PURE and APPLIED ANALYSIS

PAA

Giorgio Cipolloni, László Erdős,
Torben Krüger and Dominik Schröder

CUSP UNIVERSALITY FOR RANDOM MATRICES II: THE REAL SYMMETRIC CASE

msp

vol. 1 no. 4 2019

PURE and APPLIED
ANALYSIS

Vol. 1, No. 4, 2019
dx.doi.org/10.2140/paa.2019.1.615

CUSP UNIVERSALITY FOR RANDOM MATRICES II: THE REAL SYMMETRIC CASE

GIORGIO CIPOLLONI, LÁSZLÓ ERDŐS, TORBEN KRÜGER AND DOMINIK SCHRÖDER

We prove that the local eigenvalue statistics of real symmetric Wigner-type matrices near the cusp points of the eigenvalue density are universal. Together with the companion paper by Erdős et al. (2018, arXiv:1809.03971), which proves the same result for the complex Hermitian symmetry class, this completes the last remaining case of the Wigner–Dyson–Mehta universality conjecture after bulk and edge universalities have been established in the last years. We extend the recent Dyson Brownian motion analysis at the edge by Landon and Yau (2017, arXiv:1712.03881) to the cusp regime using the optimal local law by Erdős et al. (2018, arXiv:1809.03971) and the accurate local shape analysis of the density by Ajanki et al. (2015, arXiv:1506.05095) and Alt et al. (2018, arXiv:1804.07752). We also present a novel PDE-based method to improve the estimate on eigenvalue rigidity via the maximum principle of the heat flow related to the Dyson Brownian motion.

1. Introduction

We consider *Wigner-type* matrices, i.e., $N \times N$ Hermitian random matrices H with independent, not necessarily identically distributed entries above the diagonal; these are a natural generalisation of the standard Wigner ensembles that have i.i.d. entries. The Wigner–Dyson–Mehta (WDM) conjecture asserts that the local eigenvalue statistics are universal; i.e., they are independent of the details of the ensemble and

MSC2010: 60B20, 15B52.
Keywords: cusp universality, Dyson Brownian motion, local law.

depend only on the *symmetry type*, i.e., on whether H is real symmetric or complex Hermitian. Moreover, different statistics emerge in the bulk of the spectrum and at the spectral edges with a square-root vanishing behaviour of the eigenvalue density. The WDM conjecture for both symmetry classes has been proven for Wigner matrices; see [Erdős and Yau 2017] for complete historical references. Recently it has been extended to more general ensembles including Wigner-type matrices in the bulk and edge regimes; we refer to the companion paper [Erdős et al. 2018] for up-to-date references.

The key tool for the recent proofs of the WDM conjecture is the Dyson Brownian motion (DBM), a system of coupled stochastic differential equations. The DBM method has evolved during the last years. The original version, presented in [Erdős and Yau 2017], was in the spirit of a high-dimensional analysis of a strongly correlated Gibbs measure and its dynamics. Starting in [Erdős and Yau 2015] with the analysis of the underlying parabolic equation and its short-range approximation, the PDE component of the theory became prominent. With the coupling idea, introduced in [Bourgade et al. 2016; Bourgade and Yau 2017], the essential part of the proofs became fully deterministic, greatly simplifying the technical aspects. In the current paper we extend this trend and use PDE methods even for the proof of the rigidity bound, a key technical input, that earlier was obtained with direct random matrix methods.

The historical focus on the bulk and edge universalities has been motivated by the Wigner ensemble since, apart from the natural bulk regime, its semicircle density vanishes as a square root near the edges, giving rise to the Tracy–Widom statistics. Beyond the Wigner ensemble, however, the density profile shows a much richer structure. Already Wigner matrices with nonzero expectation on the diagonal, also called the *deformed Wigner ensemble*, may have a density supported on several intervals and a cubic root cusp singularity in the density arises whenever two such intervals touch each other as some deformation parameter varies. Since local spectral universality is ultimately determined by the local behaviour of the density near its vanishing points, the appearance of the cusp gives rise to a new type of universality. This was first observed in [Brézin and Hikami 1998b] and the local eigenvalue statistics at the cusp can be explicitly described by the Pearcey process in the complex Hermitian case [Tracy and Widom 2006]. The corresponding explicit formulas for the real symmetric case have not yet been established.

The key classification theorem [Ajanki et al. 2017a] for the density of Wigner-type matrices showed that the density may vanish only as a square root (at regular edges) or as a cubic root (at cusps); no other singularity may occur. This result has recently been extended to a large class of matrices with correlated entries [Alt et al. 2018a]. In other words, the cusp universality is the third and last universal spectral statistics for random matrix ensembles arising from natural generalisations of the Wigner matrices. We note that invariant β-ensembles may exhibit further universality classes; see [Claeys et al. 2018].

In the companion paper [Erdős et al. 2018] we established cusp universality for Wigner-type matrices in the complex Hermitian symmetry class. In the present work we extend this result to the real symmetric class and even to certain space-time correlation functions. In fact, we show the appearance of a natural one-parameter family of universal statistics associated to a family of singularities of the eigenvalue density that we call *physical cusps*. In both works we follow the *three-step strategy*, a general method developed for proving local spectral universality for random matrices; see [Erdős and Yau 2017] for a pedagogical introduction. The first step is the *local law* or *rigidity*, establishing the location of the eigenvalues with a precision slightly above the typical local eigenvalue spacing. The second step is to establish universality for ensembles with a tiny Gaussian component. The third step is a perturbative argument to remove this tiny Gaussian component relying on the optimal local law. The first and third steps are insensitive to the symmetry type; in fact the optimal local law in the cusp regime has been established for both symmetry classes in [Erdős et al. 2018] and it completes also the third step in both cases.

There are two different strategies for the second step. In the complex Hermitian symmetry class, the Brézin–Hikami formula [1998a] turns the problem into a saddle-point analysis for a contour integral. This direct path was followed in [Erdős et al. 2018], relying on the optimal local law. In the real symmetric case, lacking the Brézin–Hikami formula, only the second strategy via the analysis of Dyson Brownian motion (DBM) is feasible. This approach exploits the very fast decay to local equilibrium of DBM. It is the most robust and powerful method up to now to establish local spectral universality. In this paper we present a version of this method adjusted to the cusp situation. We will work in the real symmetric case for definiteness. The proof can easily be modified for the complex Hermitian case as well. The DBM method does not explicitly yield the local correlation kernel. Instead it establishes that the local statistics are universal and therefore can be identified from a reference ensemble that we will choose as the simplest Gaussian ensemble exhibiting a cusp singularity.

In this paper we partly follow the recent DBM analysis at the regular edges [Landon and Yau 2017] and we extend it to the cusp regime, using the optimal local law from the companion paper [Erdős et al. 2018] and the precise control of the density near the cusps [Ajanki et al. 2015; Alt et al. 2018a]. The main conceptual difference between [Landon and Yau 2017] and the current work is that we obtain the necessary local law along the time evolution of DBM via novel DBM methods in Section 6. Some other steps, such as the Sobolev inequality, heat kernel estimates from [Bourgade et al. 2014] and the finite speed of propagation [Erdős and Yau 2015; Bourgade and Yau 2017; Landon and Yau 2017], require only moderate adjustments for the cusp regime, but for completeness we include them in the Appendix. The comparison of the short-range approximation of the DBM with the full evolution, Lemma 7.2 and Lemma C.1, will be presented in detail in Section 7 and in Appendix C since it is more involved in the cusp setup, after the necessary estimates on the semicircular flow near the cusp are proven in Section 4.

We now outline the novelties and main difficulties at the cusp compared with the edge analysis in [Landon and Yau 2017]. The basic idea is to interpolate between the time evolution of two DBMs, with initial conditions given by the original ensemble and the reference ensemble, respectively, after their local densities have been matched by shift and scaling. Beyond this common idea there are several differences.

The first difficulty lies in the rigidity analysis of the DBM starting from the interpolated initial conditions. The optimal rigidity from [Erdős et al. 2018], which holds for very general Wigner-type matrices, applies for the flows of both the original and the reference matrices, but it does not directly apply to the interpolating process. The latter starts from a regular initial data but it runs for a very short time, violating the *flatness* (i.e., effective mean-field) assumption of [Erdős et al. 2018]. While it is possible to extend the analysis of [Erdős et al. 2018] to this case, here we chose a technically lighter and conceptually more interesting route. We use the maximum principle of the DBM to transfer rigidity information on the reference process to the interpolating one after an appropriate localisation. Similar ideas for proving rigidity of the β-DBM flow have been used in the bulk [Huang and Landon 2016] and at the edge [Adhikari and Huang 2018].

The second difficulty in the cusp regime is that the shape of the density is highly unstable under the semicircular flow that describes the evolution of the density under the DBM. The regular edge analysed in [Landon and Yau 2017] remains of square-root type along its dynamics and it can be simply described by its location and its multiplicative *slope parameter* — both vary regularly with time. In contrast, the evolution of the cusp is a relatively complicated process: it starts with a small gap that shrinks to zero as the cusp forms and then continues developing a small local minimum. Heavily relying on the main results of [Alt et al. 2018a], the density is described by quite involved shape functions, see (2-3c), (2-3e), that have a two-scale structure, given in terms of a total of three parameters, each varying on different time scales. For example, the location of the gap moves linearly with time, the length of gap shrinks as the $\frac{3}{2}$-th power of the time, while the local minimum after the cusp increases as the $\frac{1}{2}$-th power of the time. The scaling behaviour of the corresponding quantiles, which approximate the eigenvalues by rigidity, follows the same complicated pattern of the density. All these require a very precise description of the semicircular flow near the cusp as well as the optimal rigidity.

The third difficulty is that we need to run the DBM for a relatively long time in order to exploit the local decay; in fact this time scale, $N^{-\frac{1}{2}+\epsilon}$ is considerably longer than the characteristic time scale $N^{-\frac{3}{4}}$ on which the physical cusp varies under the semicircular flow. We need to tune the initial condition very precisely so that after a relatively long time it develops a cusp exactly at the right location with the right slope.

The fourth difficulty is that, unlike for the regular edge regime, the eigenvalues or quantiles on both sides of the (physical) cusp contribute to the short-range approximation of the dynamics, and their effect cannot be treated as mean-field. Moreover, there are two scaling regimes for quantiles corresponding to the two-scale structure of the density.

Finally, we note that the analysis of the semicircular flow around the cusp, partly completed already in the companion paper [Erdős et al. 2018], is relatively short and transparent despite its considerably more complex pattern compared to the corresponding analysis around the regular edge. This is mostly due to strong results imported from the general shape analysis [Ajanki et al. 2015]. Not only are the exact formulas for the density shapes taken over, but we also heavily rely on the $\frac{1}{3}$-Hölder continuity in space and time of the density and its Stieltjes transform, established in the strongest form in [Alt et al. 2018a].

