t}) V_1]|^p d\mu_1(\mathbf{v}_1) = |\det \mathcal{G}|^p + O(N^{-\kappa}).$$

We now claim that $|\det \mathcal{G}|^p \gtrsim \eta_{z,t}^{-D}$ for some $D = O(1)$, since we can take $\eta_{z,t} = N^{-\epsilon_0}$ for $\epsilon_0 > 0$ small enough so that $\eta_{z,t}^{-D} \gg N^{-\kappa}$. This implies that the RHS of the previous display is $|\det \mathcal{G}|^p[1 + O(N^{-\kappa})]$ for possibly different $\kappa > 0$. But said claim is equivalent to $|\det[V_1^* G_z(\eta_{z,t}) V_1]| \gtrsim \eta_{z,t}^{-D}$ because of the previous display, and this bound is shown in the proof of Lemma 10. ☐

By the same argument, we deduce the following analog of Lemma 13 but for arbitrary $j \geq 1$.

Lemma 14. *For any $p \geq 1$ and $j \geq 1$, there exists $\kappa > 0$ such that*

$$\int_{V^2(\mathbb{R}^{N-2j+2})} |\det[V_j^* G_z^{(j-1)}(\eta_{z,t}) V_j]|^p d\mu_j(\mathbf{v}_j) = |\det \mathcal{G}_j|^p \left[1 + O(N^{-\kappa})\right],$$

where $\mathcal{G}_j$ is the following 4×4 matrix:

$$\mathcal{G}_{j,a_j,b_j} := \begin{pmatrix} \mathcal{G}_{j,a_j,b_j,11} & \mathcal{G}_{j,a_j,b_j,12} \\ \mathcal{G}_{j,a_j,b_j,21} & \mathcal{G}_{j,a_j,b_j,22} \end{pmatrix}$$

$$\mathcal{G}_{j,a_j,b_j,11} := \begin{pmatrix} i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(H_z^{(j-1)} \otimes E_{11}) & i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(H_z^{(j-1)} \otimes E_{12}) \\ i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(H_z^{(j-1)} \otimes E_{21}) & i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(H_z^{(j-1)} \otimes E_{22}) \end{pmatrix}$$

$$\mathcal{G}_{j,a_j,b_j,12} := \begin{pmatrix} \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[H_z^{(j-1)}(A^{(j-1)} - z) \otimes E_{11}] & \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[H_z^{(j-1)}(A^{(j-1)} - z) \otimes E_{12}] \\ \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[H_z^{(j-1)}(A^{(j-1)} - z) \otimes E_{21}] & \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[H_z^{(j-1)}(A^{(j-1)} - z) \otimes E_{22}] \end{pmatrix}$$

$$\mathcal{G}_{j,a_j,b_j,21} := \begin{pmatrix} \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[(A^{(j-1)} - z)^* H_z^{(j-1)} \otimes E_{11}] & \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[(A^{(j-1)} - z)^* H_z^{(j-1)} \otimes E_{12}] \\ \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[(A^{(j-1)} - z)^* H_z^{(j-1)} \otimes E_{21}] & \frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}[(A^{(j-1)} - z)^* H_z^{(j-1)} \otimes E_{22}] \end{pmatrix}$$

$$\mathcal{G}_{j,a_j,b_j,22} := \begin{pmatrix} i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(\tilde{H}_z^{(j-1)} \otimes E_{11}) & i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(\tilde{H}_z^{(j-1)} \otimes E_{12}) \\ i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(\tilde{H}_z^{(j-1)} \otimes E_{21}) & i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}(\tilde{H}_z^{(j-1)} \otimes E_{22}) \end{pmatrix},$$

$$E_{11} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{22} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad E_{12} := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad E_{21} := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The final step in this computation is to show that in $\mathcal{G}_{j,a_j,b_j}$, we can replace a_j, b_j by a, b (i.e. replace λ_j by z) and remove the $(j-1)$-superscripts. For the exact formula for $\mathcal{G}_{j,a,b}$ appearing in Lemma 15 below, see (4.11).

Lemma 15. *Define $\mathcal{G}_{j,a,b}$ in the same way as in Lemma 14 but replacing (a_j, b_j) by (a, b) and θ_j by $\pi/4$ and removing the $(j-1)$ superscript. Then we have $|\det \mathcal{G}_{j,a_j,b_j}|^p = |\det \mathcal{G}_{j,a,b}|^p[1 + O(N^{-\kappa})]$ for some $\kappa > 0$.*

Proof. Resolvent perturbation implies that $\|\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j} - \tilde{H}^{(j-1)}_{a,b,\frac{\pi}{4}}\|_{\mathrm{op}} \lesssim \eta_{z,t}^{-C}[|a_j - a| + |b_j - b| + |\theta_j - \frac{\pi}{4}|]$. But $a_j - a, b_j - b = O(N^{-1/2})$ and $|\theta_j - \frac{\pi}{4}| = O(N^{-1/2+\kappa}\eta_{z,t}^{-1})$. Moreover, if we use Cauchy interlacing as in (4.1), then for any $k, \ell \in \{1, 2\}$, we have

$$i\eta_{z,t}\frac{t}{N}\left|\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\frac{\pi}{4}}(\tilde{H}_z^{(j-1)} \otimes E_{11}) - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\frac{\pi}{4}}(\tilde{H}_z \otimes E_{11})\right| \lesssim \tfrac{t}{N}\|\tilde{H}_{a_j,b_j,\theta_j}\|_{\mathrm{op}}\|H_z\|_{\mathrm{op}} \lesssim N^{-1}\eta_{z,t}^{-C}.$$

A similar estimate for removing $(j-1)$-superscripts in entries of $\mathcal{G}_{j,a_j,b_j}$ also holds. Thus, the entries of $\mathcal{G}_{j,a_j,b_j} - \mathcal{G}_{j,a,b}$ are $O(N^{-1/2+\kappa+C\epsilon_0})$. Since the determinant is smooth in its entries, we deduce that $|\det \mathcal{G}_{j,a_j,b_j}|^p = |\det \mathcal{G}_{j,a,b}|^p + O(N^{-1/2+4\epsilon_0})$. Now, again use the lower bound $|\det \mathcal{G}_{j,a_j,b_j}|^p \gtrsim \eta_{z,t}^{-D}$ for some $D = O(1)$ to get $|\det \mathcal{G}_{j,a_j,b_j}|^p \gtrsim \eta_{z,t}^{-D}$. This implies that $|\det \mathcal{G}_{j,a,b}|^p + O(N^{-1/2+4\epsilon_0}) = |\det \mathcal{G}_{j,a,b}|^p[1 + O(N^{-1/2+4\epsilon_0})]$ and completes the proof. $\square$

4.3 The dP integration over $M_2^{sa}(\mathbb{R})$

We start by a similar computation to the dP integration in the proof of Lemma 13. By $\det[1+A] = \exp \operatorname{Tr} \log[1+A]$ and Taylor expansion of log, we have

$$
\begin{aligned}
&\det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} \\
&= \exp\left\{-\frac{i}{2}\operatorname{Tr}[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}\right\} \\
&\times \exp\left\{-\frac{1}{2}\operatorname{Tr}\left([\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}](P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}\right)\right\} \\
&\times \exp\left\{O(N\eta_{z,t}^{-6}\|P\|_{\mathrm{op}}^3)\right\}.
\end{aligned} \tag{4.6}
$$

We also have the following bound (by an essentially identical Taylor expansion) which is an inequality version of the previous display:

$$
\begin{aligned}
&\det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} \\
&\lesssim \exp\left\{-\frac{1}{2}\operatorname{Tr}\left([\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}](P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}\right)\right\}.
\end{aligned} \tag{4.7}
$$

If we write $P = \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix}$, then an elementary computation shows

$$
\begin{aligned}
&\exp\left\{-\frac{1}{2}\operatorname{Tr}\left([\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}](P \otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}\right)\right\} \\
&= \exp\left\{-\begin{pmatrix} P_{11} \\ P_{12} \\ P_{22} \end{pmatrix}^* \mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j} \begin{pmatrix} P_{11} \\ P_{12} \\ P_{22} \end{pmatrix}\right\},
\end{aligned}
$$

in which we use the notation

$$\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j} = \begin{pmatrix} \tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11} & \tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \\ \tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} & \tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22} \end{pmatrix}$$

$$\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j} := \begin{pmatrix} [\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]_1 & [\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]_2 & [\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]_3 \end{pmatrix},$$

$$[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]_1 = \begin{pmatrix} \mathrm{Tr}\,[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}]^2 \\ 2\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \\ \mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}]^T \end{pmatrix}$$

$$[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]_2 = \begin{pmatrix} 2\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \\ 2\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22} + 2\mathrm{Tr}\,[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}]^2 \\ 2\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \end{pmatrix}$$

$$[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]_3 = \begin{pmatrix} \mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}]^T \\ 2\mathrm{Tr}\,\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \\ \mathrm{Tr}\,[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}]^2 \end{pmatrix}$$

We clarify that $\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}$ is a 3×3 matrix. To continue with this computation, we need the following auxiliary bound. (It is a refinement of (4.5).)

Lemma 16. *If $\theta \in \mathcal{I}_0$, then $\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j} \gtrsim N\eta_{z,t}^{-2}$.*

Proof. This is shown for $\theta_j = \pi/4$ in Appendix B.1. For general $\theta_j \in \mathcal{I}_0$, we will perform resolvent perturbation in the entries with respect to θ_j. We show one example; we claim that

$$\mathrm{Tr}\,[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}]^2 = \mathrm{Tr}\,[\tilde{H}^{(j-1)}_{a_j,b_j,\frac{\pi}{4},11}]^2 + O(N^{\frac{1}{2}+\delta})$$

for $\delta > 0$ small. For this, note $\|\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\|_{\mathrm{op}} \lesssim \eta_{z,t}^{-2}$. Thus, resolvent perturbation shows

$$\|[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}]^2 - [\tilde{H}^{(j-1)}_{a_j,b_j,\frac{\pi}{4},11}]^2\|_{\mathrm{op}} \lesssim \eta_{z,t}^{-D}\left\|(\Lambda_{a_j,b_j,\theta_j} - \Lambda_{a_j,b_j,\frac{\pi}{4}}) \otimes I_{N-2j}\right\|_{\mathrm{op}}$$

for some $D = O(1)$. Since $\theta \in \mathcal{I}_0$, we know $\|\Lambda_{a_j,b_j,\theta_j} - \Lambda_{a_j,b_j,\frac{\pi}{4}}\|_{\mathrm{op}} \lesssim |\theta - \frac{\pi}{4}| \lesssim N^{-\frac{1}{2}+\tau}\eta_{z,t}^{-1}$ with $\tau > 0$ small. So the RHS of the previous display is $\lesssim N^{-1/2+\tau}\eta_{z,t}^{-D}$ for different $D = O(1)$ and $\tau > 0$ small. We deduce

$$|\mathrm{Tr}\,[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}]^2 - \mathrm{Tr}\,[\tilde{H}^{(j-1)}_{a_j,b_j,\frac{\pi}{4},11}]^2| \lesssim N\|[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}]^2 - [\tilde{H}^{(j-1)}_{a_j,b_j,\frac{\pi}{4},11}]^2\|_{\mathrm{op}} \lesssim N^{\frac{1}{2}+\tau}\eta_{z,t}^{-D}$$

at which point the claim follows if $\epsilon_0 > 0$ in $\eta_{z,t} = N^{-\epsilon_0}$ and $\tau > 0$ are both small enough. A similar estimate holds for all entries of $\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}$. This implies

$$\|\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j} - \mathbf{Q}^{(j-1)}_{a_j,b_j,\frac{\pi}{4}}\|_{\mathrm{op}} \lesssim N^{\frac{1}{2}+\delta}$$

with $\delta > 0$ small. Since the lemma is true for $\theta_j = \pi/4$ as noted at the beginning of this proof, and since $N\eta_{z,t}^{-2} \gg N^{1/2+\delta}$ for $\delta > 0$ small, the lemma must be true for all $\theta_j \in \mathcal{I}_0$. $\square$

By Lemma 16 and (4.7), we deduce the following for any $\upsilon > 0$

$$\left|\int_{\|P\|_{\mathrm{op}} \geq N^{-\frac{1}{2}+\upsilon}\eta_{z,t}} e^{i\frac{N}{2t}\mathrm{Tr}\,P}\det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}dP\right| \lesssim \exp[-CN^{2\upsilon}].$$

In particular, we can restrict to the region $\|P\|_{\mathrm{op}} \leq N^{-1/2+\upsilon}\eta_{z,t}$. In this region, we can use the identity (4.6) and control the last line therein (the cubic term) by $N^{-1/2+3\tau}$ for $\tau > 0$ small. In particular, we have

$$\begin{aligned}
&\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} dP \\
&= [1+O(N^{-\kappa})] \int_{\|P\|_{\mathrm{op}} \leq N^{-1/2+\upsilon}\eta_{z,t}} e^{i\frac{N}{2t}P} \exp\left\{-\frac{i}{2}\operatorname{Tr}[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}\right\} \\
&\qquad \times \exp\left\{-\frac{1}{2}\operatorname{Tr}\left([\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}](P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}\right)\right\} dP + O(\exp[-CN^{2\upsilon}]).
\end{aligned}$$

A similar argument lets us extend the integration back to all $M_2^{sa}(\mathbb{R})$. We ultimately have

$$\begin{aligned}
&\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} dP \\
&= [1+O(N^{-\kappa})] \int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \exp\left\{-\frac{i}{2}\operatorname{Tr}[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}\right\} \\
&\qquad \exp\left\{-\begin{pmatrix} P_{11} \\ P_{12} \\ P_{22}\end{pmatrix}^* \mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j} \begin{pmatrix} P_{11} \\ P_{12} \\ P_{22}\end{pmatrix}\right\} dP + O(\exp[-CN^{2\upsilon}]).
\end{aligned}$$

We can compute the dP integration above since it is just a Gaussian Fourier transform; this gives

$$\begin{aligned}
&\int_{M_2^{sa}(\mathbb{R})} e^{i\frac{N}{2t}\operatorname{Tr} P} \det[I + i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}} dP \\
&= \mathbf{C}_1 \left[1 + O(e^{-CN^{2\upsilon}})\right] |\det \mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}|^{-\frac{1}{2}} \\
&\times \exp\left\{-\mathbf{C}_2 \begin{pmatrix} \frac{N}{t} - \operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11} \\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix}^* [\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1} \begin{pmatrix} \frac{N}{t} - \operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11} \\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\} \\
&+ O(\exp[-CN^{2\upsilon}]).
\end{aligned} \tag{4.8}$$

Above, $\mathbf{C}_1$ and $\mathbf{C}_2$ are constants coming from the Gaussian integration. We now remove superscripts $(j-1)$. By Cauchy interlacing, we know $\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,k\ell} = \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,k\ell} + O(\eta_{z,t}^{-2})$, where the big-O comes from a trivial bound on the operator norm of $\tilde{H}_{a_j,b_j,\theta_j}$.

Now, note $\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j} \gtrsim N\eta_{z,t}^{-2}$. We now claim that

$$\left\| \begin{pmatrix} \frac{N}{t} - \operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11} \\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix} \right\| \lesssim N^{\frac{1}{2}+\kappa}\eta_{z,t}^{-D}, \tag{4.9}$$

where $\kappa > 0$ is small and $D = O(1)$. To see this, we first remark that if $a_j = a$ and $b_j = b$ and $\theta_j = \frac{\pi}{4}$ and $j-1 = 0$, then the vector on the LHS is the zero vector; see Appendix B. We now use Cauchy interlacing and $a_j - a = O(N^{-1/2})$ and $b_j - b = O(N^{-1/2})$ and $\theta_j - \frac{\pi}{4} = O(N^{-1/2+\kappa}\eta_{z,t}^{-1})$ for $\theta_j \in \mathcal{I}_0$ and standard Green's function perturbations to conclude.