Notations and conventions. We now introduce some custom notation we use throughout the paper. For integers n we define $[n] := \{1, \dots, n\}$. For positive quantities f, g, we write $f \lesssim g$ and $f \sim g$ if $f \le Cg$ or, respectively, $cg \le f \le Cg$ for some constants c, C that depend only on the *model parameters*, i.e., on the constants appearing in the basic Assumptions (A)–(C) listed in Section 2 below. Similarly, we write $f \ll g$ if $f \le cg$ for some tiny constant $c > 0$ depending on the model parameters. We denote vectors by bold-faced lower case Roman letters $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^N$, and matrices by upper case Roman letters $A, B \in \mathbb{C}^{N\times N}$. We write $\langle A\rangle := N^{-1}\operatorname{Tr} A$ and $\langle \boldsymbol{x}\rangle := N^{-1}\sum_{a\in[N]} x_a$ for the averaged trace and the average of a vector. We often identify a diagonal matrix with the vector of its diagonal elements. Accordingly, for any matrix R, we denote by $\operatorname{diag}(R)$ the vector of its diagonal elements, and for any vector $\boldsymbol{r}$ we denote by $\operatorname{diag}(\boldsymbol{r})$ the corresponding diagonal matrix.

We will frequently use the concept of "with very high probability", meaning that for any fixed $D > 0$ the probability of the event is bigger than $1 - N^{-D}$ if $N \geq N_0(D)$.

2. Main results

For definiteness we consider the real symmetric case $H \in \mathbb{R}^{N\times N}$. With small modifications, the proof presented in this paper works for the complex Hermitian case as well, but this case was already considered in [Erdős et al. 2018] with a contour integral analysis. Let $W = W^* \in \mathbb{R}^{N\times N}$ be a symmetric random matrix and $A = \mathrm{diag}(\boldsymbol{a})$ be a deterministic diagonal matrix with entries $\boldsymbol{a} = (a_i)_{i=1}^N \in \mathbb{R}^N$. We say that W is of *Wigner type* [Ajanki et al. 2017b] if its entries w_{ij} for $i \leq j$ are centred, $\mathbb{E}\, w_{ij} = 0$, independent random variables. We define the *variance matrix* or *self-energy matrix* $S = (s_{ij})_{i,j=1}^N$, $s_{ij} := \mathbb{E}\, w_{ij}^2$. In [Ajanki et al. 2017b] it was shown that as N tends to infinity, the resolvent $G(z) := (H - z)^{-1}$ of the *deformed Wigner-type matrix* $H = A + W$ entrywise approaches a diagonal matrix $M(z) := \mathrm{diag}(\boldsymbol{m}(z))$ for $z \in \mathbb{H} := \{z \in \mathbb{C} : \Im z > 0\}$. The entries $\boldsymbol{m} = (m_1, \dots, m_N) \colon \mathbb{H} \to \mathbb{H}^N$ of M have positive imaginary parts and solve the *Dyson equation*

$$-\frac{1}{m_i(z)} = z - a_i + \sum_{j=1}^N s_{ij} m_j(z), \quad z \in \mathbb{H} := \{z \in \mathbb{C} : \Im z > 0\},\ i \in [N]. \tag{2-1}$$

We call M or $\boldsymbol{m}$ the *self-consistent Green's function*. The normalised trace $\langle M \rangle$ of M is the Stieltjes transform $\langle M(z)\rangle = \int_{\mathbb{R}} (\tau - z)^{-1} \rho(\mathrm{d}\tau)$ of a unique probability measure ρ on $\mathbb{R}$ that approximates the empirical eigenvalue distribution of $A + W$ increasingly well as $N \to \infty$. We call ρ the *self-consistent density of states* (scDOS). Accordingly, its support $\mathrm{supp}\,\rho$ is called the *self-consistent spectrum*. It was proven in [Ajanki et al. 2015] that under very general conditions, $\rho(\mathrm{d}\tau)$ is an absolutely continuous measure with a $\frac{1}{3}$-Hölder continuous density, $\rho(\tau)$. Furthermore, the self-consistent spectrum consists of finitely many intervals with square root growth of ρ at the *edges*, i.e., at the points in $\partial\, \mathrm{supp}\,\rho$.

We call a point $\mathfrak{c} \in \mathbb{R}$ a cusp of ρ if $\mathfrak{c} \in \mathrm{int}\,\mathrm{supp}\,\rho$ and $\rho(\mathfrak{c}) = 0$. Cusps naturally emerge when we consider a one-parameter family of ensembles and two support intervals of ρ merge as the parameter value changes. The cusp universality phenomenon is not restricted to the exact cusp; it also occurs for situations shortly before and after the merging of two such support intervals, giving rise to a one-parameter family of universal statistics. More precisely, universality emerges if ρ has a *physical cusp*. The terminology indicates that all these singularities become indistinguishable from the exact cusp if the density is resolved with a local precision above the typical eigenvalue spacing. We say that ρ exhibits a physical cusp if it has a small gap $(\mathfrak{e}_-, \mathfrak{e}_+) \subset \mathbb{R} \setminus \mathrm{supp}\,\rho$ with $\mathfrak{e}_+, \mathfrak{e}_- \in \mathrm{supp}\,\rho$ in its support of size $\mathfrak{e}_+ - \mathfrak{e}_- \lesssim N^{-\frac{3}{4}}$ or a local minimum $\mathfrak{m} \in \mathrm{int}\,\mathrm{supp}\,\rho$ of size $\rho(\mathfrak{m}) \lesssim N^{-\frac{1}{4}}$; see Figure 1. Correspondingly, we call the points $\mathfrak{b} := \frac{1}{2}(\mathfrak{e}_+ + \mathfrak{e}_-)$ and $\mathfrak{b} := \mathfrak{m}$ *physical cusp points*, respectively. One of the simplest models exhibiting a physical cusp point is the deformed Wigner matrix

$$H = \mathrm{diag}(1, \dots, 1, -1, \dots, -1) + \sqrt{1+t}\,W, \tag{2-2}$$

with equal numbers of ± 1, and where W is a Wigner matrix of variance $E|w_{ij}|^2 = N^{-1}$. The ensemble H from (2-2) exhibits an exact cusp if $t = 0$ and a physical cusp if $|t| \lesssim N^{-\frac{1}{2}}$, with $t > 0$ corresponding to a

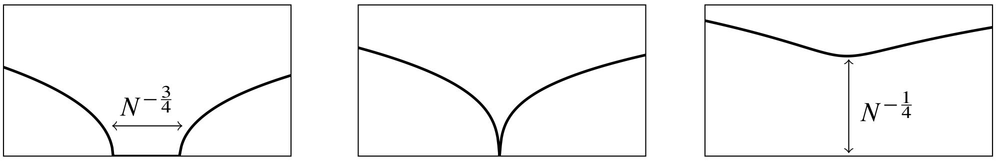

Figure 1. The cusp universality class can be observed in a one-parameter family of *physical cusps*.

small nonzero local minimum and $t < 0$ corresponding to a small gap in the support of the self-consistent density. For the proof of universality in the real symmetric symmetry class we will use (2-2) with $W \sim \text{GOE}$ as a Gaussian reference ensemble.

Our main result is cusp universality under the real symmetric analogues of the assumptions of [Erdős et al. 2018]. Throughout this paper we make the following three assumptions:

Assumption (A) (bounded moments). *The entries of the matrix $\sqrt{N}W$ have bounded moments and the expectation A is bounded*; *i.e., there are positive C_k such that*

$$|a_i| \le C_0, \qquad \mathbb{E}|w_{ij}|^k \le C_k N^{-\frac{1}{2}k}, \quad k \in \mathbb{N}.$$

Assumption (B) (flatness). *We assume that the matrix S is flat in the sense $s_{ij} = \mathbb{E}\, w_{ij}^2 \ge c/N$ for some constant $c > 0$.*

Assumption (C) (bounded self-consistent Green's function). *The scDOS ρ has a physical cusp point $\mathfrak{b}$, and in a neighbourhood of the physical cusp point $\mathfrak{b} \in \mathbb{R}$ the self-consistent Green's function is bounded*; *i.e., for positive C, κ we have*

$$|m_i(z)| \le C, \quad z \in [\mathfrak{b} - \kappa, \mathfrak{b} + \kappa] + \mathrm{i}\mathbb{R}^+.$$

We call the constants appearing in Assumptions (A)–(C) *model parameters*. All generic constants in this paper may implicitly depend on these model parameters. Dependence on further parameters, however, will be indicated.

Remark 2.1. The boundedness of $\boldsymbol{m}$ in Assumption (C) can be, for example, ensured by assuming some regularity of the variance matrix S. For more details we refer to [Ajanki et al. 2015, Chapter 6].

According to the extensive analysis in [Ajanki et al. 2015; Alt et al. 2018a] it follows[1] that there exists some small $\delta_* \sim 1$ such that the self-consistent density ρ around the points where it is small exhibits one of the following three types of behaviours:

(i) *Exact cusp*. There is a cusp point $\mathfrak{c} \in \mathbb{R}$ in the sense that $\rho(\mathfrak{c}) = 0$ and $\rho(\mathfrak{c} \pm \delta) > 0$ for $0 \neq \delta \ll 1$. In this case the self-consistent density is locally around $\mathfrak{c}$ given by

$$\rho(\mathfrak{c} + \omega) = \frac{\sqrt{3}\gamma^{\frac{4}{3}}|\omega|^{\frac{1}{3}}}{2\pi}[1 + \mathcal{O}(|\omega|^{\frac{1}{3}})] \tag{2-3a}$$

for $\omega \in [-\delta_*, \delta_*]$ and some $\gamma > 0$.

[1]The claimed expansions (2-3a) and (2-3d) follow directly from [Alt et al. 2018a, Theorem 7.2(c),(d)]. The error term in (2-3b) follows from [Alt et al. 2018a, Theorem 7.1(a)], where we define γ according to h therein.

(2) *Small gap.* There is a maximal interval $[\mathfrak{e}_-, \mathfrak{e}_+]$ of size $0 < \Delta := \mathfrak{e}_+ - \mathfrak{e}_- \ll 1$ such that $\rho|_{[\mathfrak{e}_-,\mathfrak{e}_+]} \equiv 0$. In this case the density around $\mathfrak{e}_\pm$ is, for some $\gamma > 0$, locally given by

$$\rho(\mathfrak{e}_\pm \pm \omega) = \frac{\sqrt{3}(2\gamma)^{\frac{4}{3}}\Delta^{\frac{1}{3}}}{2\pi}\Psi_{\text{edge}}\Big(\frac{\omega}{\Delta}\Big)\Big[1 + \mathcal{O}\Big(\min\Big\{\omega^{\frac{1}{3}}, \frac{\omega^{\frac{1}{2}}}{\Delta^{\frac{1}{6}}}\Big\}\Big)\Big] \tag{2-3b}$$

for $\omega \in [0, \delta_*]$, where

$$\Psi_{\text{edge}}(\lambda) := \frac{\sqrt{\lambda(1+\lambda)}}{(1+2\lambda+2\sqrt{\lambda(1+\lambda)})^{\frac{2}{3}} + (1+2\lambda-2\sqrt{\lambda(1+\lambda)})^{\frac{2}{3}} + 1}, \quad \lambda \geq 0. \tag{2-3c}$$

(3) *Nonzero local minimum.* There is a local minimum at $\mathfrak{m} \in \mathbb{R}$ of ρ such that $0 < \rho(\mathfrak{m}) \ll 1$. In this case there exists some $\gamma > 0$ such that

$$\begin{aligned}&\rho(\mathfrak{m}+\omega)\\ &\quad = \rho(\mathfrak{m}) + \rho(\mathfrak{m})\Psi_{\min}\Big(\frac{3\sqrt{3}\gamma^4\omega}{2(\pi\rho(\mathfrak{m}))^3}\Big)\Big[1 + \mathcal{O}\Big(\min\Big\{\rho(\mathfrak{m})^{\frac{1}{2}}, \frac{\rho(\mathfrak{m})^4}{|\omega|}\Big\} + \min\Big\{\frac{\omega^2}{\rho(\mathfrak{m})^5}, |\omega|^{\frac{1}{3}}\Big\}\Big)\Big]\end{aligned} \tag{2-3d}$$

for $\omega \in [-\delta_*, \delta_*]$, where

$$\Psi_{\min}(\lambda) := \frac{\sqrt{1+\lambda^2}}{(\sqrt{1+\lambda^2}+\lambda)^{\frac{2}{3}} + (\sqrt{1+\lambda^2}-\lambda)^{\frac{2}{3}} - 1} - 1, \quad \lambda \in \mathbb{R}. \tag{2-3e}$$

We note that the choices for the *slope* parameter γ in (2-3b)–(2-3d) are consistent with (2-3a) in the sense that in the regimes $\Delta \ll \omega \ll 1$ and $\rho(\mathfrak{m})^3 \ll |\omega| \ll 1$ the respective formulae asymptotically agree. The precise form of the prefactors in (2-3) is also chosen such that in the universality statement γ is a linear rescaling parameter.

It is natural to express universality in terms of a rescaled k-point function $p_k^{(N)}$ which we define implicitly by

$$\mathbb{E}\binom{N}{k}^{-1} \sum_{\{i_1,\ldots,i_k\}\subset[N]} f(\lambda_{i_1},\ldots,\lambda_{i_k}) = \int_{\mathbb{R}^k} f(\boldsymbol{x})p_k^{(N)}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{2-4}$$

for test functions f, where the summation is over all subsets of k distinct integers from $[N]$.

Theorem 2.2. *Let H be a real symmetric or complex Hermitian deformed Wigner-type matrix whose scDOS ρ has a physical cusp point $\mathfrak{b}$ such that Assumptions* (A)–(C) *are satisfied. Let $\gamma > 0$ be the slope parameter at $\mathfrak{b}$, i.e., such that ρ is locally around $\mathfrak{b}$ given by* (2-3)*. Then the local k-point correlation function at $\mathfrak{b}$ is universal*; *i.e., for any $k \in \mathbb{N}$ there exists a k-point correlation function $p_{k,\alpha}^{\mathrm{GOE/GUE}}$ such that for any test function $F \in C_c^1(\overline{\Omega})$, with $\Omega \subset \mathbb{R}^k$ some bounded open set, it holds that*

$$\int_{\mathbb{R}^k} F(\boldsymbol{x})\Big[\frac{N^{\frac{1}{4}k}}{\gamma^k}p_k^{(N)}\Big(\mathfrak{b}+\frac{\boldsymbol{x}}{\gamma N^{\frac{3}{4}}}\Big) - p_{k,\alpha}^{\mathrm{GOE/GUE}}(\boldsymbol{x})\Big]\mathrm{d}\boldsymbol{x} = \mathcal{O}_{k,\Omega}(N^{-c(k)}\|F\|_{C^1}),$$

where the parameter α and the physical cusp $\mathfrak{b}$ are given by

$$\alpha := \begin{cases} 0 & \textit{in case (i)},\\ 3\big(\frac{1}{4}\gamma\Delta\big)^{\frac{2}{3}}N^{\frac{1}{2}} & \textit{in case (ii)},\\ -(\pi\rho(\mathfrak{m})/\gamma)^2N^{\frac{1}{2}} & \textit{in case (iii)},\end{cases} \qquad \mathfrak{b} := \begin{cases} \mathfrak{c} & \textit{in case (i)},\\ \frac{1}{2}(\mathfrak{e}_-+\mathfrak{e}_+) & \textit{in case (ii)},\\ \mathfrak{m} & \textit{in case (iii)},\end{cases} \tag{2-5}$$

and $c(k)>0$ is a small constant only depending on k. The implicit constant in the error term depends on k and the diameter of the set Ω.

Remark 2.3. (i) In the complex Hermitian symmetry class the k-point function is given by

$$p_{k,\alpha}^{\mathrm{GUE}}(\boldsymbol{x}) = \det(K_{\alpha,\alpha}(x_i, x_j))_{i,j=1}^k.$$

Here the *extended Pearcey kernel* $K_{\alpha,\beta}$ is given by

$$K_{\alpha,\beta}(x,y) = \frac{1}{(2\pi \mathrm{i})^2}\int_\Xi \mathrm{d}z \int_\Phi \mathrm{d}w \, \frac{\exp\bigl(-\frac14 w^4 + \frac12 \beta w^2 - yw + \frac14 z^4 - \frac12 \alpha z^2 + xz\bigr)}{w-z} - \frac{\mathbb{1}_{\beta>\alpha}}{\sqrt{2\pi(\beta-\alpha)}} \exp\Bigl(-\frac{(y-x)^2}{2(\beta-\alpha)}\Bigr), \tag{2-6}$$

where Ξ is a contour consisting of rays from $\pm\infty e^{\frac14 \mathrm{i}\pi}$ to 0 and rays from 0 to $\pm\infty e^{-\frac14 \mathrm{i}\pi}$, and Φ is the ray from $-\mathrm{i}\infty$ to $\mathrm{i}\infty$. For more details we refer to [Tracy and Widom 2006; Adler et al. 2010; Brézin and Hikami 1998b].

(ii) The real symmetric k-point function (possibly only a distribution) $p_{k,\alpha}^{\mathrm{GOE}}$ is not known explicitly. In fact, it is not even known whether $p_{k,\alpha}^{\mathrm{GOE}}$ is Pfaffian. We will nevertheless establish the existence of $p_{k,\alpha}^{\mathrm{GOE}}$ as a distribution in the dual of the C^1 functions in Section 3 as the limit of the correlation functions of a one-parameter family of Gaussian comparison models.

Theorem 2.2 is a universality result about the spatial correlations of eigenvalues. Our method also allows us to prove the corresponding statement on space-time universality when we consider the time evolution of eigenvalues $(\lambda_i^t)_{i\in[N]}$ according to the Dyson Brownian motion $\mathrm{d}H^{(t)} = \mathrm{d}\mathfrak{B}_t$ with initial condition $H^{(0)} = H$, where, depending on the symmetry class, $\mathfrak{B}_t$ is a complex Hermitian or real symmetric matrix-valued Brownian motion. For any ordered k-tuple $\boldsymbol{\tau} = (\tau_1,\dots,\tau_k)$ with $0\le\tau_1\le\dots\le\tau_k\lesssim N^{-\frac12}$ we then define the *time-dependent k-point function* as follows. Denote the unique values in the tuple $\boldsymbol{\tau}$ by $\sigma_1<\dots<\sigma_l$ such that $\{\tau_1,\dots,\tau_k\}=\{\sigma_1,\dots,\sigma_l\}$ and denote the multiplicity of σ_j in $\boldsymbol{\tau}$ by k_j and note that $\sum k_j = k$. We then define $p_{k,\boldsymbol{\tau}}^{(N)}$ implicitly via

$$\mathbb{E}\prod_{j=1}^l \biggl[\binom{N}{k_j}^{-1} \sum_{\{i_1^j,\dots,i_{k_j}^j\}\subset[N]}\biggr] f(\lambda_{i_1^1}^{\sigma_1},\dots,\lambda_{i_{k_1}^1}^{\sigma_1},\dots,\lambda_{i_1^l}^{\sigma_l},\dots,\lambda_{i_{k_l}^l}^{\sigma_l}) = \int_{\mathbb{R}^k} f(\boldsymbol{x}) p_{k,\boldsymbol{\tau}}^{(N)}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{2-7}$$

for test functions f and note that (2-7) reduces to (2-4) in the case $\tau_1=\dots=\tau_k=0$. We note that in (2-7) coinciding indices are allowed only for eigenvalues at different times. If the scDOS ρ of H has a physical cusp in $\mathfrak{b}$, then for $\tau\lesssim N^{-\frac12}$ the scDOS ρ_τ of $H^{(\tau)}$ also has a physical cusp $\mathfrak{b}_\tau$ close to $\mathfrak{b}$ and we can prove space-time universality in the sense of the following theorem, whose proof we defer to Appendix A.

Theorem 2.4. *Let H be a real symmetric or complex Hermitian deformed Wigner-type matrix whose scDOS ρ has a physical cusp point $\mathfrak{b}$ such that Assumptions (A)–(C) are satisfied. Let $\gamma>0$ be the slope parameter at $\mathfrak{b}$, i.e., such that ρ is locally around $\mathfrak{b}$ given by (2-3). Then there exists a k-point correlation* *function* $p_{k,\boldsymbol{\alpha}}^{\mathrm{GOE/GUE}}$ *such that for any* $0 \le \tau_1 \le \cdots \le \tau_k \lesssim N^{-\frac{1}{2}}$ *and for any test function* $F \in C_c^1(\overline{\Omega})$, *with* $\Omega \subset \mathbb{R}^k$ *some bounded open set, it holds that*

$$\int_{\mathbb{R}^k} F(\boldsymbol{x})\Big[\frac{N^{\frac{1}{4}k}}{\gamma^k} p_{k,\boldsymbol{\tau}/\gamma^2}^{(N)}\Big(\mathfrak{b}_{\boldsymbol{\tau}/\gamma^2} + \frac{\boldsymbol{x}}{\gamma N^{\frac{3}{4}}}\Big) - p_{k,\boldsymbol{\alpha}}^{\mathrm{GOE/GUE}}(\boldsymbol{x})\Big]\,\mathrm{d}\boldsymbol{x} = \mathcal{O}_{k,\Omega}(N^{-c(k)}\|F\|_{C^1}),$$

where $\boldsymbol{\tau} = (\tau_1, \dots, \tau_k)$, $\mathfrak{b}_{\boldsymbol{\tau}} = (\mathfrak{b}_{\tau_1}, \dots, \mathfrak{b}_{\tau_k})$, $\boldsymbol{\alpha} = \alpha - \boldsymbol{\tau} N^{\frac{1}{2}}$ *with* α *from* (2-5), *and* $c(k) > 0$ *is a small constant only depending on* k. *In the case of the complex Hermitian symmetry class the* k*-point correlation function is known to be determinantal of the form*

$$p_{\alpha_1,\dots,\alpha_k}^{\mathrm{GUE}}(\boldsymbol{x}) = \det(K_{\alpha_i,\alpha_j}(x_i, x_j))_{i,j=1}^k,$$

with $K_{\alpha,\beta}$ *as in* (2-6).