We now combine the previous display with $\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,k\ell} = \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,k\ell} + O(\eta_{z,t}^{-2})$ and Lemma

16 to get

$$
\begin{aligned}
&\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\frac{1}{2}\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}\\
&=\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}\\
&\times\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1}\left[\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\frac{1}{2}\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix}-\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right]\right\}\\
&\times\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1}\left[\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\frac{1}{2}\operatorname{Tr}\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j,22}\end{pmatrix}-\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right]\right\}\\
&=\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}[1+O(N^{-\kappa})].
\end{aligned}
$$

for some $\kappa>0$. Let us now move to removing $(j-1)$ from the $\mathbf{Q}$ matrix. To this end, use Cauchy interlacing as in (4.1) and $\|\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}\|_{\mathrm{op}}\lesssim\eta_{z,t}^{-2}$ to get $\|\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}-\mathbf{Q}_{a_j,b_j,\theta_j}\|_{\mathrm{op}}\lesssim\eta_{z,t}^{-D}$ for some $D>0$. Resolvent perturbation and Lemma 16 then give $\|[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1}-\mathbf{Q}^{-1}_{a_j,b_j,\theta_j}\|_{\mathrm{op}}\lesssim\eta_{z,t}^{-D}N^{-2}\lesssim N^{-2+D\epsilon_0}$. Therefore, we have

$$
\begin{aligned}
&\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}]^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}\\
&=\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}_{a_j,b_j,\theta_j}]^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}\left[1+O(N^{-\kappa})\right].
\end{aligned}
$$

A similar argument shows that $|\det\mathbf{Q}^{(j-1)}_{a_j,b_j,\theta_j}|^{-\frac{1}{2}}=|\det\mathbf{Q}_{a_j,b_j,\theta_j}|^{-\frac{1}{2}}[1+O(N^{-\kappa})]$. Ultimately,

$$
\begin{aligned}
&\int_{M_2^{sa}(\mathbb{R})}e^{i\frac{N}{2t}\operatorname{Tr}P}\det[I+i[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}(P\otimes I)[\tilde{H}^{(j-1)}_{a_j,b_j,\theta_j}]^{\frac{1}{2}}]^{-\frac{1}{2}}dP\\
&=O(\exp[-CN^{2\upsilon}])+\mathbf{C}_1|\det\mathbf{Q}_{a_j,b_j,\theta_j}|^{-\frac{1}{2}}\\
&\times\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^{*}[\mathbf{Q}_{a_j,b_j,\theta_j}]^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}\left[1+O(N^{-\kappa})\right].
\end{aligned}
$$

4.4 Putting it altogether

By our computations of the ratio of determinants and the dP integration, we have the following estimate locally uniformly in a_j, b_j:

$$\begin{aligned}
&\rho_{\mathrm{main}}(z,\mathbf{z};A) = \frac{16N^7 b^4[1+O(N^{-\kappa})]}{\pi^4 t^{12}\gamma_{z,t}^2\sigma_{z,t}^3\upsilon_{z,t}^2}\rho_{\mathrm{GinUE}}^{(2)}(z_1,z_2)\int_{\mathcal{I}_0^2} d\theta_1 d\theta_2 \frac{|\cos 2\theta_1||\cos 2\theta_2|}{\sin^2\theta_1\sin^2\theta_2}\\
&\times\prod_{j=1,2}\int_{V^2(\mathbb{R}^{N-2j+2})} d\mu_j(\mathbf{v}_j)|\det[V_j^* G_z^{(j-1)}(\eta_{z,t})V_j]|^j\\
&\times\exp\left\{-b_j^2(\tan\theta_j-\tan^{-1}\theta_j)^2\eta_{z,t}^2\operatorname{Tr}H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\right\}\\
&\times\mathbf{C}_1|\det\mathbf{Q}_{a_j,b_j,\theta_j}|^{-\frac12}\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^*\mathbf{Q}_{a_j,b_j,\theta_j}^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}\\
&+O(\exp[-CN^\kappa]).
\end{aligned}$$

By (2.47), Proposition 8, and our computation of the remaining 4×4 determinants (see Lemmas 14 and 15), we deduce

$$\begin{aligned}
&\rho_t(z,\mathbf{z};A) = \frac{16N^7 b^4[1+O(N^{-\kappa})]}{\pi^4 t^{12}\gamma_{z,t}^2\sigma_{z,t}^3\upsilon_{z,t}^2}\rho_{\mathrm{GinUE}}^{(2)}(z_1,z_2)\int_{\mathcal{I}_0^2} d\theta_1 d\theta_2 \frac{|\cos 2\theta_1||\cos 2\theta_2|}{\sin^2\theta_1\sin^2\theta_2}\\
&\times\prod_{j=1,2}|\det\mathcal{G}_{j,a,b}|^j\times\exp\left\{-b_j^2(\tan\theta_j-\tan^{-1}\theta_j)^2\eta_{z,t}^2\operatorname{Tr}H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\right\}\\
&\times\mathbf{C}_1|\det\mathbf{Q}_{a_j,b_j,\theta_j}|^{-\frac12}\exp\left\{-\mathbf{C}_2\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^*\mathbf{Q}_{a_j,b_j,\theta_j}^{-1}\begin{pmatrix}\frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\right\}\\
&+O(\exp[-CN^\kappa]),
\end{aligned}\tag{4.10}$$

where, to be totally explicit, $\mathcal{G}_{j,a,b}$ is the 4×4 matrix

$$\begin{aligned}
\mathcal{G}_{j,a,b} &:= \begin{pmatrix}\mathcal{G}_{j,a,b,11} & \mathcal{G}_{j,a,b,12}\\ \mathcal{G}_{j,a,b,21} & \mathcal{G}_{j,a,b,22}\end{pmatrix}\\
\mathcal{G}_{j,a,b,11} &:= \begin{pmatrix}i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(H_z\otimes E_{11}) & i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(H_z\otimes E_{12})\\ i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(H_z\otimes E_{21}) & i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(H_z\otimes E_{22})\end{pmatrix}\\
\mathcal{G}_{j,a,b,12} &:= \begin{pmatrix}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[H_z(A-z)\otimes E_{11}] & \frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[H_z(A-z)\otimes E_{12}]\\ \frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[H_z(A-z)\otimes E_{21}] & \frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[H_z(A-z)\otimes E_{22}]\end{pmatrix}\\
\mathcal{G}_{j,a,b,21} &:= \begin{pmatrix}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[(A-z)^*H_z\otimes E_{11}] & \frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[(A-z)^*H_z\otimes E_{12}]\\ \frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[(A-z)^*H_z\otimes E_{21}] & \frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}[(A-z)^*H_z\otimes E_{22}]\end{pmatrix}\\
\mathcal{G}_{j,a,b,22} &:= \begin{pmatrix}i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(\tilde{H}_z\otimes E_{11}) & i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(\tilde{H}_z\otimes E_{12})\\ i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(\tilde{H}_z\otimes E_{21}) & i\eta_{z,t}\frac{t}{N}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4}}(\tilde{H}_z\otimes E_{22})\end{pmatrix},\\
E_{11} &:= \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\quad E_{22} := \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\quad E_{12} := \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\quad E_{21} := \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},
\end{aligned}\tag{4.11}$$

and $\mathbf{Q}_{a_j,b_j,\theta_j}$ is given by

$$
\begin{aligned}
\mathbf{Q}_{a_j,b_j,\theta_j} &:= \left([\mathbf{Q}_{a_j,b_j,\theta_j}]_1 \quad [\mathbf{Q}_{a_j,b_j,\theta_j}]_2 \quad [\mathbf{Q}_{a_j,b_j,\theta_j}]_3\right), \\
[\mathbf{Q}_{a_j,b_j,\theta_j}]_1 &= \begin{pmatrix} \mathrm{Tr}\, \tilde{H}^2_{a_j,b_j,\theta_j,11} \\ 2\mathrm{Tr}\, \tilde{H}_{a_j,b_j,\theta_j,11}\tilde{H}_{a_j,b_j,\theta_j,12} \\ \mathrm{Tr}\, \tilde{H}_{a_j,b_j,\theta_j,12}\tilde{H}^T_{a_j,b_j,\theta_j,12} \end{pmatrix} \\
[\mathbf{Q}_{a_j,b_j,\theta_j}]_2 &= \begin{pmatrix} 2\mathrm{Tr}\, \tilde{H}_{a_j,b_j,\theta_j,11}\tilde{H}_{a_j,b_j,\theta_j,12} \\ 2\mathrm{Tr}\, \tilde{H}_{a_j,b_j,\theta_j,11}\tilde{H}_{a_j,b_j,\theta_j,22} + 2\mathrm{Tr}\, \tilde{H}^2_{a_j,b_j,\theta_j,12} \\ 2\mathrm{Tr}\, \tilde{H}_{a_j,b_j,\theta_j,22}\tilde{H}_{a_j,b_j,\theta_j,12} \end{pmatrix} \\
[\mathbf{Q}_{a_j,b_j,\theta_j}]_3 &= \begin{pmatrix} \mathrm{Tr}\, \tilde{H}_{a_j,b_j,\theta_j,12}\tilde{H}^T_{a_j,b_j,\theta_j,12} \\ 2\mathrm{Tr}\, \tilde{H}_{a_j,b_j,\theta_j,22}\tilde{H}_{a_j,b_j,\theta_j,12} \\ \mathrm{Tr}\, \tilde{H}^2_{a_j,b_j,\theta_j,22} \end{pmatrix}
\end{aligned}
$$

5 Replacing resolvents in (4.10) by universal local law approximations

The RHS of (4.10) is $\rho^{(2)}_{\mathrm{GinUE}}(z_1, z_2)$ times a quantity that depends only on traces of resolvents. Thus, we can replace each such trace by its local approximation to obtain a quantity of the form $\Phi_{z,t}(z_1, z_2)$ as in Theorem 2.

5.1 Replacing the constants, e.g. $\eta_{z,t}$

We start with the following result, which approximates constants appearing in the introduction with universal quantities.

Lemma 17. *There exists a constant $\kappa > 0$, which is independent of $\epsilon_0 > 0$, as well as constants $\eta_{\mathrm{univ},t}, \alpha_{\mathrm{univ},z,t}, \beta_{\mathrm{univ},z,t}, \gamma_{\mathrm{univ},z,t}, \sigma_{\mathrm{univ},z,t}$, which do not depend on the distribution of the entries of A, such that with high probability. we have*

$$
\begin{aligned}
\eta_{z,t} &= \eta_{\mathrm{univ},z,t}[1 + O(N^{-\kappa})], \\
\alpha_{z,t} &= \alpha_{\mathrm{univ},z,t}[1 + O(N^{-\kappa})], \\
\beta_{z,t} &= \beta_{\mathrm{univ},z,t}[1 + O(N^{-\kappa})], \\
\gamma_{z,t} &= \gamma_{\mathrm{univ},z,t}[1 + O(N^{-\kappa})], \\
\sigma_{z,t} &= \sigma_{\mathrm{univ},z,t}[1 + O(N^{-\kappa})], \\
\upsilon_{z,t} &= \upsilon_{\mathrm{univ},z,t}[1 + O(N^{-\kappa})].
\end{aligned}
$$

Proof. We will give details for the $\eta_{z,t}$ bound. By definition, $t\langle H_z(\eta_{z,t})\rangle = -it\eta^{-1}_{z,t}\langle G_z(\eta_{z,t})E_{11}\rangle = 1$. Let $\eta_{\mathrm{univ},z,t}$ be such that $-it\eta^{-1}_{\mathrm{univ},z,t}\langle M_{a,b,\frac{\pi}{4}}(\eta_{\mathrm{univ},z,t})E_{11}\rangle = 1$, where $M_{a,b,\frac{\pi}{4}}(\eta)$ is the local law approximation to $G_z(\eta)$ from Lemma 26. Now, suppose $\eta_{\mathrm{univ},z,t} \asymp t$ (i.e. it is bounded above and below by a constant times t). We will show this shortly. In this case, by Lemma 26, we have the estimate $\langle G_z(\eta_{\mathrm{univ},z,t})E_{11} - M_{a,b,\frac{\pi}{4}}(\eta_{\mathrm{univ},z,t})E_{11}\rangle \prec N^{-1}\eta^{-2}_{\mathrm{univ},z,t}$, where $\prec$ means bounded above up to a factor of N^τ for any fixed $\tau > 0$ with high probability; see Definition 25. Thus, we have

$$
-it\eta^{-1}_{\mathrm{univ},z,t}\langle G_z(\eta_{\mathrm{univ},z,t})E_{11}\rangle = 1 + O_\prec(N^{-1}\eta^{-2}_{\mathrm{univ},z,t}),
$$

where, similarly, $O_\prec(N^{-1}\eta^{-2}_{\mathrm{univ},z,t]})$ means something whose absolute value is $\prec N^{-1}\eta^{-2}_{\mathrm{univ},z,t}$.

The LHS of the above display is $t\langle H_z(\eta_{\mathrm{univ},z,t})\rangle$, so $t(\langle H_z(\eta_{\mathrm{univ},z,t})\rangle - \langle H_z(\eta_{z,t})\rangle) \prec N^{-1}\eta_{\mathrm{univ},z,t}^{-2}$. Standard resolvent perturbation and the operator bound $\|H_z(\eta_{\mathrm{univ},z,t})\|_{\mathrm{op}} \gtrsim 1$ (which holds with high probability since $\|A\|_{\mathrm{op}} \lesssim 1$ with high probability) imply that

$$|t\langle H_z(\eta_{\mathrm{univ},z,t})\rangle - t\langle H_z(\eta_{z,t})\rangle| = \left|\frac{t(\eta_{z,t}^2 - \eta_{\mathrm{univ},z,t}^2)}{N}\mathrm{Tr}\, H_z(\eta_{\mathrm{univ},z,t})H_z(\eta_{z,t})\right| \gtrsim t^2|\eta_{z,t} - \eta_{\mathrm{univ},z,t}|.$$

We deduce that $|\eta_{z,t} - \eta_{\mathrm{univ},z,t}| \lesssim N^{-1}t^{-4} \lesssim N^{-\kappa}t$ for some $\kappa > 0$, since $t = N^{-\epsilon_0}$ with $\epsilon_0 > 0$ small. This proves the desired $\eta_{z,t}$ estimate, provided we can show that $\eta_{\mathrm{univ},z,t} \asymp t$. This follows by $\eta_{z,t} \asymp t$ and $|\eta_{z,t} - \eta_{\mathrm{univ},z,t}| \lesssim N^{-1}t^{-4} \lesssim N^{-\kappa}t$. (To avoid a circular argument, one can use a standard continuity argument. In particular, in this proof so far, replace $M_{a,b,\frac{\pi}{4}}$ by $sM_{a,b,\frac{\pi}{4}} + (1-s)G_z$, and use continuity in $s \in [0,1]$ to show that $\eta_{\mathrm{univ},z,t} \asymp t$.) The other estimates follow similarly, though $\alpha_{\mathrm{univ},z,t}$, for example, is defined instead using the two-term local law in Lemma 29. $\square$

5.2 Replacing $|\det \mathcal{G}_{j,a,b}|$

The matrix $\mathcal{G}_{j,a,b}$ has size 4×4, and its entries are given by normalized traces of products of two resolvents. Thus, Lemma 29 will allow us to replace its determinant by that of a matrix which does not depend on the distribution of the entries of A.