The analogous version of Theorem 2.4 for fixed energy bulk multitime universality has been proven in [Landon et al. 2019, Section 2.3.1].

Remark 2.5. The extended Pearcey kernel $K_{\alpha,\beta}$ in Theorem 2.4 has already been observed for the double-scaling limit of nonintersecting Brownian bridges [Adler et al. 2010; Tracy and Widom 2006]. However, in the random matrix setting our methods also allow us to prove that the space-time universality of Theorem 2.4 extends beyond the Gaussian DBM flow. If the times $0 \le \tau_1 \le \cdots \le \tau_k \lesssim N^{-\frac{1}{2}}$ are ordered, then the k-point correlation function of the DBM flow asymptotically agrees with the k-point correlation function of eigenvalues of the matrices

$$H + \sqrt{\tau_1}W_1, \quad H + \sqrt{\tau_1}W_1 + \sqrt{\tau_2 - \tau_1}W_2, \quad \dots, \quad H + \sqrt{\tau_1}W_1 + \cdots + \sqrt{\tau_k - \tau_{k-1}}W_k$$

for independent standard Wigner matrices $W_1, \dots, W_k$.

3. Ornstein–Uhlenbeck flow

Starting from this section we consider a more general framework that allows for random matrix ensembles with certain correlation among the entries. In this way we stress that our proofs regarding the semicircular flow and the Dyson Brownian motion are largely model-independent, assuming the optimal local law holds. The independence assumption on the entries of W is made only because we rely on the local law from [Erdős et al. 2018] that was proven for deformed Wigner-type matrices. We therefore present the flow directly in the more general framework of the *matrix Dyson equation* (MDE)

$$1 + (z - A + \mathcal{S}[M(z)])M(z) = 0, \qquad A := \mathbb{E}\,H, \quad \mathcal{S}[R] := \mathbb{E}\,WRW, \tag{3-1}$$

with spectral parameter in the complex upper half-plane, $\Im z > 0$, and positive definite imaginary part, $\frac{1}{2\mathrm{i}}(M(z) - M(z)^*) > 0$, of the solution M. The MDE generalises (2-1). Note that in the deformed Wigner-type case the *self-energy operator* $\mathcal{S}\colon \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$ is related to the variance matrix S by $\mathcal{S}[\operatorname{diag}\boldsymbol{r}] = \operatorname{diag}(S\boldsymbol{r})$.

As in [Erdős et al. 2018] we consider the Ornstein–Uhlenbeck flow

$$\mathrm{d}\widetilde{H}_s = -\tfrac{1}{2}(\widetilde{H}_s - A)\,\mathrm{d}s + \Sigma^{\frac{1}{2}}[\mathrm{d}\mathfrak{B}_s], \qquad \Sigma[R] := \tfrac{1}{2}\beta\,\mathbb{E}\,W\operatorname{Tr}WR, \quad \widetilde{H}_0 := H, \tag{3-2}$$

which preserves expectation and the self-energy operator $\mathcal{S}$. Since we consider real symmetric H, the parameter β indicating the symmetry class is $\beta = 1$. In (3-2) by $\mathfrak{B}_s \in \mathbb{R}^{N\times N}$ we denote a real symmetric matrix-valued standard (GOE) Brownian motion; i.e., $(\mathfrak{B}_s)_{ij}$ for $i < j$ and $(\mathfrak{B}_s)_{ii}/\sqrt{2}$ are independent standard Brownian motions and $(\mathfrak{B}_s)_{ji} = (\mathfrak{B}_s)_{ij}$. If H were complex Hermitian, we would have $\beta = 2$ and $\mathrm{d}\mathfrak{B}_s$ would be an infinitesimal GUE matrix. This was the setting in [Erdős et al. 2018]. The OU flow effectively adds a small Gaussian component of size $\sqrt{s}$ to $\widetilde{H}_s$. More precisely, we can construct a Wigner-type matrix H_s, satisfying Assumptions (A)–(C), such that, for any fixed s,

$$\widetilde{H}_s = H_s + \sqrt{cs}U, \quad \mathcal{S}_s = \mathcal{S} - cs\mathcal{S}^{\mathrm{GOE}}, \quad \mathbb{E}\, H_s = A, \quad U \sim \mathrm{GOE}, \tag{3-3}$$

where U is independent of H_s. Here $c > 0$ is a small universal constant which depends on the constant in Assumption (B), $\mathcal{S}_s$ is the self-energy operator corresponding to H_s and $\mathcal{S}^{\mathrm{GOE}}[R] := \langle R\rangle + R^t/N$, where $\langle \cdot \rangle := N^{-1}\operatorname{Tr}(\cdot)$ and R^t denotes the transpose of R. Since $\mathcal{S}$ is flat in the sense $\mathcal{S}[R] \gtrsim \langle R\rangle$ and s is small, it follows that also $\mathcal{S}_s$ is flat.

As a consequence of the well-established Green function comparison technique the k-point function of $H = \widetilde{H}_0$ is comparable with the one of $\widetilde{H}_s$ as long as $s \le N^{-\frac{1}{4}-\epsilon}$ for some $\epsilon > 0$. Indeed, from [Erdős et al. 2018, equation (116)] for any $F \in C_c^1(\overline{\Omega})$, a compactly supported C^1 test function on a bounded open set $\Omega \subset \mathbb{R}^k$, we find

$$\int_{\mathbb{R}^k} F(\boldsymbol{x}) N^{\frac{1}{4}k}\left[p_k^{(N)}\left(\mathfrak{b} + \frac{\boldsymbol{x}}{\gamma N^{\frac{3}{4}}}\right) - \tilde{p}_{k,s}^{(N)}\left(\mathfrak{b} + \frac{\boldsymbol{x}}{\gamma N^{\frac{3}{4}}}\right)\right]\mathrm{d}\boldsymbol{x} = \mathcal{O}_{k,\Omega}(N^{-c}\|F\|_{C^1}), \tag{3-4}$$

where $\tilde{p}_{k,s}^{(N)}$ is the k-point correlation function of $\widetilde{H}_s$ and $c = c(k) > 0$ is some constant.

It follows from the flatness assumption that the matrix H_s satisfies the assumptions of the local law from [Erdős et al. 2018, Theorem 2.5] uniformly in $s \ll 1$. Therefore [Erdős et al. 2018, Corollary 2.6] implies that the eigenvalues of H_s are rigid down to the optimal scale. It remains to prove that for long enough times s the local eigenvalue statistics of $H_s + \sqrt{cs}U$ on a scale of $1/\gamma N^{\frac{3}{4}}$ around $\mathfrak{b}$ agree with the local eigenvalue statistics of the Gaussian reference ensemble around 0 at a scale of $1/N^{\frac{3}{4}}$. By a simple rescaling, Theorem 2.2 then follows from (3-4) together with the following proposition.

Proposition 3.1. *Let $t_1 := N^{-\frac{1}{2}+\omega_1}$ with some small $\omega_1 > 0$ and let t_* be such that $|t_* - t_1| \lesssim N^{-\frac{1}{2}}$. Assume that $H^{(\lambda)}$ and $H^{(\mu)}$ are Wigner-type matrices[2] satisfying Assumptions (A)–(C) such that the scDOSs $\rho_{\lambda,t_*}, \rho_{\mu,t_*}$ of $H^{(\lambda)} + \sqrt{t_*}U^{(\lambda)}$ and $H^{(\mu)} + \sqrt{t_*}U^{(\mu)}$ with independent $U^{(\lambda)}, U^{(\mu)} \sim \mathrm{GOE}$ have cusps in some points $\mathfrak{c}_\lambda$, $\mathfrak{c}_\mu$ such that locally around $\mathfrak{c}_r$, $r = \lambda, \mu$, the densities ρ_{r,t_*} are given by (2-3a) with $\gamma = 1$. Then the local k-point correlation functions $p_{k,t_1}^{(N,r)}$ of $H^{(r)} + \sqrt{t_1}U^{(r)}$ around the respective physical cusps $\mathfrak{b}_{r,t_1}$ of ρ_{r,t_1}, $j = 1, 2$, asymptotically agree in the sense*

$$\int_{\mathbb{R}^k} F(\boldsymbol{x})\left[N^{\frac{1}{4}k} p_{k,t_1}^{(N,\lambda)}\left(\mathfrak{b}_{\lambda,t_1} + \frac{\boldsymbol{x}}{N^{\frac{3}{4}}}\right) - N^{\frac{1}{4}k} p_{k,t_1}^{(N,\mu)}\left(\mathfrak{b}_{\mu,t_1} + \frac{\boldsymbol{x}}{N^{\frac{3}{4}}}\right)\right]\mathrm{d}\boldsymbol{x} = \mathcal{O}_{k,\Omega}(N^{-c(k)}\|F\|_{C^1})$$

for any $F \in C_c^1(\overline{\Omega})$, with $\Omega \subset \mathbb{R}^k$ a bounded open set.

[2]We use the notation $H^{(\lambda)}$ and $H^{(\mu)}$ since we denote the eigenvalues of $H^{(\lambda)}$ and $H^{(\mu)}$ by λ_i and μ_i respectively, with $1 \le i \le N$.

Proof of Theorem 2.2. Set $s := t_1/c\theta^2$ and $H^{(\lambda)} := \theta H_s$, where c is the constant from (3-3) and $\theta \sim 1$ is yet to be chosen. Note that $H^{(\lambda)} + \sqrt{t}U = \theta(H_s + \sqrt{t/\theta^2}U)$, and in particular $H^{(\lambda)} + \sqrt{t_1}U = \widetilde{H}_s$. Moreover, it follows from the semicircular flow analysis in Section 4 that for some t_* with $|t_* - t_1| \lesssim N^{-\frac{1}{2}}$, the scDOS $\theta\rho_{\lambda,t_*}(\lambda\,\cdot\,)$ of $H_s + \sqrt{t_*/\theta^2}U$ and thereby also ρ_{λ,t_*}, the one of $H^{(\lambda)} + \sqrt{t_*}U$, have exact cusps in $\mathfrak{c}_\lambda/\theta$ and $\mathfrak{c}_\lambda$, respectively. It follows from the $\frac{1}{3}$-Hölder continuity of the slope parameter, see [Alt et al. 2018a, Lemma 10.5, equation (7.5a)], that locally around $\mathfrak{c}_\lambda/\theta$ the scDOS of $H_s + \sqrt{t_*/\theta^2}U$ is given by

$$\theta\rho_{\lambda,t_*}(\mathfrak{c}_\lambda + \theta\omega) = \theta\rho_{\lambda,t_*}\left(\theta\left(\frac{\mathfrak{c}_\lambda}{\theta} + \omega\right)\right) = \frac{\sqrt{3}\gamma^{\frac{4}{3}}|\omega|^{\frac{1}{3}}}{2\pi}[1 + \mathcal{O}(|\omega|^{\frac{1}{3}} + |t_* - t_1|^{\frac{1}{3}})].$$

Hence we can choose $\theta = \gamma[1 + \mathcal{O}(|t_1 - t_*|^{\frac{1}{3}})]$ appropriately such that

$$\rho_{\lambda,t_*}(\mathfrak{c}_\lambda + \omega) = \frac{\sqrt{3}|\omega|^{\frac{1}{3}}}{2\pi}[1 + \mathcal{O}(|\omega|^{\frac{1}{3}})]$$

and it follows that $H^{(\lambda)}$ satisfies the assumptions of Proposition 3.1; in particular the slope parameter of $H^{(\lambda)} + \sqrt{t_*}U$ is normalised to 1. Furthermore, the almost cusp $\mathfrak{b}_{\lambda,t_1}$ of $H^{(\lambda)} + \sqrt{t_1}U$ is given by $\mathfrak{b}_{\lambda,t_1} = \theta\mathfrak{b}$ with $\mathfrak{b}$ as in Theorem 2.2.