Lemma 18. *There exists a 4×4 matrix $\mathcal{G}_{\mathrm{univ}}$ such that $|\det \mathcal{G}_{j,a,b}| = |\det \mathcal{G}_{\mathrm{univ}}|[1+O(N^{-\kappa})]$, where $\kappa > 0$ and $\mathcal{G}_{\mathrm{univ}}$ has entries independent of the distribution of the entries of A.*

Proof. It suffices to show that for any matrix entry indices k, ℓ, we have $(\mathcal{G}_{j,a,b})_{k\ell} = (\mathcal{G}_{\mathrm{univ}})_{k\ell}[1 + O(N^{-\kappa})]$, where $(\mathcal{G}_{\mathrm{univ}})_{k\ell}$ denotes a quantity which does not depend on distribution of the entries of A. Indeed, the determinant is a polynomial in the entries, and the entries of $\mathcal{G}_{j,a,b}$ are easily checked to be $\lesssim \eta_{z,t}^{-4} \lesssim N^{4\epsilon_0}$, and we can choose $\epsilon_0 > 0$ sufficiently small. We will choose $k, \ell = 1$; the other choices of k, ℓ follow by similar arguments, since each matrix that is tensored with $E_{11}, E_{12}, E_{21}, E_{22}$ in $\mathcal{G}_{j,a,b}$ is a block in $G_z(\eta_{z,t})$. In particular, we want to show that

$$i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}_{a,b,\frac{\pi}{4}}(\eta_{z,t})(H_z(\eta_{z,t}) \otimes E_{11})$$

is equal to a universal quantity times an error $1 + O(N^{-\kappa})$. First, we replace $\eta_{z,t}$ by $\eta_{\mathrm{univ},z,t}$ from Lemma 17. Since $\tilde{H}_{a,b,\frac{\pi}{4}}$ and H_z have operator norms $\lesssim \eta_{z,t}^{-2}$, we have

$$\begin{aligned}
&i\eta_{z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}_{a,b,\frac{\pi}{4}}(\eta_{z,t})(H_z(\eta_{z,t}) \otimes E_{11}) \\
&= i\eta_{\mathrm{univ},z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}_{a,b,\frac{\pi}{4}}(\eta_{\mathrm{univ},z,t})(H_z(\eta_{\mathrm{univ},z,t}) \otimes E_{11}) + O(N^{-\kappa}\eta_{z,t}^{-D})
\end{aligned}$$

for some $D > 0$. Since $\tilde{H}_{a,b,\frac{\pi}{4}} \geq C$ and $H_z \geq C$ with high probability (it is the inverse of a covariance matrix of a shift of A), and because $\eta_{\mathrm{univ},z,t}t \asymp N^{-2\epsilon_0}$ (see the proof of Lemma 17), the RHS of the previous identity is equal to

$$i\eta_{\mathrm{univ},z,t}\frac{t}{N}\mathrm{Tr}\,\tilde{H}_{a,b,\frac{\pi}{4}}(\eta_{\mathrm{univ},z,t})(H_z(\eta_{\mathrm{univ},z,t}) \otimes E_{11})[1 + +O(N^{-\kappa})]$$

for possibly different $\kappa > 0$. Now, apply Lemma 29 for $\theta = \frac{\pi}{4}$; this shows that the first term on the RHS of the previous display is universal (i.e. independent of the distribution of the entries of A) plus an error of $O(N^{-\kappa})$. We now absorb this additive error of $O(N^{-\kappa})$ as a multiplicative error of $1 + O(N^{-\kappa})$ (for possibly different $\kappa > 0$) using the same argument that gave the previous display. This completes the proof. $\square$

5.3 Replacing $\operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})$

We move to the exponential factor in the second line in (4.10). The only term we must deal with here is $\eta_{z,t}^2$ times the trace of two resolvents. For this, we use Lemmas 17 and 29.

Lemma 19. *There exists a constant $\mathfrak{T} = \mathfrak{T}_{\lambda_j}$ that is independent of the distribution of the entries of A and such that the following holds for $\kappa > 0$ and for all $\theta_j \in \mathcal{I}_0$:*

$$\begin{aligned}&\exp\left\{-b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{z,t}^2 \operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\right\}\\ &= \exp\left\{-b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{\mathrm{univ},z,t}^2\mathfrak{T}\right\}\left[1 + O(N^{-\kappa})\right].\end{aligned}$$

Proof. We first have the decomposition

$$\begin{aligned}&b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{z,t}^2 \operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t}) \\ &= b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{\mathrm{univ},z,t}^2 \operatorname{Tr} H_{\lambda_j}(\eta_{\mathrm{univ},z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{\mathrm{univ},z,t}) \\ &+ b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2[\eta_{z,t}^2 - \eta_{\mathrm{univ},z,t}^2] \operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t}) \\ &+ b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{\mathrm{univ},z,t}^2\left[\operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t}) - \operatorname{Tr} H_{\lambda_j}(\eta_{\mathrm{univ},z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{\mathrm{univ},z,t})\right].\end{aligned} \tag{5.1}$$

We bound the last two lines in (5.1). We start with the third line. By Lemma 17, we know $\eta_{z,t}^2 - \eta_{\mathrm{univ},z,t}^2 \lesssim N^{-\kappa}$ for some $\kappa > 0$ with high probability. Moreover, $\|H_{\lambda_j}(\eta_{z,t})\|_{\mathrm{op}}, \|\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\|_{\mathrm{op}} \lesssim \eta_{z,t}^{-2}$. Finally, because $\theta_j \in \mathcal{I}_0$, we know $|\tan\theta_j - \tan^{-1}\theta_j|^2 \lesssim |\theta - \frac{\pi}{4}|^2 \lesssim N^{-1+2\tau}\eta_{z,t}^{-2}$ for small $\tau > 0$ depending on at most ϵ_0. Thus, the third line satisfies

$$\begin{aligned}b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2[\eta_{z,t}^2 - \eta_{\mathrm{univ},z,t}^2] \operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t}) &\lesssim N^{-\kappa}N^{-1+2\tau}\eta_{z,t}^{-2}N\eta_{z,t}^{-4}\\ &\lesssim N^{-\kappa+2\tau}\eta_{z,t}^{-6},\end{aligned}$$

which is $\lesssim N^{-\kappa/2}$ if we choose τ, ϵ_0 small enough. For the last line in (5.1), we have

$$\begin{aligned}&\left\|H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t}) - H_{\lambda_j}(\eta_{\mathrm{univ},z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{\mathrm{univ},z,t})\right\|_{\mathrm{op}}\\ &\lesssim \eta_{z,t}^{-2}\|H_{\lambda_j}(\eta_{z,t}) - H_{\lambda_j}(\eta_{\mathrm{univ},z,t})\|_{\mathrm{op}} + \eta_{z,t}^{-2}\|\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t}) - \tilde{H}_{\overline{\lambda}_j}(\eta_{\mathrm{univ},z,t})\|_{\mathrm{op}},\end{aligned}$$

and by resolvent perturbation, the last line is $\lesssim \eta_{z,t}^{-D}|\eta_{z,t} - \eta_{\mathrm{univ},z,t}| \lesssim N^{-\kappa}\eta_{z,t}^{-D}$ for some $\kappa > 0$ and $D = O(1)$. (The last bound follows by Lemma 17.) Using this and the bound $|\tan\theta_j - \tan^{-1}\theta_j|^2 \lesssim |\theta - \frac{\pi}{4}|^2 \lesssim N^{-1+2\tau}\eta_{z,t}^{-2}$ for small $\tau > 0$ and $\theta_j \in \mathcal{I}_0$, we deduce

$$\begin{aligned}&b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{\mathrm{univ},z,t}^2\left[\operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t}) - \operatorname{Tr} H_{\lambda_j}(\eta_{\mathrm{univ},z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{\mathrm{univ},z,t})\right]\\ &\lesssim N^{-1+2\tau}NN^{-\kappa}\eta_{z,t}^{-D} \lesssim N^{-\kappa+2\tau}\eta_{z,t}^{-D},\end{aligned}$$

which is $\lesssim N^{-\kappa/2}$ if $\tau, \epsilon_0 > 0$ are small enough. Thus, we have

$$\begin{aligned}&b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{z,t}^2 \operatorname{Tr} H_{\lambda_j}(\eta_{z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{z,t})\\ &= b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{\mathrm{univ},z,t}^2 \operatorname{Tr} H_{\lambda_j}(\eta_{\mathrm{univ},z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{\mathrm{univ},z,t}) + O(N^{-\kappa}),\end{aligned} \tag{5.2}$$

where $\kappa > 0$ is possibly different than earlier in this proof. Now, we can apply Lemma 29 to the trace in the second line in (5.2). This shows that for some $\mathfrak{T}$ as in the statement of this lemma, we have the following, where the last line uses $|\tan\theta_j - \tan^{-1}\theta_j|^2 \lesssim |\theta - \frac{\pi}{4}|^2 \lesssim N^{-1+2\tau}\eta_{z,t}^{-2}$ and holds if $\epsilon_0 > 0$ is small enough:

$$\begin{aligned}&b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{\mathrm{univ},z,t}^2 \operatorname{Tr} H_{\lambda_j}(\eta_{\mathrm{univ},z,t})\tilde{H}_{\overline{\lambda}_j}(\eta_{\mathrm{univ},z,t})\\ &= b_j^2(\tan\theta_j - \tan^{-1}\theta_j)^2\eta_{\mathrm{univ},z,t}^2\mathfrak{T} + O(|\tan\theta_j - \tan^{-1}\theta_j|^2\eta_{\mathrm{univ},z,t}^{-D}).\end{aligned}$$

If we exponentiate the previous two displays, the desired estimate follows. ∎

5.4 Replacing $\mathbf{Q}_{a_j,b_j,\theta_j}$

We now deal with the third line in (4.10). The first step is to replace both copies of $\mathbf{Q}_{a_j,b_j,\theta_j}$ therein by universal quantities. Since the entries of $\mathbf{Q}_{a_j,b_j,\theta_j}$ are given by traces of products of blocks of two resolvents, we can again use Lemma 29.

Lemma 20. *There exists a 3×3 matrix $\mathbf{Q}_{\mathrm{univ},\theta_j}$ such that $\mathbf{Q}_{\mathrm{univ},\theta_j}$ is independent of the distribution of the entries of A, and there exists $\kappa > 0$ such that the following estimates hold. First, we have $|\det \mathbf{Q}_{a_j,b_j,\theta_j}|^{-\frac{1}{2}} = |\det \mathbf{Q}_{\mathrm{univ},\theta_j}|^{-\frac{1}{2}}[1+O(N^{-\kappa})]$. We also have*

$$\exp\left\{-\mathbf{C}_2 \begin{pmatrix} \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix}^* \mathbf{Q}^{-1}_{a_j,b_j,\theta_j} \begin{pmatrix} \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix}\right\}$$
$$= \exp\left\{-\mathbf{C}_2 \begin{pmatrix} \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix}^* \mathbf{Q}^{-1}_{\mathrm{univ},\theta_j} \begin{pmatrix} \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix}\right\}\left[1+O(N^{-\kappa})\right].$$

Proof. We claim there exists a 3×3 matrix $\mathbf{Q}_{\mathrm{univ},\theta_j}$ independent of the distribution of the entries of A such that $\|\mathbf{Q}_{a_j,b_j,\theta_j} - \mathbf{Q}_{\mathrm{univ},\theta_j}\|_{\mathrm{op}} \lesssim N^{1-\kappa}\eta_{z,t}^{-D}$ with high probability for some $D = O(1)$. To prove this claim, we first show $\|\mathbf{Q}_{a_j,b_j,\theta_j} - \mathbf{Q}_{a_j,b_j,\theta_j}(\eta_{\mathrm{univ},z,t})\|_{\mathrm{op}} \lesssim N^{1-\kappa}\eta_{z,t}^{-D}$, where $\mathbf{Q}_{a_j,b_j,\theta_j}(\eta_{\mathrm{univ},z,t})$ is just $\mathbf{Q}_{a_j,b_j,\theta_j}$ but evaluating all resolvents at $\eta_{\mathrm{univ},z,t}$ instead of $\eta_{z,t}$. To see this, we first note that $\|\mathbf{Q}_{a_j,b_j,\theta_j} - \mathbf{Q}_{a_j,b_j,\theta_j}(\eta_{\mathrm{univ},z,t})\|_{\mathrm{op}} \lesssim \|\mathbf{Q}_{a_j,b_j,\theta_j} - \mathbf{Q}_{a_j,b_j,\theta_j}(\eta_{\mathrm{univ},z,t})\|_{\max}$ since matrices in question are 3×3. We now estimate the entries of $\mathbf{Q}_{a_j,b_j,\theta_j} - \mathbf{Q}_{a_j,b_j,\theta_j}(\eta_{\mathrm{univ},z,t})$. We illustrate one example; other entries are treated similarly. We are claiming that $\operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}(\eta_{z,t})^2 = \operatorname{Tr}\tilde{H}_{a_j,b_j,\theta_j,11}(\eta_{\mathrm{univ},z,t})^2 + O(N^{1-\kappa}\eta_{z,t}^{-D})$. This follows by Lemma 17, the bound $\|\tilde{H}_{a_j,b_j,\theta_j,11}\|_{\mathrm{op}} \lesssim \eta_{z,t}^{-2}$, and resolvent identities. The matrix $\mathbf{Q}_{\mathrm{univ},\theta_j}$ is then constructed by applying Lemma 29 to each entry in $\mathbf{Q}_{a_j,b_j,\theta_j}(\eta_{\mathrm{univ},z,t})$.

We now prove the two proposed estimates. For the first determinant estimate, we note that

$$\begin{aligned}|\det \mathbf{Q}_{a_j,b_j,\theta_j}|^{-\frac{1}{2}} &= |\det \mathbf{Q}_{\mathrm{univ},\theta_j}|^{-\frac{1}{2}}\left[\frac{|\det \mathbf{Q}_{\mathrm{univ},\theta_j}|}{|\det \mathbf{Q}_{a_j,b_j,\theta_j}|}\right]^{\frac{1}{2}} \\ &= |\det \mathbf{Q}_{\mathrm{univ},\theta_j}|^{-\frac{1}{2}}\left[1 + \frac{|\det \mathbf{Q}_{\mathrm{univ},\theta_j}| - |\det \mathbf{Q}_{a_j,b_j,\theta_j}|}{|\det \mathbf{Q}_{a_j,b_j,\theta_j}|}\right]^{\frac{1}{2}}.\end{aligned}$$

Since the entries in $\mathbf{Q}_{a_j,b_j,\theta_j}$ are $\lesssim N\eta_{z,t}^{-D}$ for some $D = O(1)$, the same must be true for $\mathbf{Q}_{\mathrm{univ},\theta_j}$ by the operator norm estimate $\|\mathbf{Q}_{a_j,b_j,\theta_j} - \mathbf{Q}_{\mathrm{univ},\theta_j}\|_{\mathrm{op}} \lesssim N^{1-\kappa}\eta_{z,t}^{-D}$. This same estimate combined with an elementary expansion of the determinant (as a cubic in the entries of these 3×3 matrices) then shows that $|\det \mathbf{Q}_{\mathrm{univ},\theta_j}| - |\det \mathbf{Q}_{a_j,b_j,\theta_j}| = O(N^{3-\kappa}\eta_{z,t}^{-D})$ for possibly different $D = O(1)$. On the other hand, Lemma 16 shows that $|\det \mathbf{Q}_{a_j,b_j,\theta_j}| \gtrsim N^3\eta_{z,t}^{-6}$. We deduce from this paragraph, the bound $\eta_{z,t} \asymp t = N^{-\epsilon_0}$ with $\epsilon_0 > 0$ sufficiently small, and the previous display that

$$|\det \mathbf{Q}_{a_j,b_j,\theta_j}|^{-\frac{1}{2}} = |\det \mathbf{Q}_{\mathrm{univ},\theta_j}|^{-\frac{1}{2}}\left[1+O(N^{-\kappa})\right].$$

This is the first desired estimate. We move to the exponential estimate. We must estimate the difference of inverses $\mathbf{Q}^{-1}_{a_j,b_j,\theta_j} - \mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}$. By resolvent identities, we have

$$\mathbf{Q}^{-1}_{a_j,b_j,\theta_j} - \mathbf{Q}^{-1}_{\mathrm{univ},\theta_j} = \mathbf{Q}^{-1}_{a_j,b_j,\theta_j}(\mathbf{Q}_{\mathrm{univ},\theta_j} - \mathbf{Q}_{a_j,b_j,\theta_j})\mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}.$$

Lemma 16 shows that $\mathbf{Q}_{a_j,b_j,\mathrm{univ}} \gtrsim N\eta_{z,t}^{-2}$. The operator bound $\|\mathbf{Q}_{a_j,b_j,\theta_j} - \mathbf{Q}_{\mathrm{univ},\theta_j}\|_{\mathrm{op}} \lesssim N^{1-\kappa}\eta_{z,t}^{-D}$, if we choose $\epsilon_0 > 0$ sufficiently small, then implies the same for $\mathbf{Q}_{\mathrm{univ},\theta_j}$. Thus, if we choose $\epsilon_0 > 0$ small enough, we get

$$\|\mathbf{Q}^{-1}_{a_j,b_j,\theta_j} - \mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}\|_{\mathrm{op}} \lesssim N^{-2}\eta_{z,t}^4 N^{1-\kappa}\eta_{z,t}^{-D} \lesssim N^{-1-\frac{\kappa}{2}}.$$

To derive the proposed exponential estimate, we use (4.9) and Cauchy interlacing to get

$$\left\| \begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix} \right\| \lesssim_\tau N^{\frac{1}{2}+\tau}\eta_{z,t}^{-D},$$

where $\tau > 0$ is as small as we want. (To be clear, $\lesssim_\tau$ means bounded above up to a constant depending on τ; in this case, the constant blows up as $\tau \to 0$.) Combining the previous two displays yields the quadratic form estimate

$$\begin{aligned} &\begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix}^* \mathbf{Q}_{a_j,b_j,\theta_j}^{-1} \begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix} \\ &= \begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix}^* \mathbf{Q}_{\mathrm{univ},\theta_j}^{-1} \begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix} + O(N^{-\kappa+2\tau}\eta_{z,t}^{-D}) \end{aligned}$$

for $\kappa > 0$ and $\tau > 0$ (which we can choose to be small) and $D = O(1)$. If we choose $\tau > 0$ and $\epsilon_0 > 0$ sufficiently small (depending only on κ), the error term in the second line of the previous display becomes $O(N^{-\delta})$ for some $\delta > 0$. Exponentiating the resulting estimate completes the proof. ☐

5.5 Replacing the vector in the quadratic form

The final step is to replace the traces in the vector in the third line (4.10) by universal quantities. To this end, we use Lemma 26.

Lemma 21. *There exists $\mathfrak{v} \in \mathbb{R}^3$ whose entries are independent of the distribution of the entries of A, and such that the following estimate holds for some $\kappa > 0$:*

$$\begin{aligned} &\exp\left\{ -\mathbf{C}_2 \begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix}^* \mathbf{Q}_{\mathrm{univ},\theta_j}^{-1} \begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix} \right\} \\ &= \exp\left\{ -\mathbf{C}_2 \mathfrak{v}^* \mathbf{Q}_{\mathrm{univ},\theta_j}^{-1} \mathfrak{v} \right\} \left[1 + O(N^{-\kappa})\right] \end{aligned}$$

Proof. We first claim there exists $\mathfrak{v} \in \mathbb{R}^3$ whose entries are independent of the distribution of the entries of A such that for some $\kappa > 0$ and $D = O(1)$, we have the following with high probability:

$$\left\| \begin{pmatrix} \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} \\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12} \\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22} \end{pmatrix} - \mathfrak{v} \right\| \lesssim N^{\frac{1}{2}-\kappa}. \tag{5.3}$$

We first note the following, where the last line follows by computations in Appendix B:

$$\begin{aligned} \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} &= \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\frac{\pi}{4},11} + \mathrm{Tr}\,[\tilde{H}_{a_j,b_j,\theta_j,11} - \tilde{H}_{a_j,b_j,\frac{\pi}{4},11}] \\ &= \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\frac{\pi}{4},11} + \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}[A^*_{a_j,b_j,\frac{\pi}{4}}A_{a_j,b_j,\frac{\pi}{4}} - A^*_{a_j,b_j,\theta_j}A_{a_j,b_j,\theta_j}]\tilde{H}_{a_j,b_j,\frac{\pi}{4},11} \\ &= \frac{N}{t} + \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}[A^*_{a_j,b_j,\frac{\pi}{4}}A_{a_j,b_j,\frac{\pi}{4}} - A^*_{a_j,b_j,\theta_j}A_{a_j,b_j,\theta_j}]\tilde{H}_{a_j,b_j,\frac{\pi}{4},11}. \end{aligned}$$

Since $|\theta_j - \frac{\pi}{4}| \leq N^{-1/2+\tau}$, we know $\|A^*_{a_j,b_j,\frac{\pi}{4}}A_{a_j,b_j,\frac{\pi}{4}} - A^*_{a_j,b_j,\theta_j}A_{a_j,b_j,\theta_j}\|_{\mathrm{op}} \lesssim N^{-1/2+\tau}$, where $\tau > 0$ is small. In particular, by resolvent perturbation, Lemma 17, and trivial resolvent bounds, we have

$$\begin{aligned} &\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}[A^*_{a_j,b_j,\frac{\pi}{4}}A_{a_j,b_j,\frac{\pi}{4}} - A^*_{a_j,b_j,\theta_j}A_{a_j,b_j,\theta_j}]\tilde{H}_{a_j,b_j,\frac{\pi}{4},11} \\ &= \mathrm{Tr}\,\left\{\tilde{H}_{a_j,b_j,\theta_j,11}(\eta_{\mathrm{univ},z,t})[A^*_{a_j,b_j,\frac{\pi}{4}}A_{a_j,b_j,\frac{\pi}{4}} - A^*_{a_j,b_j,\theta_j}A_{a_j,b_j,\theta_j}]\tilde{H}_{a_j,b_j,\frac{\pi}{4},11}(\eta_{\mathrm{univ},z,t})\right\} + O(N^{1/2-\kappa}), \end{aligned}$$

where $\kappa > 0$ is fixed and independent of our choice of small $\epsilon_0 > 0$. By another resolvent identity, we have

$$\begin{aligned}
&\mathrm{Tr}\left\{\tilde{H}_{a_j,b_j,\theta_j,11}(\eta_{\mathrm{univ},z,t})[A^*_{a_j,b_j,\frac{\pi}{4}}A_{a_j,b_j,\frac{\pi}{4}} - A^*_{a_j,b_j,\theta_j}A_{a_j,b_j,\theta_j}]\tilde{H}_{a_j,b_j,\frac{\pi}{4},11}(\eta_{\mathrm{univ},z,t})\right\}\\
&= \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}(\eta_{\mathrm{univ},z,t}) - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\frac{\pi}{4},11}(\eta_{\mathrm{univ},z,t}).
\end{aligned}$$

In particular, the previous three displays imply

$$\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11} = \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\frac{\pi}{4},11}(\eta_{\mathrm{univ},z,t}) - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}(\eta_{\mathrm{univ},z,t}).$$

The RHS of the previous display is equal to a quantity which does not depend on the distribution of the entries of A plus $O(\eta_{z,t}^{-D})$ for some $D = O(1)$; this is by the local law in Lemma 26. The aforementioned "universal" term is our choice of $\mathfrak{v}_1$, the first entry of $\mathfrak{v}$. The rest of $\mathfrak{v}$ is constructed using a similar argument. We now compute

$$\begin{aligned}
&\begin{pmatrix}\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^* \mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}\begin{pmatrix}\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix} - \mathfrak{v}^*\mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}\mathfrak{v} \qquad (5.4)\\
&= \left[\begin{pmatrix}\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix} - \mathfrak{v}\right]^* \mathbf{Q}^{-1}_{\mathrm{univ},j}\begin{pmatrix}\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}\\
&+ \mathfrak{v}^*\mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}\left[\begin{pmatrix}\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix} - \mathfrak{v}\right].
\end{aligned}$$

By (4.9), the last vector in the second line of (5.4) has norm $\lesssim N^{\frac12+\tau}\eta_{z,t}^{-D}$ for $\tau > 0$ small and $D = O(1)$. By (5.3), the same is true of $\mathfrak{v}$. Moreover, as shown in the proof of Lemma 20, we know that $\mathbf{Q}^{-1}_{\mathrm{univ},j} \lesssim N^{-1}\eta_{z,t}^2$ with high probability. So, the second line in (5.4) is $\lesssim N^{\frac12+\tau}N^{\frac12-\kappa}\eta_{z,t}^{-D}N^{-1}\eta_{z,t}^2 \lesssim N^{-\kappa+\tau}\eta_{z,t}^{-D}$ for possibly different $D = O(1)$ and for small $\tau > 0$. By the same token, the same is true for the last line in (5.4). Thus, we get

$$\begin{pmatrix}\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix}^* \mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}\begin{pmatrix}\frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,11}\\ -2\mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,12}\\ \frac{N}{t} - \mathrm{Tr}\,\tilde{H}_{a_j,b_j,\theta_j,22}\end{pmatrix} - \mathfrak{v}^*\mathbf{Q}^{-1}_{\mathrm{univ},\theta_j}\mathfrak{v} = O(N^{-\kappa+\tau}\eta_{z,t}^{-D}).$$

It now suffices to choose $\tau, \epsilon_0 > 0$ small enough to make the RHS $O(N^{-\kappa/2})$, at which point we then exponentiate the resulting bound to conclude the proof. ☐

6 Proof of Theorem 2

Combining (4.10) with Lemmas 17, 18, 19, 20, and 21 shows that

$$\rho_t(z, \mathbf{z}; A) = \Phi_{z,t}(z_1, z_2)\rho^{(2)}_{\mathrm{GinUE}}(z_1, z_2)\left[1 + O(N^{-\kappa})\right] + O(\exp[-CN^{\kappa}]), \qquad (6.1)$$

where $\kappa > 0$ and $\Phi_{z,t}(z_1, z_2)$ is as in the statement of Theorem 2. Recall $\rho_t(z, \mathbf{z}; A)$ is the two-point correlation function for local eigenvalue statistics near z; more precisely, we have the following (in which we recall $\mathbf{z} = (z_1, z_2)$):

$$\int_{\mathbb{C}^2} O(z_1, z_2)\rho_t(z, \mathbf{z}; A)dz_1dz_2 = \mathbb{E}\left[\sum_{i_1 \neq i_2} O(N^{\frac12}\sigma_{z,t}^{\frac12}[z - \lambda_{i_1}(t)], N^{\frac12}\sigma_{z,t}^{\frac12}[z - \lambda_{i_2}(t)])\right].$$

Above, $O(z_1, z_2) \in C_c^\infty(\mathbb{C}^2)$ is arbitrary. Since $\rho_t(z, \mathbf{z}; A)$ is a probability density (up to a scaling by a deterministic, $O(1)$ factor) with respect to $dz_1 dz_2$, the estimate (6.1) gives

$$\int_{\mathbb{C}^2} O(z_1, z_2)\rho_t(z, \mathbf{z}; A)dz_1 dz_2 = \int_{\mathbb{C}^2} O(z_1, z_2)\rho_t(z, \mathbf{z}; A)[1 + O(N^{-\kappa})]dz_1 dz_2 + O(N^{-\kappa})$$
$$= \int_{\mathbb{C}^2} O(z_1, z_2)\Phi_{z,t}(z_1, z_2)\rho^{(2)}_{\mathrm{GinUE}}(z_1, z_2)dz_1 dz_2 + O(N^{-\kappa})$$

for some $\kappa > 0$. It now suffices to combine the previous two displays. □

7 Proof of Theorem 1

It suffices to prove Theorem 1 for $\tilde{A}_t := A\sqrt{1+t}$, where $t = N^{-\epsilon_0}$ as in Theorem 2. Indeed, let $\rho_{A,z}(z_1, z_2)$ be the two-point correlation function for eigenvalues of A near z, and $\rho_{\tilde{A}_t,z}(z_1, z_2)$ is the same but for $\tilde{A}_t$. Then, for any $O \in C_c^\infty(\mathbb{C}^2)$ we have

$$\begin{aligned}\int_{\mathbb{C}^2} O(z_1, z_2)\rho_{A,z}(z_1, z_2)dz_1 dz_2 &= \frac{1}{(1+t)^2}\int_{\mathbb{C}^2} O\left(\frac{z_1}{\sqrt{1+t}}, \frac{z_1}{\sqrt{1+t}}\right)\rho_{\tilde{A}_t,z}(z_1, z_2)dz_1 dz_2 \\ &= \int_{\mathbb{C}^2} O\left(\frac{z_1}{\sqrt{1+t}}, \frac{z_1}{\sqrt{1+t}}\right)\rho_{\tilde{A}_t,z}(z_1, z_2)dz_1 dz_2 + O(t) \\ &= \int_{\mathbb{C}^2} O(z_1, z_2)\rho_{\tilde{A}_t,z}(z_1, z_2)dz_1 dz_2 + O(t)\end{aligned}$$

by change of variables on $\mathbb{C}^2$. Since $t = N^{-\epsilon_0} \to 0$, we have reduced to proving Theorem 1 for $\tilde{A}_t$ instead of A. To this end, the main ingredient we require is the following "three-and-a-half moment matching" theorem, which we state in more generality. Before we state this theorem, we first say that $X, \tilde{X}$ match up to three and a half moments if:

1. X is an $N \times N$ matrix whose entries X_{ij} are real i.i.d. variables that satisfy $\mathbb{E}|X_{ij}|^p \lesssim_p N^{-p/2}$ for $1 \leq p \leq 4$.
2. $\tilde{X}$ is an $N \times N$ matrix whose entries $\tilde{X}_{ij}$ satisfy the same properties.
3. $\mathbb{E}X_{ij}^p = \mathbb{E}\tilde{X}_{ij}^p$ for $p = 1, 2, 3$, and $\mathbb{E}|X_{ij}|^4 = \mathbb{E}|\tilde{X}_{ij}|^4 + O(N^{-2-\delta})$ for some $\delta > 0$. (Note that $N^{-2-\delta}$ is the below the natural scale of N^{-2} for fourth moments.)

Lemma 22. *Suppose that X and $\tilde{X}$ are $N \times N$ matrices that match up to three and a half moments. Let $\{\lambda_i^X\}_i$ and $\{\lambda^{\tilde{X}}\}_i$ be eigenvalues of X and $\tilde{X}$, respectively. Fix any $k \geq 1$ and $z \in \mathbb{C}$ such that $|z| \leq 1 - \tau$ and $\mathrm{Im}(z) \geq \tau$ for some $\tau > 0$ fixed. For any $O \in C_c^\infty(\mathbb{C}^k)$, we have*

$$\mathbb{E}\left\{\sum_{i_1 \neq i_2 \neq \ldots \neq i_k}\left[O(N^{\frac{1}{2}}[z - \lambda_{i_1}^X], \ldots, N^{\frac{1}{2}}[z - \lambda_{i_k}^X]) - O(N^{\frac{1}{2}}[z - \lambda_{i_1}^{\tilde{X}}], \ldots, N^{\frac{1}{2}}[z - \lambda_{i_k}^{\tilde{X}}])\right]\right\} \to_{N\to\infty} 0.$$

Proof of Theorem 1 given Lemma 22. Let $t = N^{-\epsilon_0}$ as in Theorem 2. By Lemma 3.4 in [18], we can find $N \times N$ matrices $\mathbf{A}$ and B such that B is real Ginibre (entries are independent $N(0, N^{-1})$ random variables), such that $\mathbf{A}$ and B are independent, such that $\mathbf{A} + \sqrt{t}B$ and $\tilde{A}_t$ match up to three and a half moments, and such that $\mathbb{E}|\mathbf{A}_{ij}|^p \lesssim_p N^{-p/2}$ for all $p \geq 1$. (Indeed, for this last property, note that it is true for $\tilde{A}_t$ by assumption and for B.) We necessarily have $\mathbb{E}|\mathbf{A}_{ij}|^2 = N^{-1}$ (again, since $\mathbb{E}|\tilde{A}_t|^2 = (1+t)N^{-1}$ and $\mathbb{E}|B_{ij}|^2 = N^{-1}$. Ultimately, by Lemma 22 (for the test function O precomposed with scaling by a universal $O(1)$ constant), it suffices to prove Theorem

1 for $\mathbf{A} + \sqrt{t}B$ instead of for $\tilde{A}_t$. This amounts to replacing $\sigma_{z,t}$ in Theorem 1 by its universal approximation $\sigma_{\mathrm{univ},z,t}$ from Lemma 17, and then using Theorem 2. We give the details below.