We now choose our Gaussian comparison model. For $\alpha \in \mathbb{R}$ we consider the *reference ensemble*

$$U_\alpha = U_\alpha^{(N)} := \operatorname{diag}(1, \dots, 1, -1, \dots, -1) + \sqrt{1 - \alpha N^{-\frac{1}{2}}}U \in \mathbb{R}^{N\times N}, \quad U \sim \mathrm{GOE}, \tag{3-5}$$

with $\lfloor \frac{1}{2}N \rfloor$ and $\lceil \frac{1}{2}N \rceil$ times ± 1 in the deterministic diagonal. An elementary computation shows that for even N and $\alpha = 0$, the self-consistent density of U_α has an exact cusp of slope $\gamma = 1$ in $\mathfrak{c} = 0$; i.e., it is given by (2-3a). For odd N the exact cusp is at distance $\lesssim N^{-1}$ away from 0, which is well below the natural scale of order $N^{-\frac{3}{4}}$ of the eigenvalue fluctuation and therefore has no influence on the k-point correlation function. The reference ensemble U_α has for $0 \neq |\alpha| \sim 1$ a small gap of size $N^{-\frac{3}{4}}$ or small local minimum of size $N^{-\frac{1}{4}}$ at the physical cusp point $|\mathfrak{b}| \lesssim 1/N$, depending on the sign of α. Using the definition in (3-5), let $H^{(\mu)} := U_{N^{1/2}t_*}$, from which it follows that $H^{(\mu)} + \sqrt{t_*}U \sim U_0$ has an exact cusp in 0 whose slope is 1 by an easy explicit computation in the case of even N. For odd N the cusp emerges at a distance of $\lesssim N^{-1}$ away from 0, which is well below the investigated scale. Thus also $H^{(2)}$ satisfies the assumptions of Proposition 3.1. The almost cusp $\mathfrak{b}_{\mu,t_1}$ is given by $\mathfrak{b}_{\mu,t_1} = 0$ by symmetry of the density ρ_{μ,t_1} in the case of even N and at a distance of $|\mathfrak{b}_{\mu,t_1}| \lesssim N^{-1}$ in the case of odd N. This fact follows, for example, from explicitly solving a quadratic equation in two variables. The perturbation of size $1/N$ is not visible on the scale of the k-point correlation functions.

Now Proposition 3.1 together with (3-4) and $s \sim N^{-\frac{1}{2}+\omega_1}$ implies

$$\int_{\mathbb{R}^k} F(\boldsymbol{x})\left[\frac{N^{\frac{1}{4}k}}{\theta^k} p_k^{(N)}\left(\mathfrak{b} + \frac{\boldsymbol{x}}{\theta N^{\frac{3}{4}}}\right) - N^{\frac{1}{4}k} p_{k,\alpha,\mathrm{GOE}}^{(N)}\left(\frac{\boldsymbol{x}}{N^{\frac{3}{4}}}\right)\right] \mathrm{d}\boldsymbol{x} = \mathcal{O}_{k,\Omega}(N^{-c(k)}\|F\|_{C^1(\Omega)}), \tag{3-6}$$

with $\alpha = N^{\frac{1}{2}}(t_* - t_1)$, where $p_{k,\alpha,\mathrm{GOE}}^{(N)}$ denotes the k-point function of the comparison model U_α. This completes the proof of Theorem 2.2 modulo the comparison of $p_{k,\alpha,\mathrm{GOE}}^{(N)}$ with its limit by relating $t_* - t_1$ to the size of the gap and the local minimum of ρ via [Erdős et al. 2018, Lemma 5.1] (or (4-6a)–(4-6c)) and recalling that $\theta = \gamma[1 + \mathcal{O}(|t_1 - t_*|^{\frac{1}{3}})]$.

To complete the proof we claim that for any fixed k and α there exists a distribution $p^{\mathrm{GOE}}_{k,\alpha}$ on $\mathbb{R}^k$, locally in the dual of $C^1_c(\overline{\Omega})$ for every open bounded $\Omega \subset \mathbb{R}^k$, such that

$$\int_{\mathbb{R}^k} F(\boldsymbol{x}) \left[N^{\frac{1}{4}k} p^{(N)}_{k,\alpha,\mathrm{GOE}}\left(\frac{\boldsymbol{x}}{N^{\frac{3}{4}}}\right) - p^{\mathrm{GOE}}_{k,\alpha}(\boldsymbol{x}) \right] \mathrm{d}\boldsymbol{x} = \mathcal{O}_{k,\Omega}(N^{-c(k)} \|F\|_{C^1}) \tag{3-7}$$

holds for any $F \in C^1_c(\Omega)$. We now show that (3-7) is a straightforward consequence of (3-6).

First notice that, for notational simplicity, we gave the proof of (3-6) only for the case when H and U_α are of the same dimension, but it works without any modification when their dimensions are only comparable; see Remark 5.2. Hence, applying this result to a sequence of GOE ensembles $U^{(N_n)}_\alpha$ with $N_n := \left(\frac{4}{3}\right)^n$, for any compactly supported $F \in C^1_c(\overline{\Omega})$ we have

$$\int_{\mathbb{R}^k} F(\boldsymbol{x}) \left[N_n^{\frac{1}{4}k} p^{(N_n)}_{k,\alpha,\mathrm{GOE}}\left(\frac{\boldsymbol{x}}{N_n^{\frac{3}{4}}}\right) - N_{n+1}^{\frac{1}{4}k} p^{(N_{n+1})}_{k,\alpha,\mathrm{GOE}}\left(\frac{\boldsymbol{x}}{N_{n+1}^{\frac{3}{4}}}\right) \right] \mathrm{d}\boldsymbol{x} = \mathcal{O}_{k,\Omega}\left(\left(\tfrac{3}{4}\right)^{nc(k)} \|F\|_{C^1}\right). \tag{3-8}$$

Fix a bounded open set $\Omega \subset \mathbb{R}^k$ and define the sequence of functionals $\{\mathcal{J}_n\}_{n\in\mathbb{N}}$ in the dual space $C^1_c(\overline{\Omega})^*$ as

$$\mathcal{J}_n(F) := \int_{\mathbb{R}^k} F(\boldsymbol{x}) N_n^{\frac{1}{4}k} p^{(N_n)}_{k,\alpha,\mathrm{GOE}}\left(\frac{\boldsymbol{x}}{N_n^{\frac{3}{4}}}\right) \mathrm{d}\boldsymbol{x}$$

for any $F \in C^1_c(\overline{\Omega})$. Then, by (3-8) it easily follows that $\{\mathcal{J}_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence on $C^1_c(\overline{\Omega})^*$. Indeed, for any $M > L$ we have by a telescopic sum

$$\begin{aligned} |(\mathcal{J}_M - \mathcal{J}_L)(F)| &= \left| \sum_{n=L}^{M-1} \int_{\mathbb{R}^k} F(\boldsymbol{x}) \left[N_{n+1}^{\frac{1}{4}k} p^{(N_{n+1})}_{k,\alpha,\mathrm{GOE}}\left(\frac{\boldsymbol{x}}{N_{n+1}^{\frac{3}{4}}}\right) - N_n^{\frac{1}{4}k} p^{(N_n)}_{k,\alpha,\mathrm{GOE}}\left(\frac{\boldsymbol{x}}{N_n^{\frac{3}{4}}}\right) \right] \mathrm{d}\boldsymbol{x} \right| \\ &\le C_{k,\Omega}\left(\tfrac{3}{4}\right)^{Lc(k)} \|F\|_{C^1}. \end{aligned} \tag{3-9}$$

Thus, we conclude that there exists a unique $\mathcal{J}_\infty \in C^1_c(\overline{\Omega})^*$ such that $\mathcal{J}_n \to \mathcal{J}_\infty$ as $n \to \infty$ in norm. Then, (3-9) clearly concludes the proof of (3-7), identifying $\mathcal{J}_\infty = \mathcal{J}^{(\Omega)}_\infty$ with $p^{\mathrm{GOE}}_{k,\alpha}$ restricted to $\overline{\Omega}$. Since this holds for any open bounded set $\Omega \subset \mathbb{R}^k$, the distribution $p^{\mathrm{GOE}}_{k,\alpha}$ can be identified with the inductive limit of the consistent family of functionals $\{\mathcal{J}^{(\Omega_m)}_\infty\}_{m\ge 1}$, where, say, Ω_m is the ball of radius m. This completes the proof of Theorem 2.2. □

4. Semicircular flow analysis

In this section we analyse various properties of the semicircular flow in order to prepare the Dyson Brownian motion argument in Sections 6 and 7. If ρ is a probability density on $\mathbb{R}$ with Stieltjes transform m, then the free semicircular evolution $\rho^{\mathrm{fc}}_t = \rho \boxplus \sqrt{t}\rho_{\mathrm{sc}}$ of ρ is defined as the unique probability measure whose Stieltjes transform m^{fc}_t solves the implicit equation

$$m^{\mathrm{fc}}_t(\zeta) = m(\zeta + t m^{\mathrm{fc}}_t(\zeta)), \quad \zeta \in \mathbb{H},\ t \ge 0. \tag{4-1}$$

Here $\sqrt{t}\rho_{\mathrm{sc}}$ is the semicircular distribution of variance t.