Let $\{\lambda_j^{\mathbf{A}+\sqrt{t}B}\}_j$ denote the eigenvalues of $\mathbf{A} + \sqrt{t}B$. Fix $z \in \mathbb{C}$ as in the statement of Theorem 1, and let $\sigma_{z,t}$ be defined as in the introduction but for $\mathbf{A}$ and its resolvents. We first claim that

$$\begin{aligned} &\mathbb{E}\left[\sum_{i_1 \neq i_2} O(N^{\frac{1}{2}}\sigma_{z,t}^{\frac{1}{2}}[z-\lambda_{i_1}], N^{\frac{1}{2}}\sigma_{z,t}^{\frac{1}{2}}[z-\lambda_{i_1}])\right] \\ &= \mathbb{E}\left[\sum_{i_1 \neq i_2} O(N^{\frac{1}{2}}\sigma_{\mathrm{univ},z,t}^{\frac{1}{2}}[z-\lambda_{i_1}], N^{\frac{1}{2}}\sigma_{\mathrm{univ},z,t}^{\frac{1}{2}}[z-\lambda_{i_1}])\right] + o(1), \end{aligned} \tag{7.1}$$

where $\sigma_{\mathrm{univ},z,t}$ is from Lemma 17. Indeed, with high probability, we know that $\sigma_{z,t} = \sigma_{\mathrm{univ},z,t}[1 + O(N^{-\kappa})]$ for some $\kappa > 0$. If $\rho_{\mathbf{A}+\sqrt{t}B,z}(z_1, z_2)$ denotes the two-point correlation function for $\mathbf{A}+\sqrt{t}B$ near z, then

$$\begin{aligned} &\mathbb{E}\left[\sum_{i_1 \neq i_2} O(N^{\frac{1}{2}}\sigma_{\mathrm{univ},z,t}^{\frac{1}{2}}[z-\lambda_{i_1}], N^{\frac{1}{2}}\sigma_{\mathrm{univ},z,t}^{\frac{1}{2}}[z-\lambda_{i_1}])\right] \\ &= \int_{\mathbb{C}^2} O\left(\sigma_{\mathrm{univ},z,t}^{\frac{1}{2}}\sigma_{z,t}^{-\frac{1}{2}} z_1, \sigma_{\mathrm{univ},z,t}^{\frac{1}{2}}\sigma_{z,t}^{-\frac{1}{2}} z_2\right) \frac{1}{N^2\sigma_{z,t}^2}\rho_{\mathbf{A}+\sqrt{t}B}\left(z + \frac{z_1}{\sqrt{N\sigma_{z,t}}}, z + \frac{z_2}{\sqrt{N\sigma_{z,t}}}\right) dz_1 dz_2, \end{aligned}$$

at which point (7.1) follows by $\sigma_{\mathrm{univ},z,t}\sigma_{z,t}^{-1} = 1 + O(N^{-\kappa})$ and an elementary Taylor expansion of O. To conclude the proof, we now combine (7.1) with Theorem 2 (applied to $\mathbf{A}$). $\square$

7.1 Proof of Lemma 22

By a standard approximation procedure, it suffices to prove Lemma 22 by functions of the form $O(z_1, \ldots, z_k) = f^{(1)}(z_1) \ldots f^{(k)}(z_k)$. The idea behind the following calculation is to use Girko's Hermitization formula to bring Lemma 22 into the realm of real symmetric matrices; this computation was also explored in [10]. Before we proceed, we introduce some notation. Define

$$\mathsf{G}_z(\eta) := \begin{pmatrix} -i\eta & X - z \\ X^* - \overline{z} & -i\eta \end{pmatrix}, \quad \tilde{\mathsf{G}}_z(\eta) := \begin{pmatrix} -i\eta & X - z \\ \tilde{X}^* - \overline{z} & -i\eta \end{pmatrix}.$$

We use different font to distinguish from G_z, which was meant for A earlier in this paper. We also define $m^z(w)$ to be the unique solution to

$$-\frac{1}{m^z(w)} = w + m^z(w) - \frac{|z|^2}{w + m^z(w)}, \quad \mathrm{Im}[m^z(w)]\mathrm{Im}[w] > 0. \tag{7.2}$$

By the inclusion-exclusion principle, it suffices to show that if σ_i and $\tilde{\sigma}_i$ denote eigenvalues of X and $\tilde{X}$, respectively, we have

$$\mathbf{E}\prod_{j=1}^{k}\left(\frac{1}{N}\sum_{i=1}^{N} f_{z_j}^{(j)}(\sigma_i) - \frac{1}{\pi}\int_{\mathbf{D}} f_{z_j}^{(j)}(z)d^2z\right) = \mathbf{E}\prod_{j=1}^{k}\left(\frac{1}{N}\sum_{i=1}^{N} f_{z_j}^{(j)}(\tilde{\sigma}_i) - \frac{1}{\pi}\int_{\mathbf{D}} f_{z_j}^{(j)}(z)d^2z\right) + O(N^{-c})$$

where we introduced the rescaled test functions

$$f_{z_j}^{(j)}(z) := Nf^{(j)}\left(\sqrt{N}(z - z_j)\right), \quad z \in \mathbf{C},$$

and the implicit constant in $O(\cdot)$ depends on $\|\Delta f^{(j)}\|_{L^1(\mathbb{C})}$ for $j=1,\ldots,k$. We adjusted notation slightly to fit that of [10]. By Theorem 2.4 in [10], we have

$$\left(\frac{1}{N}\sum_{i=1}^{N} f_{z_j}^{(j)}(\sigma_i) - \frac{1}{\pi}\int_{\mathbf{D}} f_{z_j}^{(j)}(z)d^2z\right) = \mathcal{I}_\epsilon(X, f_{z_j}^{(j)}) + \mathcal{E}_\epsilon$$

where $\epsilon>0$ is arbitrary, where $\mathbb{E}|\mathcal{E}_\epsilon| \lesssim_\tau N^{-\frac{1}{4}\epsilon}\|\Delta f^{(j)}\|_{L^1(\mathbb{C})}$, and where

$$\mathcal{I}_\epsilon\left(X, f_{z_j}^{(j)}\right) := \frac{1}{2\pi}\int_{\mathbf{C}} \Delta f_{z_j}^{(j)}(z)\int_{N^{-1-\epsilon}}^{N^{-1+\epsilon}} \langle \mathrm{Im}\mathsf{G}_z(\eta) - \mathrm{Im} m^z(i\eta)\rangle\, d\eta d^2z.$$

(Recall that $\langle\rangle$ denotes normalized trace.) The same computations and estimates hold for $\tilde{X}$ in place of X. Thus, the proof of Lemma 22 reduces to proving the following. (In Lemma 23 below, the constant $\delta>0$ is the exponent in the moment matching $\mathbb{E}X_{ij}^p = \mathbb{E}\tilde{X}_{ij}^p$ for $p=1,2,3$, and $\mathbb{E}|X_{ij}|^4 = \mathbb{E}|\tilde{X}_{ij}|^4 + O(N^{-2-\delta})$.)

Lemma 23. *Take the assumptions in Lemma 22. Fix any $k\geq 1$. There exists small $\epsilon=\epsilon(k,\delta)>0$ such that uniformly over $N^{-1-\epsilon}\leq \eta_\ell \leq N^{-1+\epsilon}$ for $\ell=1,\ldots,k$, we have*

$$\mathbb{E}\prod_{l=1}^{k}\langle \mathrm{Im}\mathsf{G}_z(i\eta_\ell) - \mathrm{Im} m^{z_l}(i\eta_\ell)\rangle - \mathbb{E}\prod_{l=1}^{k}\left\langle \mathrm{Im}\widetilde{\mathsf{G}}_{z_\ell}(i\eta_\ell) - \mathrm{Im} m^{z_\ell}(i\eta_\ell)\right\rangle = \mathcal{O}\left(N^{-c\delta}\right). \tag{7.3}$$

The proof of Lemma 23 uses a standard Green's function comparison argument. We adopt the continuous comparison method introduced in [22], which is based on the following construction.

Definition 24 (Interpolating matrices)**.** *Define the following matrices:*

$$\mathcal{H}^0 := \widetilde{\mathcal{H}} = \begin{pmatrix} 0 & \widetilde{X} \\ \widetilde{X}^T & 0 \end{pmatrix}, \quad \mathcal{H}^1 := \mathcal{H} = \begin{pmatrix} 0 & X \\ X^T & 0 \end{pmatrix}$$

Let ρ_{ij}^0 and ρ_{ij}^1 denote the laws of $\widetilde{\mathcal{H}}_{ij}$ and $\mathcal{H}_{ij}$, respectively. For $\theta\in[0,1]$, we define the interpolated laws $\rho_{ij}^\theta := (1-\theta)\rho_{ij}^0 + \theta\rho_{ij}^1$. Let $\{\mathcal{H}^\theta : \theta\in(0,1)\}$ be a collection of random matrices that satisfy the following properties. For any fixed $\theta\in(0,1)$, the triple $(\mathcal{H}^0,\mathcal{H}^\theta,\mathcal{H}^1)$ of $2N\times 2N$ random matrices are jointly independent, and the matrix $\mathcal{H}^\theta = (\mathcal{H}_{ij}^\theta)$ has law

$$\prod_{i\leq j}\rho_{ij}^\theta\left(\mathrm{d}\mathcal{H}_{ij}^\theta\right)$$

(We do not require $\mathcal{H}^{\theta_1}$ to be independent of $\mathcal{H}^{\theta_2}$ for $\theta_1\neq\theta_2\in(0,1)$.) For $\lambda\in\mathbb{C}$ and indices i,j, we define the matrix $\mathcal{H}_{(ij)}^{\theta,\lambda}$ as

$$\left(\mathcal{H}_{(ij)}^{\theta,\lambda}\right)_{kl} := \begin{cases} \mathcal{H}_{ij}^\theta, & \text{if } \{k,l\}\neq\{i,j\}, \\ \lambda, & \text{if } \{k,l\}=\{i,j\}. \end{cases}$$

Correspondingly, we define the resolvents

$$\mathsf{G}_\ell^\theta := \left(\mathcal{H}^\theta - \begin{pmatrix} 0 & z_\ell \\ z_\ell^* & 0 \end{pmatrix} - i\eta_\ell\right)^{-1}, \quad \mathsf{G}_{\ell,(ij)}^{\theta,\lambda}(z) := \left(\mathcal{H}_{(ij)}^{\theta,\lambda} - \begin{pmatrix} 0 & z_\ell \\ z_\ell^* & 0 \end{pmatrix} - i\eta_\ell\right)^{-1}.$$

For any function $F:\mathbb{R}^{2N\times 2N}\to\mathbb{C}$, we have the basic interpolation formula

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbb{E}F\left(\mathcal{H}^\theta\right) = \sum_{i,j}\left[\mathbb{E}F\left(\mathcal{H}_{(ij)}^{\theta,\mathcal{H}_{ij}^1}\right) - \mathbb{E}F\left(\mathcal{H}_{(ij)}^{\theta,\mathcal{H}_{ij}^0}\right)\right] \tag{7.4}$$

provided all the expectations exist.

Proof of Lemma 23. We will use (7.4) with the choice

$$F\left(\mathcal{H}^{\theta}\right)=\prod_{l=1}^{k}\left\langle \mathrm{Im}\mathsf{G}_{\ell}^{\theta}-\mathrm{Im}m^{z_\ell}\right\rangle \tag{7.5}$$

We omit θ from the notation m^{z_ℓ} since the entries of the $\mathcal{H}^{\theta}$ have the same variances for all θ. Hence m_ℓ^{θ} is independent of θ. In the remaining, we will prove the following bound for our choice of F in (7.5), which implies Lemma 23: for all $\theta\in[0,1]$ and $1\le i,j\le 2N$,

$$\mathbb{E}F\left(\mathcal{H}_{(i,j)}^{\theta,\mathcal{H}_{ij}^{1}}\right)-\mathbb{E}F\left(\mathcal{H}_{(i,j)}^{\theta,\mathcal{H}_{ij}^{0}}\right)\prec N^{-2+C\,\epsilon-\delta}. \tag{7.6}$$

Combining (7.4) and (7.6), if we choose $\epsilon>0$ small enough, then we deduce Lemma 23. In particular, are left to prove (7.6). By Theorem 5.2 in [4] (see also [10]), we have the following for any $\tau>0$ with very high probability (i.e. probability at least $1-O(N^{-D})$ for any $D=O(1)$):

$$\|\mathsf{G}_{\ell}^{\theta}-m^{z_\ell}\|_{\max}\lesssim N^{\epsilon+\tau}.$$

The max-norm is sup-norm over entries. By resolvent expansion, for any $\lambda,\ \lambda'\in\mathbb{R}$, we have the following (in which we drop subscript ℓ for simplicity):

$$\mathsf{G}_{(ij)}^{\theta,\lambda'}=\mathsf{G}_{(ij)}^{\theta,\lambda}+\sum_{k=1}^{K}\mathsf{G}_{(ij)}^{\theta,\lambda}\left\{\left[(\lambda-\lambda')\,\Delta_{ij}\right]\mathsf{G}_{(ij)}^{\theta,\lambda}\right\}^{k}+\mathsf{G}_{(ij)}^{\theta,\lambda'}\left\{\left[(\lambda-\lambda')\,\Delta_{ij}\right]\mathsf{G}_{(ij)}^{\theta,\lambda}\right\}^{K+1}. \tag{7.7}$$

Above, Δ_{ij} are $2N\times 2N$ matrices defined as

$$\left(\Delta_{ij}\right)_{kl}:=\delta_{ki}\delta_{lj}+\delta_{kj}\delta_{li}$$

Combining the above two statements and the trivial bound $\|\mathsf{G}_\ell\|\le(1/\eta_\ell)$, we get

$$\left\|\mathsf{G}_{l,(i,j)}^{\theta,0}-m^{z_\ell}\right\|_{\max}\lesssim N^{\epsilon+\tau}$$

for any $\tau>0$ with very high probability. Now, use (7.7) with $\lambda=0$ and $\lambda'=\mathcal{H}_{ij}^{\gamma}$ and $K=7$. In doing so, we use $\mathbb{E}|\mathcal{H}_{ij}^{\gamma}|^p\lesssim_p N^{-p/2}$. We also apply the local law (Lemma 26) to control the Green's function entries. Ultimately, we get the following in which we again drop the subscript ℓ:

$$\left[\mathsf{G}_{(ij)}^{\theta,H_{ij}^{\gamma}}\right]_{xy}=\left[\mathsf{G}_{(ij)}^{\theta,0}\right]_{xy}+\sum_{k=1}^{4}\mathcal{X}_{xy}^{(ij)}(\gamma,k)+\mathcal{E},\quad 1\le x,y\le 2N,\gamma\in\{0,1\}. \tag{7.8}$$

Above, $\mathcal{E}=O(N^{-5/2+\tau})$ with very high probability (for any fixed $\tau>0$), and

$$\mathcal{X}^{(ij)}(\gamma,k):=\left(-\mathcal{H}_{ij}^{\gamma}\right)^{k}\mathsf{G}_{(ij)}^{\theta,0}\left[\Delta_{ij}\mathsf{G}_{(ij)}^{\theta,0}\right]^{k},\qquad\text{with}\qquad\|\mathcal{X}(\gamma,k)\|_{\max}\prec N^{-k/2+C\epsilon}.$$

Notice that $\mathsf{G}_{(ij)}^{\theta,0}$ is independent of $\mathcal{H}_{ij}^{\gamma},\gamma\in\{0,1\}$. Moreover, note that

$$\mathbb{E}\left|\left(\mathsf{G}_{(ij)}^{\theta,0}\left[\Delta_{ij}\mathsf{G}_{(ij)}^{\theta,0}\right]^{k}\right)_{\mathfrak{a}\mathfrak{b}}\right|\lesssim N^{C\epsilon+\tau}$$

for any entry indices $\mathfrak{a},\mathfrak{b}$, since we have shown that the entries on the LHS are $\prec N^{C\epsilon+\tau}$ for any $\tau>0$, but they are also deterministically bounded by $O(\eta^{-D})$ for some $D=O(1)$, and $\eta\ge N^{-2}$. Thus, with the expansion (7.8), we can write

$$\mathbb{E}F\left(\mathcal{H}_{(ij)}^{\theta,X_{ij}^{\gamma}}\right)=\sum_{s}\sum_{k_1,k_2\cdots k_s}1\left(\sum_{t=1}^{s}k_t\le 4\right)\mathbb{E}[\mathcal{C}_{k_1,k_2\cdots k_s}]\prod_{t=1}^{s}\left(\mathbb{E}(\mathcal{H}_{ij}^{\gamma})^{k_t}\right)+O(N^{-5/2+C\epsilon})$$

where $\mathcal{C}_{k_1,k_2\cdots k_s} = O(N^{C\epsilon+\tau})$ for any $\tau > 0$, and $\mathcal{C}_{k_1,k_2\cdots k_s}$ does not depend on γ. We now use this for $\gamma = 0, 1$. The LHS of (7.6) is thus given by

$$\sum_s \sum_{k_1,k_2\cdots k_s} 1\left(\sum_{t=1}^s k_t \le 4\right) \mathcal{C}_{k_1,k_2\cdots k_s} \left[\prod_{t=1}^s \left(\mathbb{E}(\mathcal{H}_{ij}^1)^{k_t}\right) - \prod_{t=1}^s \left(\mathbb{E}(\mathcal{H}_{ij}^0)^{k_t}\right)\right] + O(N^{-5/2+C\epsilon}).$$

By the moment matching condition, the first term above is $O(N^{-2+C\epsilon+\delta})$, so (7.6) follows. □

A Resolvent estimates for $G_{a,b,\theta}(\eta)$

For $\kappa > 0$ let $\theta \in \left[\kappa, \frac{\pi}{2} - \kappa\right]$, $a \in (-1, 1)$, $b \in [\kappa, 1)$. Suppose $A \in M_N(\mathbb{R})$ is a random matrix with i.i.d., centered, variance $1/N$ entries. Define

$$G(\eta) = G_{a,b,\theta}(\eta) = \begin{pmatrix} -i\eta & I_2 \otimes A - \Lambda_{a,b,\theta} \otimes I_N \\ I_2 \otimes A^T - \Lambda_{a,b,\theta}^T \otimes I_N & -i\eta \end{pmatrix}^{-1},$$

where

$$\Lambda = \Lambda_{a,b,\theta} = \begin{pmatrix} a & b\tan\theta \\ -\frac{b}{\tan\theta} & a \end{pmatrix}$$

and let $w = a + ib$. We will also use the notation

$$W = \begin{pmatrix} 0 & I_2 \otimes A \\ I_2 \otimes A^T & 0 \end{pmatrix}, \quad Z = Z_{a,b,\theta} = \begin{pmatrix} 0 & \Lambda_{a,b,\theta} \otimes I_N \\ \Lambda_{a,b,\theta}^T \otimes I_N & 0 \end{pmatrix}.$$

In this notation $G_{a,b,\theta}(\eta) = (W - Z_{a,b,\theta} - i\eta)^{-1}$. We recall

$$A_{a,b,\theta} := I_2 \otimes A - \Lambda_{a,b,\theta} \otimes I_N$$

Then

$$G_{a,b,\theta}(\eta) = \begin{pmatrix} i\eta H_{a,b,\theta} & H_{a,b,\theta} A_{a,b,\theta} \\ A_{a,b,\theta}^T H_{a,b,\theta} & i\eta \tilde{H}_{a,b,\theta} \end{pmatrix},$$

where

$$\tilde{H}_{a,b,\theta} := \tilde{H}_{a,b,\theta}(\eta) = \left(A_{a,b,\theta}^T A_{a,b,\theta} + \eta^2\right)^{-1},$$
$$H_{a,b,\theta} := H_{a,b,\theta}(\eta) = \left(A_{a,b,\theta} A_{a,b,\theta}^T + \eta^2\right)^{-1}.$$

Define a linear operator $\mathcal{S} : M_{4N}(\mathbb{C}) \to M_4(\mathbb{C})$ as follows. Given a matrix $T \in M_{4N}(\mathbb{C})$ consisting of 16 blocks $T_{ij} \in M_N(\mathbb{C})$ for $i, j \in [[1, 4]]$, let

$$\mathcal{S}(T) = \begin{pmatrix} \langle T_{33}\rangle & \langle T_{34}\rangle & 0 & 0 \\ \langle T_{43}\rangle & \langle T_{44}\rangle & 0 & 0 \\ 0 & 0 & \langle T_{11}\rangle & \langle T_{12}\rangle \\ 0 & 0 & \langle T_{21}\rangle & \langle T_{22}\rangle \end{pmatrix}.$$

In this Appendix we use the standard technique of cumulant expansion to derive the deterministic approximations of the resolvent $G_{a,b,\theta}$ and products of two resolvents $G_{a,b,\theta} T G_{a,b,\theta}$ for some deterministic matrices T. See e.g. [12] for a similar argument for the resolvent of Wigner ensemble. We are not concerned with the optimality of the error bounds is η since we only apply these estimates with $\eta = N^{-\epsilon}$. To state the estimates we use the notion of stochastic domination defined here.

Definition 25 (Stochastic domination)**.** *Suppose that* $X = \{X_N(s) : N \in \mathbb{Z}_+, s \in S_N\}$ *and* $Y = \{Y_N(s) : N \in \mathbb{Z}_+, s \in S_N\}$ *are sequences of random variables, possibly parametrized by* s*. We say that* X *is stochastically dominated by* Y *uniformly in* s *and write* $X \prec Y$ *or* $X = \mathcal{O}(Y)$ *if for any* $\varepsilon, D > 0$ *we have*

$$\sup_{S_N} \mathbb{P}\left(X_N(s) > N^{\varepsilon} Y_N(s)\right) < N^{-D}$$

for large enough N*.*

Lemma 26. *For* $\eta > N^{-1/2+\varepsilon}$*, and any* $\alpha, \beta \in [[1,4]]$*,*

$$|\langle (G_{a,b,\theta}(\eta) - M_{a,b,\theta}(\eta))\, E_{\alpha,\beta}\rangle| \prec \frac{1}{N\eta^2}$$

uniformly in a, b, θ*, where* $E_{\alpha,\beta} = e_{\alpha,\beta} \otimes I_N \in M_{4N}(\mathbb{R})$ *and*

$$e_{\alpha,\beta} \in M_4(\mathbb{R}), \qquad (e_{\alpha,\beta})_{\alpha',\beta'} = \delta_{\alpha'=\alpha}\delta_{\beta'=\beta}$$

and $M(\eta) = M_{a,b,\theta}(\eta) \in M_{4N}(\mathbb{C})$ *is the solution to the matrix Dyson equation (MDE)*

$$[i\eta + Z + \mathcal{S}(M(\eta))]\, M(\eta) + I = 0, \tag{A.1}$$

satisfying $\eta \mathrm{Im}\, M(\eta) > 0$*.*

The existence and uniqueness of the solution to matrix Dyson equation (A.1) satisfying $\eta \mathrm{Im}\, M(\eta) > 0$ was shown in [20]. Before we prove this lemma we introduce the notation used to control the error terms in the proof.

Definition 27 (Renormalized term)**.** *For any smooth function* $f : M_{2N}(\mathbb{R}) \to M_{2N}(\mathbb{R})$ *define*

$$\underline{Wf(W)} = Wf(W) - \mathbb{E}_{\widetilde{W}} \widetilde{W}\left(\partial_{\widetilde{W}} f\right)(W),$$

where $\widetilde{W}$ *is an independent copy of* W*.*

Note that $\mathbb{E}_{\widetilde{W}} \widetilde{W}\left(\partial_{\widetilde{W}} f\right)(W)$ is the first order term in the cumulant expansion of $\mathbb{E}Wf(W)$ with respect to W. The following lemma lets us control the renormalized terms. The proof is deferred until the end of this Appendix.

Lemma 28 (Renormalized term bounds)**.** *For* $\eta \geq N^{-\frac{1}{2}+\varepsilon}$ *and any* $\alpha, \beta, \alpha', \beta' \in [[1,4]]$*, we have*

$$|\langle \underline{WG} E_{\alpha,\beta}\rangle| \prec \frac{1}{N\eta^2}, \tag{A.2}$$

$$\left|\left\langle \underline{WGE_{\alpha,\beta}G} E_{\alpha',\beta'}\right\rangle\right| \prec \frac{1}{N\eta^3} \tag{A.3}$$

uniformly in a, b, θ*.*

Proof of Lemma 26. From the definition of renormalized term we see that

$$WG = -\mathcal{S}(G)G - \frac{1}{N}\mathcal{T}(G)G + \underline{WG}.$$

Then

$$-I = -(W - Z - i\eta)G = (Z + i\eta + \mathcal{S}(G))\, G + \frac{1}{N}\mathcal{T}(G)G - \underline{WG},$$

where

$$\mathcal{T}(T) = \begin{pmatrix} 0 & 0 & T_{13} & T_{23} \\ 0 & 0 & T_{14} & T_{24} \\ T_{31} & T_{41} & 0 & 0 \\ T_{32} & T_{42} & 0 & 0 \end{pmatrix}.$$

We take a normalized trace separately for every $N \times N$ block in this equation to get

$$-\frac{1}{4}\delta_{\alpha\beta} = \sum_{\gamma=1}^{4} (Z + i\eta + \mathcal{S}(G))_{\alpha\gamma} \langle GE_{\gamma\beta}\rangle + \frac{1}{N} \langle \mathcal{T}(G)GE_{\alpha\beta}\rangle + \langle \underline{WG}E_{\alpha\beta}\rangle .$$

Note that $\frac{1}{N} \langle \mathcal{T}(G)GE_{\alpha\beta}\rangle \prec \frac{1}{N\eta^2}$ by a trivial bound and $\langle \underline{WG}E_{\alpha\beta}\rangle \prec \frac{1}{N\eta^2}$ by Lemma 28. Then

$$-\frac{1}{4}\delta_{\alpha\beta} = \sum_{\gamma=1}^{4} (Z + i\eta + \mathcal{S}(G))_{\alpha\gamma} \langle GE_{\gamma\beta}\rangle + \mathcal{O}\left(\frac{1}{N\eta^2}\right)$$

and the result of Lemma 26 follows by stability of the solution of MDE (A.1). ☐

Lemma 29. *For $\eta > N^{-\frac{1}{6}+\varepsilon}$, and any $\alpha, \beta, \alpha', \beta' \in [[1,4]]$,*

$$\langle G_{a,b,\theta}E_{\alpha\beta}G_{a,b,\theta}E_{\alpha'\beta'}\rangle = \left\langle M_{a,b,\theta}E_{\alpha\beta}M_{a,b,\theta}\mathcal{X}^{-1}(E_{\alpha'\beta'})\right\rangle + \mathcal{O}\left(\frac{1}{N\eta^6}\right)$$

uniformly in a, b, θ, where operator linear $\mathcal{X} = \mathcal{X}_{a,b,\theta}$ acting on $M_4(\mathbb{R})$ by

$$\mathcal{X}_{a,b,\theta}(\cdot) = I - \mathcal{S}(M_{a,b,\theta} \cdot M_{a,b,\theta}).$$

Proof. From the definition of G and (A.1) we see that

$$G = M + M\mathcal{S}(G-M)G - M\underline{WG} + \frac{1}{N}\mathcal{T}(G)G, \tag{A.4}$$

Now we consider $GE_{\alpha,\beta}G$. The integration by parts formula gives

$$\begin{aligned} GE_{\alpha,\beta}G = {} & ME_{\alpha,\beta}M + M\mathcal{S}(GE_{\alpha,\beta}G)M + ME_{\alpha,\beta}(G-M) \\ & + M\mathcal{S}(G-M)GE_{\alpha,\beta}G + M\mathcal{S}(GE_{\alpha,\beta}G)(G-M) \\ & + \frac{1}{N}M\mathcal{T}(G)GE_{\alpha,\beta}G + \frac{1}{N}\mathcal{T}(GE_{\alpha,\beta}G)G - M\underline{WGE_{\alpha,\beta}G}. \end{aligned} \tag{A.5}$$

Define a linear operator $\mathcal{B}$ acting on $M_{4N}(\mathbb{C})$ by

$$\mathcal{B}(\cdot) = I - M\mathcal{S}(\cdot)M.$$

We move the second term of (A.5) to the left, apply $\mathcal{B}^{-1}$ on both sides, multiply by $E_{\alpha',\beta'}$ on the right and take the normalized trace to get

$$\begin{aligned} \langle GE_{\alpha,\beta}GE_{\alpha',\beta'}\rangle = {} & \left\langle \mathcal{B}^{-1}(ME_{\alpha,\beta}M)E_{\alpha',\beta'}\right\rangle \\ & + \left\langle \mathcal{B}^{-1}(ME_{\alpha,\beta}(G-M))E_{\alpha',\beta'}\right\rangle \\ & + \left\langle \mathcal{B}^{-1}(M\mathcal{S}(G-M)GE_{\alpha,\beta}G)E_{\alpha',\beta'}\right\rangle \\ & + \left\langle \mathcal{B}^{-1}(M\mathcal{S}(GE_{\alpha,\beta}G)(G-M))E_{\alpha',\beta'}\right\rangle \\ & + \frac{1}{N}\left\langle \mathcal{B}^{-1}(M\mathcal{T}(G)GE_{\alpha,\beta}G)E_{\alpha',\beta'}\right\rangle \\ & + \frac{1}{N}\left\langle \mathcal{B}^{-1}(\mathcal{T}(GE_{\alpha,\beta}G)G)E_{\alpha',\beta'}\right\rangle \\ & - \left\langle \mathcal{B}^{-1}(M\underline{WGE_{\alpha,\beta}G})E_{\alpha',\beta'}\right\rangle . \end{aligned}$$

Now we notice that operators $\mathcal{B}$ and $\mathcal{X}$ are related thorough the following identity. For any $B_1, B_2 \in M_{4N}(\mathbb{C})$, we have

$$\left\langle \mathcal{B}^{-1}(B_1)B_2\right\rangle = \left\langle B_1\mathcal{X}^{-1}(B_2)\right\rangle .$$