We now prepare the Dyson Brownian motion argument in Section 7 by providing a detailed analysis of the scDOS along the semicircular flow. As in Proposition 3.1 we consider the setting of two densities ρ_λ, ρ_μ

whose semicircular evolutions reach a cusp of the same slope at the same time. Within the whole section we shall assume the following setup: Let ρ_λ, ρ_μ be densities associated with solutions M_λ, M_μ to some Dyson equations satisfying Assumptions (A)–(C) (or their matrix counterparts). We consider the free convolutions $\rho_{\lambda,t} := \rho_\lambda \boxplus \sqrt{t}\rho_{\mathrm{sc}}$ and $\rho_{\mu,t} := \rho_\mu \boxplus \sqrt{t}\rho_{\mathrm{sc}}$ of ρ_λ, ρ_μ with semicircular distributions of variance t and assume that after a time $t_* \sim N^{-\frac{1}{2}+\omega_1}$ both densities $\rho_{\lambda,t_*}, \rho_{\mu,t_*}$ have cusps in points $\mathfrak{c}_\lambda, \mathfrak{c}_\mu$ around which they can be approximated by (2-3a) with the same $\gamma = \gamma_\lambda(t_*) = \gamma_\mu(t_*)$. It follows from the semicircular flow analysis in [Erdős et al. 2018, Lemma 5.1] that for $0 \le t \le t_*$ both densities have small gaps $[\mathfrak{e}^-_{r,t}, \mathfrak{e}^+_{r,t}]$, $r = \lambda, \mu$ in their supports, while for $t_* \le t \le 2t_*$ they have nonzero local minima in some points $\mathfrak{m}_{r,t}$, $r = \lambda, \mu$. Instead of comparing the eigenvalue flows corresponding to ρ_λ, ρ_μ directly, we rather consider a continuous interpolation ρ_α for $\alpha \in [0, 1]$ of ρ_λ and ρ_μ. For technical reasons we define this interpolated density $\rho_{\alpha,t}$ as an interpolation of $\rho_{\lambda,t}$ and $\rho_{\mu,t}$ separately for each time t, rather than considering the evolution $\rho_{\alpha,0} \boxplus \sqrt{t}\rho_{\mathrm{sc}}$ of the initial interpolation $\rho_{\alpha,0}$. We warn the reader that semicircular evolution and interpolation do not commute; i.e., $\rho_{\alpha,t} \ne \rho_{\alpha,0} \boxplus \sqrt{t}\rho_{\mathrm{sc}}$. We now define the concept of *interpolating densities* following [Landon and Yau 2017, Section 3.1.1].

Definition 4.1. For $\alpha \in [0, 1]$ define the *α-interpolating density* $\rho_{\alpha,t}$ as follows. For any $0 \le E \le \delta_*$ and $r = \lambda, \mu$ let

$$n_{r,t}(E) := \int_{\mathfrak{e}^+_{r,t}}^{\mathfrak{e}^+_{r,t}+E} \rho_{r,t}(\omega)\,\mathrm{d}\omega, \quad 0 \le t \le t_*, \qquad n_{r,t}(E) := \int_{\mathfrak{m}_{r,t}}^{\mathfrak{m}_{r,t}+E} \rho_{r,t}(\omega)\,\mathrm{d}\omega, \quad t_* \le t \le 2t_*,$$

be the counting functions and $\varphi_{\lambda,t}, \varphi_{\mu,t}$ their inverses; i.e., $n_{r,t}(\varphi_{r,t}(s)) = s$. Define now

$$\varphi_{\alpha,t}(s) := \alpha\varphi_{\lambda,t}(s) + (1-\alpha)\varphi_{\mu,t}(s) \tag{4-2}$$

for $s \in [0, \delta_{**}]$, where $\delta_{**} \sim 1$ depends on δ_* and is chosen in such a way that $\varphi_{\alpha,t}$ is invertible.[3] We thus define $n_{\alpha,t}(E)$ to be the inverse of $\varphi_{\alpha,t}(s)$ near zero. Furthermore, for $0 \le t \le t_*$ set

$$\mathfrak{e}^\pm_{\alpha,t} := \alpha\mathfrak{e}^\pm_{\lambda,t} + (1-\alpha)\mathfrak{e}^\pm_{\mu,t}, \tag{4-3}$$

$$\rho_{\alpha,t}(\mathfrak{e}^+_{\alpha,t} + E) := \frac{\mathrm{d}}{\mathrm{d}E} n_{\alpha,t}(E), \quad E \in [0, \delta_*], \tag{4-4}$$

and for $t \ge t_*$ set

$$\begin{aligned} \mathfrak{m}_{\alpha,t} &:= \alpha\mathfrak{m}_{\lambda,t} + (1-\alpha)\mathfrak{m}_{\mu,t}, \\ \rho_{\alpha,t}(\mathfrak{m}_{\alpha,t} + E) &:= \alpha\rho_{\lambda,t}(\mathfrak{m}_{\lambda,t}) + (1-\alpha)\rho_{\mu,t}(\mathfrak{m}_{\mu,t}) + \frac{\mathrm{d}}{\mathrm{d}E} n_{\alpha,t}(E), \quad E \in [-\delta_*, \delta_*]. \end{aligned} \tag{4-5}$$

We define $\rho_{\alpha,t}(E)$ for $0 \le t \le t_*$ and $E \in [\mathfrak{e}^-_{\alpha,t} - \delta_*, \mathfrak{e}^-_{\alpha,t}]$ analogously.

The motivation for the interpolation mode in Definition 4.1 is that (4-2) ensures that the quantiles of $\rho_{\alpha,t}$ are the convex combination of the quantiles of $\rho_{\lambda,t}$ and $\rho_{\mu,t}$; see (4-13c). The following two lemmas collect various properties of the interpolating density. Recall that $\rho_{\lambda,t}$ and $\rho_{\mu,t}$ are asymptotically close near the cusp regime, up to a trivial shift, since they develop a cusp with the same slope at the same time.

[3] Invertibility in a small neighbourhood follows from the form of the explicit shape functions in (2-3b) and (2-3d)

In Lemma 4.2 we show that $\rho_{\alpha,t}$ shares this property. Lemma 4.3 shows that $\rho_{\alpha,t}$ inherits the regularity properties of $\rho_{\lambda,t}$ and $\rho_{\mu,t}$ from [Alt et al. 2018a].

Lemma 4.2 (size of gaps and minima along the flow). *For $t \le t_*$ and $r = \alpha, \lambda, \mu$ the supports of $\rho_{r,t}$ have small gaps $[\mathfrak{e}^-_{r,t}, \mathfrak{e}^+_{r,t}]$ near $\mathfrak{c}_*$ of size*

$$\Delta_{r,t} := \mathfrak{e}^+_{r,t} - \mathfrak{e}^-_{r,t} = (2\gamma)^2\Big(\frac{t_* - t}{3}\Big)^{\frac{3}{2}}\big[1 + \mathcal{O}((t_* - t)^{\frac{1}{3}})\big], \quad \Delta_{r,t} = \Delta_{\mu,t}[1 + \mathcal{O}((t_* - t)^{\frac{1}{3}})], \tag{4-6a}$$

and the densities are close in the sense

$$\rho_{r,t}(\mathfrak{e}^\pm_{r,t} \pm \omega) = \rho_{\mu,t}(\mathfrak{e}^\pm_{\mu,t} \pm \omega)\bigg[1 + \mathcal{O}\bigg((t_* - t)^{\frac{1}{3}} + \min\Big\{\omega^{\frac{1}{3}}, \frac{\omega^{\frac{1}{2}}}{(t_* - t)^{\frac{1}{4}}}\Big\}\bigg)\bigg] \tag{4-6b}$$

for $0 \le \omega \le \delta_*$. For $t_* < t \le 2t_*$ the densities $\rho_{r,t}$ have small local minima $\mathfrak{m}_{r,t}$ of size

$$\rho_{r,t}(\mathfrak{m}_{r,t}) = \frac{\gamma^2\sqrt{t - t_*}}{\pi}[1 + \mathcal{O}((t - t_*)^{\frac{1}{2}})], \quad \rho_{r,t}(\mathfrak{m}_{r,t}) = \rho_{\mu,t}(\mathfrak{m}_{\mu,t})[1 + \mathcal{O}((t - t_*)^{\frac{1}{2}})], \tag{4-6c}$$

and the densities are close in the sense

$$\frac{\rho_{r,t}(\mathfrak{m}_{r,t} + \omega)}{\rho_{\mu,t}(\mathfrak{m}_{\mu,t} + \omega)} = 1 + \mathcal{O}\bigg((t - t_*)^{\frac{1}{2}} + \min\Big\{(t - t_*)^{\frac{1}{4}}, \frac{(t - t_*)^2}{|\omega|}\Big\} + \min\Big\{\frac{\omega^2}{(t - t_*)^{\frac{5}{2}}}, |\omega|^{\frac{1}{3}}\Big\}\bigg) \tag{4-6d}$$

for $\omega \in [-\delta_*, \delta_*]$. Here $\delta_*, \delta_{**} \sim 1$ are small constants depending on the model parameters in Assumptions (A)–(C).

Lemma 4.3. *The density $\rho_{\alpha,t}$ from Definition 4.1 is well-defined and is a $\frac{1}{3}$-Hölder continuous density. More precisely, in the precusp regime, i.e., for $t \le t_*$, we have*

$$|\rho'_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm x)| \lesssim \frac{1}{\rho_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm x)(\rho_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm x) + \Delta^{\frac{1}{3}}_{\alpha,t})} \tag{4-7a}$$

for $0 \le x \le \delta_*$. Moreover, the Stieltjes transform $m_{\alpha,t}$ satisfies the bounds

$$\begin{aligned} |m_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm x)| &\lesssim 1, \\ |m_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm (x + y)) - m_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm x)| &\lesssim \frac{|y||\log|y||}{\rho_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm x)(\rho_{\alpha,t}(\mathfrak{e}^\pm_{\alpha,t} \pm x) + \Delta^{\frac{1}{3}}_{\alpha,t})} \end{aligned} \tag{4-7b}$$

for $|x| \le \frac{1}{2}\delta_*$, $|y| \ll x$. In the small minimum case, i.e., for $t \ge t_*$, we similarly have

$$|\rho'_{\alpha,t}(\mathfrak{m}_{\alpha,t} + x)| \lesssim \frac{1}{\rho^2_{\alpha,t}(\mathfrak{m}_{\alpha,t} + x)} \tag{4-8a}$$

for $|x| \le \delta_*$ and

$$|m_{\alpha,t}(\mathfrak{m}_{\alpha,t} + x)| \lesssim 1, \quad |m_{\alpha,t}(\mathfrak{m}_{\alpha,t} + (x + y)) - m_{\alpha,t}(\mathfrak{m}_{\alpha,t} + x)| \lesssim \frac{|y||\log|y||}{\rho^2_{\alpha,t}(\mathfrak{m}_{\alpha,t} + x)} \tag{4-8b}$$

for $|x| \le \delta_*$ and $|y| \ll |x|$.