Thus, we have

$$
\begin{aligned}
\langle GE_{\alpha,\beta}GE_{\alpha',\beta'}\rangle = &\left\langle ME_{\alpha,\beta}M\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \\
&+ \left\langle ME_{\alpha,\beta}(G-M)\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \\
&+ \left\langle M\mathcal{S}(G-M)GE_{\alpha,\beta}G\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \\
&+ \left\langle M\mathcal{S}(GE_{\alpha,\beta}G)(G-M)\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \\
&+ \frac{1}{N}\left\langle M\mathcal{T}(G)GE_{\alpha,\beta}G\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \\
&+ \frac{1}{N}\left\langle \mathcal{T}(GE_{\alpha,\beta}G)G\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \\
&- \left\langle M\underline{WGE_{\alpha,\beta}G}\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle
\end{aligned}
$$

To bound the error terms we use that $\|\mathcal{S}\|_{op} \lesssim 1$, $\|\mathcal{T}\|_{op} \lesssim 1$ and $\|\mathcal{X}\|_{op} \gtrsim \eta$ (see section A.1). Note that in the 2nd and 4th terms $G-M$ appears and it is multiplied by a 4×4 block matrix. In the 3rd term $G-M$ appears inside of the operator $\mathcal{S}$. Thus every instance of $G-M$ on the right can be bound using Lemma 26 by $\frac{1}{N\eta^2}$. Every other instance of G or M on the right we control using the operator norm bound $\|G\|, \|M\| \prec \frac{1}{\eta}$. In the final term we use Lemma 28 for the renormalized term. Putting this together we get

$$
\begin{aligned}
&\left\langle ME_{\alpha,\beta}(G-M)\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \prec \frac{1}{N\eta^4}, \\
&\left\langle M\mathcal{S}(G-M)GE_{\alpha,\beta}G\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \prec \frac{1}{N\eta^6}, \\
&\left\langle M\mathcal{S}(GE_{\alpha,\beta}G)(G-M)\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \prec \frac{1}{N\eta^6}, \\
&\frac{1}{N}\left\langle M\mathcal{T}(G)GE_{\alpha,\beta}G\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \prec \frac{1}{N\eta^5}, \\
&\frac{1}{N}\left\langle \mathcal{T}(GE_{\alpha,\beta}G)G\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \prec \frac{1}{N\eta^4}, \\
&\left\langle M\underline{WGE_{\alpha,\beta}G}\mathcal{X}^{-1}(E_{\alpha',\beta'})\right\rangle \prec \frac{1}{N\eta^5}.
\end{aligned}
$$

This concludes the proof. □

Proof of Lemma 28. We show the proof of (A.2) here. The proof of (A.3) is analogous. Every extra G in the trace gives rise to an additional $\frac{1}{\eta}$ in the bound.

The proof of (A.2) follows closely the proof of Theorem 4.1 of [12], so we outline the differences. Similarly to [12], we use the cumulant expansion

$$
\mathbb{E}A_{ij}f(W) = \sum_{m\geq 1}\frac{\kappa_{m+1}}{m!N^{\frac{1}{2}(m+1)}}\mathbb{E}\partial^m_{A_{ij}}f(W),
$$

where κ_m is mth cumulant of $\sqrt{N}A_{ij}$.

We introduce matrices $\Delta^{ij}\in M_{4N}(\mathbb{R})$ for $i,j\in[[1,N]]$, such that

$$
(\Delta^{ij})_{xy} = \delta_{x=2N+i}\delta_{y=j} + \delta_{x=3N+i}\delta_{y=N+j} + \delta_{x=j}\delta_{y=2N+i} + \delta_{x=N+j}\delta_{y=3N+i}.
$$

Then $W = \sum_{i,j=1}^N A_{ij}\Delta^{ij}$. Consider the second moment of $\langle \underline{WG}E_{\alpha,\beta}\rangle$ and use the cumulant expansion with respect to both W. The first derivative terms of of the second moment are

$$\mathbb{E}|\langle \underline{WG}E_{\alpha,\beta}\rangle|^2 = \sum_{i,j=1}^{N}\frac{\kappa_2}{N}\mathbb{E}\left\langle \Delta^{ij}GE_{\alpha,\beta}\right\rangle\left\langle E^*_{\alpha,\beta}G^*\Delta^{ij}\right\rangle$$
$$+\sum_{i,j=1}^{N}\sum_{i',j'=1}^{N}\frac{\kappa_2^2}{N^2}\mathbb{E}\left\langle \Delta^{ij}G\Delta^{i'j'}GE_{\alpha,\beta}\right\rangle\left\langle E^*_{\alpha,\beta}G^*\Delta^{ij}G^*\Delta^{i'j'}\right\rangle+\dots$$

The first term can be written as a sum of 16 terms of the type

$$\frac{\kappa_2}{N^2}\mathbb{E}\left\langle G^{(*)}B_1G^{(*)}B_2\right\rangle,$$

where $\|B_1\|,\|B_2\|\le 1$ and $G^{(*)}$ is indicating G or G^*. Since we do not need to control (A.2) by $\frac{1}{N\eta}$, we bound $\langle GB_1GB_2\rangle$ trivially by $\frac{1}{\eta^2}$. Similarly, the second term can be written as a sum of 16^2 terms of three types

$$\frac{\kappa_2^2}{N^3}\left\langle G^{(*)}B_1G^{(*)}B_2G^{(*)}B_3G^{(*)}B_4\right\rangle$$
$$\frac{\kappa_2^2}{N^2}\left\langle G^{(*)}B_1\right\rangle\left\langle G^{(*)}B_2G^{(*)}B_3G^{(*)}B_4\right\rangle$$
$$\frac{\kappa_2^2}{N^2}\left\langle G^{(*)}B_1G^{(*)}B_2\right\rangle\left\langle G^{(*)}B_3G^{(*)}B_4\right\rangle$$

All these terms can be bound by $\frac{\kappa_2^2}{N^2\eta^4}$. Similarly for higher order terms in the cumulant expansion each extra derivative adds another G into the term, which we bound by $\frac{1}{\eta}$ and another $\frac{1}{\sqrt{N}}$ due to taking a higher order cumulant of the entry of A. Analogous calculation works for $2p$-th moment. We refer the reader to [12] for details. □

A.1 Properties of the operator $\mathcal{X}$.

Note that matrix Dyson equation (A.1) invariant under transposition of M and under conjugation of M by a permutation matrix

$$\Pi=\begin{pmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix}.$$

The solution to MDE satisfying the condition $\eta\mathrm{Im}\,M(\eta)>0$ is unique by Theorem 2.1 of [20]. Thus for $\eta>0$ the solution is

$$M=\begin{pmatrix}M_{11}&M_{11}\Lambda(i\eta+M_{11})^{-1}\\(i\eta+M_{11})^{-1}\Lambda^TM_{11}&FM_{11}F\end{pmatrix},$$

where

$$F=\begin{pmatrix}0&1\\1&0\end{pmatrix}$$

and $M_{11}\in M_2(\mathbb{C})$ is satisfies $iM_{11}<0$ and

$$-M_{11}^{-1}=i\eta+FM_{11}F-\Lambda(i\eta+M_{11})^{-1}\Lambda^T$$

By a trivial computation we can derive the expansion of the solution in η.

$$M_{11}(\eta)=i\sqrt{1-a^2-b^2}\begin{pmatrix}\tan^{-1}\theta&0\\0&\tan\theta\end{pmatrix}+i\eta S+O(\eta^2),$$

where $S \in M_2^{sa}(\mathbb{R})$ has entries

$$S_{11} = -1 + \frac{1+\tan^{-2}\theta}{4(1-a^2-b^2)} + \frac{(1-a^2)(1-\tan^{-2}\theta)}{4b^2}$$
$$S_{22} = -1 + \frac{1+\tan^{2}\theta}{4(1-a^2-b^2)} + \frac{(1-a^2)(1-\tan^{2}\theta)}{4b^2}$$
$$S_{12} = S_{21} = -\frac{a}{2b}(\tan\theta - \tan^{-1}\theta).$$

Recall that the operator $\mathcal{X}$ acts on $M_4(\mathbb{C})$ by $\mathcal{X}(\cdot) = I - \mathcal{S}(M(\eta)\cdot M(\eta))$ and $\mathcal{X}^*(\cdot) = I - M(-\eta)\mathcal{S}(\cdot)M(-\eta)$. Using the formulas for $M(\eta)$ above, it is easy to verify the following.

- Operator $\mathcal{X}$ has an eigenvalue 1 with multiplicity 8. In particular, the left eigenspace of $\mathcal{X}$ corresponding to this eigenvalue is the space of 2×2 block matrices with 0 diagonal blocks.
- Define
$$\beta_+ = \frac{1}{2\sqrt{1-|w|^2}}(\tan\theta + \tan^{-1}\theta)\eta, \quad \beta_- = 2(1-a^2-b^2)$$
and
$$\gamma_\pm = 1 - a^2 + b^2 \pm \sqrt{(1-a^2+b^2)^2 - 4b^2}.$$
Operator $\mathcal{X}$ has two eigenvalues of the form $\beta_- + O(\eta)$, two eigenvalues of the form $\beta_+ + O(\eta)$, two eigenvalues of the form $\gamma_- + O(\eta)$, two eigenvalues of the form $\gamma_+ + O(\eta)$.
- As a consequence, we have $\|\mathcal{X}\|_{op} \gtrsim \eta$.

B Computation of resolvent quantities at $\theta = \frac{\pi}{4}$.

In this section we compute certain functions of $G_{a,b,\frac{\pi}{4}}(\eta)$ explicitly based on the observation that at $\theta = \frac{\pi}{4}$ this resolvent reduces to the regular resolvent $G_w(\eta)$ as follows. It is straightforward to check that for $w = a + ib$

$$G_{a,b,\frac{\pi}{4}}(\eta) = U^* \begin{pmatrix} G_w(\eta) & 0 \\ 0 & G_{\overline{w}}(\eta) \end{pmatrix} U,$$

where

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i & 0 & 0 \\ 0 & 0 & 1 & -i \\ 1 & i & 0 & 0 \\ 0 & 0 & 1 & i \end{pmatrix} \tag{B.1}$$

Note also that

$$UE_{3,3}U^* = \frac{1}{2}\left(E_{2,2} + E_{2,4} + E_{4,2} + E_{4,4}\right)$$
$$UE_{4,4}U^* = \frac{1}{2}\left(E_{2,2} - E_{2,4} - E_{4,2} + E_{4,4}\right)$$
$$UE_{3,4}U^* = \frac{i}{2}\left(E_{2,2} - E_{2,4} + E_{4,2} - E_{4,4}\right)$$
$$UE_{4,3}U^* = \frac{i}{2}\left(-E_{2,2} - E_{2,4} + E_{4,2} + E_{4,4}\right)$$

Suppose $z=a+ib$ and $\eta_{z,t}$ is defined by $t\left\langle H_z(\eta_{z,t})\right\rangle=1$. We show here that

$$\begin{pmatrix} \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4},11} \\ -2\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4},12} \\ \frac{N}{t}-\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4},22}\end{pmatrix}=0.$$

Indeed,

$$\begin{aligned}\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4},11}&=\frac{1}{i\eta_{z,t}}\operatorname{Tr}G_{a,b,\frac{\pi}{4}}E_{3,3}=\frac{1}{i\eta_{z,t}}\operatorname{Tr}\begin{pmatrix}G_w(\eta) & 0\\ 0 & G_{\overline{w}}(\eta)\end{pmatrix}UE_{3,3}U^*\\ &=\frac{1}{2i\eta_{z,t}}\operatorname{Tr}\begin{pmatrix}G_z(\eta) & 0\\ 0 & G_{\overline{z}}(\eta)\end{pmatrix}(E_{2,2}+E_{2,4}+E_{4,2}+E_{4,4})\\ &=\frac{1}{2}\operatorname{Tr}\tilde{H}_z+\frac{1}{2}\operatorname{Tr}\tilde{H}_{\overline{z}}=\frac{N}{t}.\end{aligned}$$

Similarly, $\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4},12}=0$ and $\operatorname{Tr}\tilde{H}_{a,b,\frac{\pi}{4},22}=\frac{N}{t}$.

B.1 Computing $\mathbf{Q}_{a,b,\frac{\pi}{4}}$.

Lemma 30. *For $a\in[-1,1]$, $b\in[\kappa,1)$ and $\eta>N^{-\delta}$ we have $\mathbf{Q}_{a,b,\frac{\pi}{4}}\gtrsim\frac{N}{\eta^2}$ uniformly in a,b.*

Proof. Note that as a quadratic form on $M_2^{sa}(\mathbb{R})$,

$$\mathbf{Q}_{a,b,\frac{\pi}{4}}=-\frac{4N}{\eta^2}\left\langle\left(G_{a,b,\frac{\pi}{4}}\begin{pmatrix}0&0\\0&P\end{pmatrix}\right)^2\right\rangle$$

Thus, we can compute the entries of Q in the basis

$$\begin{pmatrix}1&0\\0&0\end{pmatrix},\quad\frac{1}{2}\begin{pmatrix}0&1\\1&0\end{pmatrix},\quad\begin{pmatrix}0&0\\0&1\end{pmatrix}.$$

We get

$$\mathbf{Q}_{a,b,\frac{\pi}{4}}=\frac{N}{4}\begin{pmatrix}\left\langle\tilde{H}_z^2\right\rangle+\left\langle\tilde{H}_z\tilde{H}_{\bar{z}}\right\rangle & 0 & \left\langle\tilde{H}_z^2\right\rangle-\left\langle\tilde{H}_z\tilde{H}_{\bar{z}}\right\rangle\\ 0 & 4\left\langle\tilde{H}_z\tilde{H}_{\bar{z}}\right\rangle & 0\\ \left\langle\tilde{H}_z^2\right\rangle-\left\langle\tilde{H}_z\tilde{H}_{\bar{z}}\right\rangle & 0 & \left\langle\tilde{H}_z^2\right\rangle+\left\langle\tilde{H}_z\tilde{H}_{\bar{z}}\right\rangle\end{pmatrix}.$$

This matrix has eigenvalues $2N\left\langle\tilde{H}_z^2\right\rangle, 2N\left\langle\tilde{H}_z\tilde{H}_{\bar{z}}\right\rangle, N\left\langle\tilde{H}_z\tilde{H}_{\bar{z}}\right\rangle$, all of which are bounded below by $N\eta^{-2}$ by [13]. □

C Jacobian calculation

Let us recall the Schur decomposition construction. Define the manifold

$$\Omega=\mathbb{R}\times\mathbb{R}_+\times[0,\pi/2)\times V^2(\mathbb{R}^N)\times M_{(N-2)\times 2}(\mathbb{R}),$$

where $V^2(\mathbb{R}^N)$ is a Stiefel manifold, i.e.

$$V^2(\mathbb{R}^N)=\mathbf{O}(N)/\mathbf{O}(N-2)=\{(v_1,v_2)\in\mathbb{S}^{N-1}\times\mathbb{S}^{N-1}:\ v_1^Tv_2=0\}.$$

Choose a smooth map $R : V^2(\mathbb{R}^N) \to \mathbf{O}(N)$ such that for any $\mathbf{v} = (v_1, v_2) \in V^2(\mathbb{R}^N)$ we have $R(\mathbf{v})\mathbf{e}_i = v_i$ for $i = 1, 2$. For any $(a, b, \theta) \in \mathbb{R} \times \mathbb{R}_+ \times [0, \pi/2)$ define

$$\Lambda_{a,b,\theta} = \begin{pmatrix} a & b\tan\theta \\ -\frac{b}{\tan\theta} & a \end{pmatrix}.$$

Define a map

$$\Phi : \Omega \times M_{N-2}(\mathbb{R}) \to M_N(\mathbb{R})$$

such that

$$\Phi(a, b, \theta, \mathbf{v}, W, M^{(1)}) = M = R(\mathbf{v}) \begin{pmatrix} \Lambda_{a,b,\theta} & W^T \\ 0 & M^{(1)} \end{pmatrix} R(\mathbf{v})^T.$$

Note that if $\mathbf{v} = (v_1, v_2)$, then $\cos\theta v_1 \pm i \sin\theta v_2$ are eigenvectors of the right-hand side with corresponding eigenvalues $\lambda, \overline{\lambda} = a \pm bi$.

Lemma 31. *The Jacobian of Φ is*

$$J(\Phi) = 16b^2 \frac{|\cos 2\theta|}{\sin^2 2\theta} \left|\det\left(M^{(1)} - \lambda\right)\right|^2 \tag{C.1}$$

Proof. Consider a smooth atlas on $V^2(\mathbb{R})$ and let $\varphi : U \to \tilde{U} \subset \mathbb{R}^{2N-3}$, where $U \subset V^2(\mathbb{R})$, be a chart in this atlas. Denote the standard coordinates in $\mathbb{R}^{2N-3}$ by $(u^1, \ldots, u^{2N-3})$. By abusing notation, in the the rest of the proof we view v_i as

$$v_i = v_i(u^1, \ldots, u^{2N-3}) = \left(\varphi^{-1}(u^1, \ldots, u^{2N-3})\right)_i$$

for $i = 1, 2$. Similarly, we will view $R(\mathbf{v})$ as a function from $\tilde{U}$ to $M_N(\mathbb{R})$. Note that

$$J(\Phi) = J(\Phi \circ \varphi^{-1}) J(\varphi).$$

We start by computing $J(\Phi \circ \varphi^{-1})$. By differentiating the definition of Φ, we get

$$\begin{aligned} dM = {} & dR(\mathbf{v}) \begin{pmatrix} \Lambda_{a,b,\theta} & W^T \\ 0 & M^{(1)} \end{pmatrix} R(\mathbf{v})^T + R(\mathbf{v}) \begin{pmatrix} \Lambda_{a,b,\theta} & W^T \\ 0 & M^{(1)} \end{pmatrix} dR(\mathbf{v})^T \\ & + R(\mathbf{v}) \begin{pmatrix} d\Lambda_{a,b,\theta} & dW^T \\ 0 & dM^{(1)} \end{pmatrix} R(\mathbf{v})^T. \end{aligned}$$

Consider $d\tilde{M} = R(\mathbf{v})^T dM R(\mathbf{v})$. Since $R(\mathbf{v})$ is orthogonal, the volume form on $M_N(\mathbb{R})$ can be expressed in terms of $d\tilde{M}$ as $\bigwedge_{i=1}^N \bigwedge_{j=1}^N dM_{ij} = \bigwedge_{i=1}^N \bigwedge_{j=1}^N d\tilde{M}_{ij}$. Note that

$$\begin{aligned} d\tilde{M} = {} & R(\mathbf{v})^T dR(\mathbf{v}) \begin{pmatrix} \Lambda_{a,b,\theta} & W^T \\ 0 & M^{(1)} \end{pmatrix} + \begin{pmatrix} \Lambda_{a,b,\theta} & W^T \\ 0 & M^{(1)} \end{pmatrix} dR(\mathbf{v})^T R(\mathbf{v}) \\ & + \begin{pmatrix} d\Lambda_{a,b,\theta} & dW^T \\ 0 & dM^{(1)} \end{pmatrix}. \end{aligned}$$

Since $R(\mathbf{v})\mathbf{e}_i = v_i$ for $i = 1, 2$, the first two columns of $R(\mathbf{v})$ are v_1 and v_2. Denote the rest of the matrix $R(\mathbf{v})$ by $\tilde{R}(\mathbf{v})$. Since $R(\mathbf{v})$ is orthogonal, $\tilde{R}(\mathbf{v})^T v_i = 0$ for $i = 1, 2$ and, thus, $d\tilde{R}(\mathbf{v})^T v_i = -\tilde{R}(\mathbf{v})^T dv_i$. Then

$$R(\mathbf{v})^T dR(\mathbf{v}) = -dR(\mathbf{v})^T R(\mathbf{v}) = \begin{pmatrix} df E & -dH^T \\ dH & \tilde{R}(\mathbf{v})^T d\tilde{R}(\mathbf{v}) \end{pmatrix},$$

where

$$\begin{aligned}
df &= v_2^T dv_1, \\
E &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\
dH &= (dH_1, dH_2) = (\tilde{R}(\mathbf{v})^T dv_1, \tilde{R}(\mathbf{v})^T dv_2).
\end{aligned}$$

Plugging this into the expression for $d\tilde{M}$, we get

$$d\tilde{M} = \begin{pmatrix} d\tilde{M}^{(11)} & d\tilde{M}^{(12)} \\ d\tilde{M}^{(21)} & d\tilde{M}^{(22)} \end{pmatrix},$$

where

$$\begin{aligned}
d\tilde{M}^{(11)} &= d\Lambda_{a,b,\theta} + df(E\Lambda_{a,b,\theta} - \Lambda_{a,b,\theta}E) - W^T dH, \\
d\tilde{M}^{(12)} &= dW^T + df EW^T - dH^T M^{(1)} + \Lambda_{a,b,\theta} dH^T - W^T \tilde{R}(\mathbf{v})^T d\tilde{R}(\mathbf{v}), \\
d\tilde{M}^{(21)} &= dH\Lambda_{a,b,\theta} - M^{(1)} dH, \\
d\tilde{M}^{(22)} &= dM^{(1)} + dHW^T + \tilde{R}(\mathbf{v})^T d\tilde{R}(\mathbf{v}) M^{(1)} - M^{(1)} \tilde{R}(\mathbf{v})^T d\tilde{R}(\mathbf{v}).
\end{aligned}$$

Now we compute the volume form by changing the order and noticing that dW and $dM^{(1)}$ appear only in $d\tilde{M}^{(12)}$ and $d\tilde{M}^{(22)}$ respectively.

$$\begin{aligned}
\bigwedge_{i=1}^{N}\bigwedge_{j=1}^{N} d\tilde{M}_{ij} &= \bigwedge_{i=1}^{2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(11)} \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(21)} \wedge \bigwedge_{i=1}^{2}\bigwedge_{j=1}^{N-2} d\tilde{M}_{ij}^{(12)} \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{N-2} d\tilde{M}_{ij}^{(22)} \\
&= \bigwedge_{i=1}^{2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(11)} \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(21)} \wedge \bigwedge_{i=1}^{2}\bigwedge_{j=1}^{N-2} dW_{ji} \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{N-2} dM_{ij}^{(1)}.
\end{aligned}$$

Notice that

$$E\Lambda_{a,b,\theta} - \Lambda_{a,b,\theta}E = b(\tan\theta - \tan^{-1}\theta)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and

$$d\Lambda_{a,b,\theta} = \begin{pmatrix} da & \tan\theta db + \frac{b}{\cos^2\theta} d\theta \\ -\tan^{-1}\theta db + \frac{b}{\sin^2\theta} d\theta & da \end{pmatrix}.$$

Thus we can simplify the first two terms of the volume form:

$$\begin{aligned}
&\bigwedge_{i=1}^{2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(11)} \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(21)} = (da + b(\tan\theta - \tan^{-1}\theta)df) \\
&\wedge \left(\tan\theta db + \frac{b}{\cos^2\theta} d\theta\right) \wedge (da - b(\tan\theta - \tan^{-1}\theta)df) \\
&\wedge \left(-\tan^{-1}\theta db + \frac{b}{\sin^2\theta} d\theta\right) \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(21)} \\
&= 4b^2 \frac{\tan\theta - \tan^{-1}\theta}{\sin\theta\cos\theta} da \wedge db \wedge d\theta \wedge df \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(21)} \\
&= -16b^2 \frac{\cos 2\theta}{\sin^2 2\theta} da \wedge db \wedge d\theta \wedge df \wedge \bigwedge_{i=1}^{N-2}\bigwedge_{j=1}^{2} d\tilde{M}_{ij}^{(21)}.
\end{aligned}$$

To simplify $d\tilde{M}^{(22)}$ consider its two columns $d\tilde{M}_1^{(22)}$ and $d\tilde{M}_2^{(22)}$. Then

$$\begin{pmatrix} d\tilde{M}_1^{(21)} \\ d\tilde{M}_2^{(21)} \end{pmatrix} = \begin{pmatrix} a - M^{(1)} & -b\tan^{-1}\theta \\ b\tan\theta & a - M^{(1)} \end{pmatrix} \begin{pmatrix} dH_1 \\ dH_2 \end{pmatrix}$$

Thus

$$\bigwedge_{j=1}^{2}\bigwedge_{i=1}^{N-2} d\tilde{M}_{ij}^{(21)} = \det\left\{\Lambda_{a,b,\theta}^T \otimes I_{N-2} - I_2 \otimes M^{(1)}\right\} \bigwedge_{j=1}^{2}\bigwedge_{i=1}^{N-2} dH_{ij}.$$

Consider matrices $L_\theta, R_\theta \in M_2(\mathbb{C})$ consisting of left and right eigenvectors of $\Lambda_{a,b,\theta}$ such that $L_\theta^* R_\theta = I$. Then $\Lambda_{a,b,\theta}^T = L_\theta^* \begin{pmatrix} \overline{\lambda} & 0 \\ 0 & \lambda \end{pmatrix} R_\theta$. Then the determinant above is equal to

$$\begin{aligned} &\det\left\{\Lambda_{a,b,\theta}^T \otimes I_{N-2} - I_2 \otimes M^{(1)}\right\} \\ &= \det\left\{L_\theta^* \otimes I_{N-2} \begin{pmatrix} \overline{\lambda} - M^{(1)} & 0 \\ 0 & \lambda - M^{(1)} \end{pmatrix} R_\theta \otimes I_{N-2}\right\} \\ &= \left|\det(M^{(1)} - \lambda)\right|^2. \end{aligned}$$

It remains to show that

$$J(\varphi) df \wedge \bigwedge_{j=1}^{2}\bigwedge_{i=1}^{N-2} dH_{ij} = d\mathbf{v},$$

where $d\mathbf{v}$ is the rotationally invariant volume form on $V^2(\mathbb{R})$. First, choose a translation invariant local coordinate chart, so that $J(\varphi)$ is constant in $\mathbf{v}$ (we can always do this because the action of $\mathbf{O}(N)$ on $V^2(\mathbb{R}^N) = \mathbf{O}(N)/\mathbf{O}(N-2)$ is transitive). For any $(v_1, v_2) \in V^2(\mathbb{R}^N)$, choose a smooth lift to $(v_1, v_2, \ldots, v_N) \in \mathbf{O}(N)$. We have

$$df \wedge \bigwedge_{j=1,2}\bigwedge_{i=1}^{N-2} dH_{ij} = \bigwedge_{\substack{i=1,\ldots,N \\ j=1,2 \\ j<i}} v_i^T dv_j$$

We claim this is translation invariant; this would complete the proof. To this end, note

$$\bigwedge_{\substack{i,j=1,\ldots,N \\ j<i}} v_i^T dv_j$$

is translation invariant (see the bottom of page 16 of [16]). But this is a smooth lift of $\bigwedge_{\substack{i=1,\ldots,N \\ j=1,2 \\ j<i}} v_i^T dv_j$, so the proposed translation invariance follows. $\square$

References

[1] A. Adhikari, S. Dubova, C. Xu, and J. Yin. Eigenstate thermalization hypothesis for generalized wigner matrices. *arXiv preprint arXiv:2302.00157*, 2023.

[2] A. Aggarwal. Bulk universality for generalized wigner matrices with few moments. *Probability Theory and Related Fields*, 173:375–432, 2019.

[3] J. Alt, L. Erdős, and T. Krüger. Local inhomogeneous circular law. *Communications in Mathematical Physics*, 388:1005–1048, 2021.

[4] J. Alt, L. Erdős, and T. Krüger. Spectral radius of random matrices with independent entries. *Probability and Mathematical Physics*, 2(2):221–280, 2021.

[5] A. Borodin and C. Sinclair. The ginibre ensemble of real random matrices and its scaling limits. *Communications in Mathematical Physics*, 291:177–224, 2009.

[6] P. Bourgade and G. Dubach. The distribution of overlaps between eigenvectors of ginibre matrices. *Probability Theory and Related Fields*, 177:397–464, 2020.

[7] P. Bourgade, H.-T. Yau, and J. Yin. Local circular law for random matrices. *Probability Theory and Related Fields*, 159(3-4):545–595, 2014.

[8] G. Cipolloni, L. Erdős, and J. Henheik. Eigenstate thermalisation at the edge for wigner matrices. *arXiv preprint arXiv:2309.05488*, 2023.

[9] G. Cipolloni, L. Erdős, J. Henheik, and D. Schröder. Optimal lower bound on eigenvector overlaps for non-hermitian random matrices. *arXiv preprint arXiv:2301.03549*, 2023.

[10] G. Cipolloni, L. Erdős, and D. Schröder. Towards the bulk universality of non-hermitian random matrices. *arXiv preprint arXiv:1909.06350*, 2020.

[11] G. Cipolloni, L. Erdős, and D. Schröder. Edge universality for non-hermitian random matrices. *Probability Theory and Related Fields*, 179(1-2):1–28, 2021.

[12] G. Cipolloni, L. Erdős, and D. Schröder. Eigenstate thermalization hypothesis for wigner matrices. *Communications in Mathematical Physics*, 388:1005–1048, 2021.

[13] G. Cipolloni, L. Erdős, and D. Schröder. Optimal multi-resolvent local laws for wigner matrices. *Electronic Journal of Probability*, 27:1–38, 2022.

[14] G. Cipolloni, L. Erdős, and D. Schröder. Mesoscopic central limit theorem for non-hermitian random matrices. *Probability Theory and Related Fields*, pages 1–52, 2023.

[15] J. A. Dı, R. Gutiérrez-Jáimez, et al. On wishart distribution: some extensions. *Linear Algebra and its Applications*, 435(6):1296–1310, 2011.

[16] A. Edelman. The probability that a random real gaussian matrix has k real eigenvalues, related distributions, and the circular law. *Journal of Multivariate Analysis*, 60(2):203–232, 1997.

[17] L. Erdős, S. Péché, J. A. Ramírez, B. Schlein, and H.-T. Yau. Bulk universality for wigner matrices. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 63(7):895–925, 2010.

[18] L. Erdős, H.-T. Yau, and J. Yin. Universality for generalized wigner matrices with bernoulli distribution. *Journal of Combinatorics*, 2(1):15–81, 2011.

[19] V. L. Girko. The circular law. *Teoriya Veroyatnostei i ee Primeneniya*, 29(4):669–679, 1984.

[20] J. W. Helton, R. R. Far, and R. Speicher. Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints. *International Mathematics Research Notices*, 2007(9):rnm086–rnm086, 2007.

[21] J. Huang, B. Landon, and H.-T. Yau. Bulk universality of sparse random matrices. *Journal of Mathematical Physics*, 56(12), 2015.

[22] A. Knowles and J. Yin. Anisotropic local laws for random matrices. *Probability Theory and Related Fields*, 169:257–352, 2016.

[23] A. Maltsev and M. Osman. Bulk universality for complex non-hermitian matrices with independent and identically distributed entries. *arXiv preprint arXiv:2310.11429*, 2023.

[24] M. Mehta. Random matrices and the statistical theory of energy levels, 1967.

[25] M. Osman. Universality for weakly non-hermitian matrices: Bulk limit. *arXiv preprint arXiv:2310.15001*, 2023.

[26] M. Osman. Bulk universality for real matrices with independent and identically distributed entries. *arXiv preprint arXiv:2402.04071*, 2024.

[27] T. Tao and V. Vu. Universality of local eigenvalue statistics. *Acta Mathematica*, 206(1):127–204, 2011.

[28] T. Tao and V. Vu. Random matrices: Universality of local spectral statistics of non-hermitian matrices. *The Annals of Probability*, 43(2):782–874, 2015.