Proof of Lemma 4.2. We first consider the two densities $r=\lambda,\mu$ only. The first claims in (4-6a) and (4-6c) follow directly from [Erdős et al. 2018, Lemma 5.1], while the second claims follow immediately from the first ones. For the proof of (4-6b) and (4-6d) we first note that by elementary calculus

$$\Psi_{\mathrm{edge}}((1+\epsilon)\lambda)=\Psi_{\mathrm{edge}}(\lambda)[1+\mathcal{O}(\epsilon)],\quad \Psi_{\min}((1+\epsilon)\lambda)=\Psi_{\min}(\lambda)[1+\mathcal{O}(\epsilon)]$$

so that

$$\Delta_{\lambda,t}^{\frac{1}{3}}\Psi_{\mathrm{edge}}\left(\frac{\omega}{\Delta_{\lambda,t}}\right)=\Delta_{\mu,t}^{\frac{1}{3}}\Psi_{\mathrm{edge}}\left(\frac{\omega}{\Delta_{\mu,t}}\right)[1+\mathcal{O}((t_*-t)^{\frac{1}{3}})]$$

and the claimed approximations follow together with (2-3b) and (2-3d). Here the exact cusp case $t=t_*$ is also covered by interpreting $0^{\frac{1}{3}}\Psi_{\mathrm{edge}}(\omega/0)=\omega^{\frac{1}{3}}/2^{\frac{4}{3}}$.

In order to prove the corresponding statements for the interpolating densities $\rho_{\alpha,t}$, we first have to establish a quantitative understanding of the counting function $n_{r,t}$ and its inverse. We claim that for $r=\alpha,\lambda,\mu$ they satisfy, for $0\le E\le\delta_*,\ 0\le s\le\delta_{**}$,

$$n_{r,t}(E)\sim\min\left\{\frac{E^{\frac{3}{2}}}{\Delta_{r,t}^{\frac{1}{6}}},E^{\frac{4}{3}}\right\},\quad \varphi_{r,t}(s)\sim\max\{s^{\frac{3}{4}},s^{\frac{2}{3}}\Delta_{r,t}^{\frac{1}{9}}\},\quad \frac{\varphi_{r,t}(s)}{\varphi_{\lambda,t}(s)}\sim\min\left\{\varphi_{\lambda,t}^{\frac{1}{3}}(s),\frac{\varphi_{\lambda,t}^{\frac{1}{2}}(s)}{\Delta_{\lambda,t}^{\frac{1}{6}}}\right\}\tag{4-9a}$$

for $t\le t_*$ and

$$\begin{gathered}n_{r,t}(E)\sim\max\{E^{\frac{4}{3}},E\rho_{r,t}(\mathfrak{m}_{r,t})\},\quad \varphi_{r,t}(s)\sim\min\left\{s^{\frac{3}{4}},\frac{s}{\rho_{r,t}(\mathfrak{m}_{r,t})}\right\},\\ \frac{\varphi_{r,t}(s)}{\varphi_{\lambda,t}(s)}\sim\min\left\{\varphi_{\lambda,t}^{\frac{1}{3}}(s),\frac{\varphi_{\lambda,t}(s)}{\rho_{r,t}^{2}(\mathfrak{m}_{r,t})},\frac{\varphi_{\lambda,t}^{2}(s)}{\rho_{r,t}^{\frac{11}{2}}(\mathfrak{m}_{r,t})}\right\}\end{gathered}\tag{4-9b}$$

for $t\ge t_*$.

Proof of (4-9). We begin with the proof of (4-9a) for $r=\lambda,\mu$. Recall that the shape function Ψ_{edge} satisfies the scaling $\Delta^{\frac{1}{3}}\Psi_{\mathrm{edge}}(\omega/\Delta)\sim\min\{\omega^{\frac{1}{3}},\omega^{\frac{1}{2}}/\Delta^{\frac{1}{6}}\}$. We first find by elementary integration that

$$\int_0^q\min\left\{\omega^{\frac{1}{3}},\frac{\omega^{\frac{1}{2}}}{\Delta^{\frac{1}{6}}}\right\}\mathrm{d}\omega=\frac{9q^{\frac{4}{3}}\min\{q,\Delta\}^{\frac{1}{6}}-\min\{q,\Delta\}^{\frac{3}{2}}}{12\Delta^{\frac{1}{6}}}\sim\min\left\{\frac{q^{\frac{3}{2}}}{\Delta^{\frac{1}{6}}},q^{\frac{4}{3}}\right\},$$

from which we conclude the first relation in (4-9a), and by inversion also the second relation. Together with the estimate for the error integral for $\rho_{\lambda,t}(\mathfrak{e}_{\lambda,t}^{+}+\omega)-\rho_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\omega)\lesssim\min\{\omega^{\frac{2}{3}},\omega/\Delta_{\lambda,t}^{\frac{1}{3}}\}$,

$$\int_0^q\min\left\{\omega^{\frac{2}{3}},\frac{\omega}{\Delta^{\frac{1}{3}}}\right\}\mathrm{d}\omega=\frac{6q^{\frac{5}{3}}\min\{q,\Delta\}^{\frac{1}{3}}-\min\{q,\Delta\}^{2}}{10\Delta^{\frac{1}{3}}}\sim\min\left\{\frac{q^{2}}{\Delta^{\frac{1}{3}}},q^{\frac{5}{3}}\right\},$$

we can thus conclude also the third relation in (4-9a).

We now turn to the case $t>t_*$ where both densities $\rho_{\lambda,t},\rho_{\mu,t}$ exhibit a small local minimum. We first record the elementary integral

$$\int_0^q\left(\rho+\min\left\{\omega^{\frac{1}{3}},\frac{\omega^2}{\rho^5}\right\}\right)\mathrm{d}\omega=\frac{q^{\frac{4}{3}}\min\{\rho^3,q\}^{\frac{5}{3}}+12q\rho^6-5\min\{q,\rho^3\}^3}{12\rho^5}\sim\max\{q^{\frac{4}{3}},q\rho\}$$

for $q,\rho\ge 0$ and easily conclude the first two relations in (4-9b). For the error integral we obtain

$$\int_0^q \min\Big\{\omega^{\frac13},\frac{\omega^2}{\rho^5}\Big\}\Big(\min\Big\{\rho^{\frac12},\frac{\rho^4}{\omega}\Big\}+\min\Big\{\omega^{\frac13},\frac{\omega^2}{\rho^5}\Big\}\Big)\,\mathrm{d}\omega\sim\min\Big\{q^{\frac53},\frac{q^2}{\rho},\frac{q^3}{\rho^{\frac92}}\Big\},$$

from which the third relation in (4-9b) follows. Finally, the claims (4-9a) and (4-9b) for $r=\alpha$ follow immediately from Definition 4.1 and the corresponding statements for $r=\lambda,\mu$. This completes the proof of (4-9). □

We now turn to the density $\rho_{\alpha,t}$ for which the claims (4-6a) and (4-6c) follow immediately from Definition 4.1 and the corresponding statements for $\rho_{\lambda,t}$ and $\rho_{\mu,t}$. For $t\le t_*$ we now continue by differentiating $E=\varphi_{r,t}(n_{r,t}(E))$ to obtain

$$\begin{aligned}\rho_{\alpha,t}(\mathfrak{e}_{\alpha,t}^{+}+\varphi_{\alpha,t}(s))&=\frac{1}{\varphi'_{\alpha,t}(s)}=\frac{1}{\alpha\varphi'_{\lambda,t}(s)+(1-\alpha)\varphi'_{\mu,t}(s)}\\&=\Big(\frac{\alpha}{\rho_{\lambda,t}(\mathfrak{e}_{\lambda,t}^{+}+\varphi_{\lambda,t}(s))}+\frac{1-\alpha}{\rho_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\varphi_{\mu,t}(s))}\Big)^{-1}\\&=\rho_{\lambda,t}(\mathfrak{e}_{\lambda,t}^{+}+\varphi_{\lambda,t}(s))\Big(\alpha+(1-\alpha)\frac{\rho_{\lambda,t}(\mathfrak{e}_{\lambda,t}^{+}+\varphi_{\lambda,t}(s))}{\rho_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\varphi_{\mu,t}(s))}\Big)^{-1},\end{aligned}$$

from which we can easily conclude (4-6b) for $r=\alpha$ together with (4-6b) for $r=\lambda$ and (4-9a). The proof of (4-6d) for $r=\alpha$ follows by the same argument and replacing $\mathfrak{e}_{r,t}^{+}$ by $\mathfrak{m}_{r,t}$. This finishes the proof of Lemma 4.2 □

Proof of Lemma 4.3. By differentiating we find

$$\begin{aligned}\frac{\rho'_{\alpha,t}(\mathfrak{e}_{\alpha,t}^{+}+\varphi_{\alpha,t}(s))}{\rho_{\alpha,t}(\mathfrak{e}_{\alpha,t}^{+}+\varphi_{\alpha,t}(s))}&=-\frac{\alpha\varphi''_{\lambda,t}(s)+(1-\alpha)\varphi''_{\mu,t}(s)}{(\alpha\varphi'_{\lambda,t}(s)+(1-\alpha)\varphi'_{\mu,t}(s))^2}\\&=\Big(\alpha\frac{\rho'_{\lambda,t}(\mathfrak{e}_{\lambda,t}^{+}+\varphi_{\lambda,t}(s))}{\rho^3_{\lambda,t}(\mathfrak{e}_{\lambda,t}^{+}+\varphi_{\lambda,t}(s))}+(1-\alpha)\frac{\rho'_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\varphi_{\mu,t}(s))}{\rho^3_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\varphi_{\mu,t}(s))}\Big)\\&\qquad\times\Big(\frac{\alpha}{\rho_{\lambda,t}(\mathfrak{e}_{\lambda,t}^{+}+\varphi_{\lambda,t}(s))}+\frac{1-\alpha}{\rho_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\varphi_{\mu,t}(s))}\Big)^{-2},\end{aligned}$$

from which we conclude the claimed bound (4-7a) together with the fact that the densities ρ_λ and ρ_μ fulfil the same bound according to [Alt et al. 2018a, Remark 10.7], and the estimates from Lemma 4.2. Similarly, the bound in (4-8a) follows by the same argument by replacing $\mathfrak{e}_{\alpha,t}^{\pm}$ by $\mathfrak{m}_{\alpha,t}$. The bound $|\rho'|\le\rho^{-2}$ on the derivative implies $\frac13$-Hölder continuity.

We now turn to the claimed bound on the Stieltjes transform and compute

$$m_{\alpha,t}(\mathfrak{e}_{\alpha,t}^{+}+x)=\int_0^{\delta_*}\frac{\rho_{\alpha,t}(\mathfrak{e}_{\alpha,t}^{+}+\omega)}{\omega-x}\,\mathrm{d}\omega+\int_{-\delta_*}^{0}\frac{\rho_{\alpha,t}(\mathfrak{e}_{\alpha,t}^{-}+\omega)}{\omega-\Delta_{\alpha,t}-x}\,\mathrm{d}\omega,$$

out of which for $x>0$ the first term can be bounded by

$$\int_0^{\delta_*} \frac{\rho_{\alpha,t}(\mathfrak{e}^+_{\alpha,t}+\omega)}{\omega-x}\,\mathrm{d}\omega \lesssim \int_0^{\delta_*} \frac{|\omega-x|^{\frac{1}{3}}}{\omega-x}\,\mathrm{d}\omega + \int_{2x}^{\delta_*} \frac{\rho_{\alpha,t}(\mathfrak{e}^+_{\alpha,t}+x)}{\omega-x}\,\mathrm{d}\omega \lesssim |x|^{\frac{1}{3}}|\log x| + |\delta_*-x|^{\frac{1}{3}},$$

while the second term can be bounded by

$$\left|\int_{-\delta_*}^{0} \frac{\rho_{\alpha,t}(\mathfrak{e}^-_{\alpha,t}+\omega)}{\omega-\Delta_{\alpha,t}-x}\,\mathrm{d}\omega\right| \lesssim |\delta_*-\Delta_{\alpha,t}-x|^{\frac{1}{3}} + |\Delta_{\alpha,t}+x|^{\frac{1}{3}}|\log(\Delta_{\alpha,t}+x)|,$$

both using the $\frac{1}{3}$-Hölder continuity of $\rho_{\alpha,t}$. The corresponding bounds for $x<0$ are similar, completing the proof of the first bound in (4-7b).

The proof of the first bound in (4-8b) is very similar and follows from

$$|m_{\alpha,t}(\mathfrak{m}_{\alpha,t}+x)| \lesssim \left|\int_{-\delta_*}^{\delta_*} \frac{|\omega-x|^{\frac{1}{3}}}{\omega-x}\,\mathrm{d}\omega\right| + \left|\int_{[-\delta_*,\delta_*]\setminus[x-\frac{1}{2}\delta_*,x+\frac{1}{2}\delta_*]} \frac{\rho_{\alpha,t}(\mathfrak{m}_{\alpha,t}+x)}{\omega-x}\,\mathrm{d}\omega\right| \lesssim 1.$$

We now turn to the second bound in (4-7b), which is only nontrivial in the case $x>0$. To simplify the following integrals we temporarily use the short-hand notations $m=m_{\alpha,t}$, $\mathfrak{e}^+=\mathfrak{e}^+_{\alpha,t}$, $\rho=\rho_{\alpha,t}$, $\Delta=\Delta_{\alpha,t}$ and compute

$$m(\mathfrak{e}^++x+y)-m(\mathfrak{e}^++x) = \int_{-\Delta-\delta_*}^{\delta_*} \frac{\rho(\mathfrak{e}^++\omega)}{\omega-x-y}\,\mathrm{d}\omega - \int_{-\Delta-\delta_*}^{\delta_*} \frac{\rho(\mathfrak{e}^++\omega)}{\omega-x}\,\mathrm{d}\omega,$$

where we now focus on the integration regime $\omega\ge 0$ as this is the regime containing the two critical singularities. We first observe that

$$\begin{aligned}\int_{-y}^{\delta_*-y} \frac{\rho(\mathfrak{e}^++\omega+y)}{\omega-x}\,\mathrm{d}\omega &- \int_0^{\delta_*} \frac{\rho(\mathfrak{e}^++\omega)}{\omega-x}\,\mathrm{d}\omega \\ &= \int_0^{\delta_*} \frac{\rho(\mathfrak{e}^++\omega+y)-\rho(\mathfrak{e}^++\omega)}{\omega-x}\,\mathrm{d}\omega + \int_{-y}^{0} \frac{\rho(\mathfrak{e}^++\omega+y)}{\omega-x}\,\mathrm{d}\omega + \mathcal{O}(y),\end{aligned}$$

where the second integral is easily bounded by

$$\int_{-y}^{0} \frac{\rho(\mathfrak{e}+\omega+y)}{\omega-x}\,\mathrm{d}\omega \lesssim \frac{1}{x}\min\{y^{\frac{4}{3}}, y^{\frac{3}{2}}\Delta^{-\frac{1}{6}}\} \lesssim \frac{y}{\rho(\mathfrak{e}^++x)(\rho(\mathfrak{e}^++x)+\Delta^{\frac{1}{3}})}.$$

We split the remaining integral into three regimes $\left[0,\frac{1}{2}x\right]$, $\left[\frac{1}{2}x,\frac{3}{2}\right]$ and $\left[\frac{3}{2}x,\delta_*\right]$. In the first one we use (4-7a) as well as the scaling relation $\rho(\mathfrak{e}^++\omega)\sim\min\{\omega^{\frac{1}{3}},\omega^{\frac{1}{2}}\Delta^{-\frac{1}{6}}\}$ to obtain

$$\begin{aligned}\int_0^{\frac{1}{2}x} \frac{\rho(\mathfrak{e}^++\omega+y)-\rho(\mathfrak{e}^++\omega)}{\omega-x}\,\mathrm{d}\omega &\lesssim \frac{y}{x}\int_0^{\frac{1}{2}x} \frac{1}{\rho(\mathfrak{e}^++\omega)(\rho(\mathfrak{e}^++\omega)+\Delta^{\frac{1}{3}})}\,\mathrm{d}\omega \\ &\lesssim \frac{y}{x}\min\left\{\frac{x^{\frac{1}{2}}}{\Delta^{\frac{1}{6}}}, x^{\frac{1}{3}}\right\} \sim \frac{y}{\max\{x^{\frac{2}{3}}, x^{\frac{1}{2}}\Delta^{\frac{1}{6}}\}} \\ &\lesssim \frac{y}{\rho(\mathfrak{e}^++x)(\rho(\mathfrak{e}^++x)+\Delta^{\frac{1}{3}})}.\end{aligned}$$

The integral in the regime $\left[\frac{3}{2}x,\delta_*\right]$ is completely analogous and contributes the same bound. Finally, we are left with the regime $\left[\frac{1}{2}x,\frac{3}{2}x\right]$ which we again subdivide into $[x-y,x+y]$ and $\left[\frac{1}{2}x,\frac{3}{2}x\right]\setminus[x-y,x+y]$. In the first of those we have

$$\int_{x-y}^{x+y}\frac{\rho(\mathfrak{e}^{+}+\omega+y)-\rho(\mathfrak{e}^{+}+\omega)}{\omega-x}\,\mathrm{d}\omega=\int_{x-y}^{x+y}\frac{\rho(\mathfrak{e}^{+}+\omega+y)-\rho(\mathfrak{e}^{+}+x+y)-\rho(\mathfrak{e}^{+}+\omega)+\rho(\mathfrak{e}^{+}+x)}{\omega-x}\,\mathrm{d}\omega$$
$$\lesssim\frac{y}{\rho(\mathfrak{e}^{+}+x)(\rho(\mathfrak{e}^{+}+x)+\Delta^{\frac{1}{3}})},$$

while in the second one we obtain

$$\int_{[\frac{1}{2}x,\frac{3}{2}x]\setminus[x-y,x+y]}\frac{\rho(\mathfrak{e}^{+}+\omega+y)-\rho(\mathfrak{e}^{+}+x+y)-\rho(\mathfrak{e}^{+}+\omega)+\rho(\mathfrak{e}^{+}+x)}{\omega-x}\,\mathrm{d}\omega$$
$$\lesssim\frac{y}{\rho(\mathfrak{e}^{+}+x)(\rho(\mathfrak{e}^{+}+x)+\Delta^{\frac{1}{3}})}\int_{[\frac{1}{2}x,\frac{3}{2}x]\setminus[x-y,x+y]}|\omega-x|^{-1}\,\mathrm{d}\omega$$
$$\lesssim\frac{y|\log y|}{\rho(\mathfrak{e}^{+}+x)(\rho(\mathfrak{e}^{+}+x)+\Delta^{\frac{1}{3}})}.$$

Collecting the various estimates completes the proof of (4-7b).

The second bound in (4-8b) follows by a similar argument and we focus on the most critical term

$$\int_{-\frac{1}{2}\delta_*}^{\frac{1}{2}\delta_*}\frac{\rho(\mathfrak{m}+\omega+y)-\rho(\mathfrak{m}+\omega)}{\omega-x}\,\mathrm{d}\omega=\left(\int_{-\frac{1}{2}\delta_*}^{x-y}+\int_{x-y}^{x+y}+\int_{x+y}^{\frac{1}{2}\delta_*}\right)\frac{\rho(\mathfrak{m}+\omega+y)-\rho(\mathfrak{m}+\omega)}{\omega-x}\,\mathrm{d}\omega.$$

Here we can bound the middle integral by

$$\left|\int_{x-y}^{x+y}\frac{\rho(\mathfrak{m}+\omega+y)-\rho(\mathfrak{m}+\omega)}{\omega-x}\,\mathrm{d}\omega\right|=\left|\int_{x-y}^{x+y}\frac{\rho(\mathfrak{m}+\omega+y)-\rho(\mathfrak{m}+x+y)-\rho(\mathfrak{m}+\omega)+\rho(\mathfrak{m}+x)}{\omega-x}\,\mathrm{d}\omega\right|$$
$$\lesssim\frac{|y|}{\rho^{2}(\mathfrak{m}+x)},$$

while for the first integral we have

$$\left|\int_{-\frac{1}{2}\delta_*}^{x-y}\frac{\rho(\mathfrak{m}+\omega+y)-\rho(\mathfrak{m}+x+y)-\rho(\mathfrak{m}+\omega)+\rho(\mathfrak{m}+x)}{\omega-x}\,\mathrm{d}\omega\right|\lesssim\frac{|y|}{\rho^{2}(\mathfrak{m}+x)}\int_{-\frac{1}{2}\delta_*}^{x-y}\frac{1}{|\omega-x|}\,\mathrm{d}\omega$$
$$\lesssim\frac{|y|\,|\log|y||}{\rho^{2}(\mathfrak{m}+x)}.$$

The third integral is completely analogous, completing the proof of (4-8b). □

4A. ***Quantiles.*** Finally we consider the locations of quantiles of $\rho_{r,t}$ for $r=\alpha,\lambda,\mu$ and their fluctuation scales. For $0\le t\le t_*$ we define the shifted quantiles $\hat{\gamma}_{r,i}(t)$, and for $t_*\le t\le 2t_*$ the shifted quantiles[4]

[4]We use a separate variable name $\check{\gamma}$ because in Section 8 the name $\hat{\gamma}$ is used for the quantiles with respect to the base point $\tilde{\mathfrak{m}}$ instead of $\mathfrak{m}$.

$\check{\gamma}_{r,i}(t)$ in such a way that

$$\int_0^{\hat{\gamma}_{r,i}(t)} \rho_{r,t}(\mathfrak{e}_{r,t}^+ + \omega)\,\mathrm{d}\omega = \frac{i}{N}, \quad \int_0^{\check{\gamma}_{r,i}(t)} \rho_{r,t}(\mathfrak{m}_{r,t} + \omega)\,\mathrm{d}\omega = \frac{i}{N}, \quad |i| \ll N. \tag{4-10}$$

Notice that for $i = 0$ we always have $\hat{\gamma}_{r,0}(t) = \check{\gamma}_{r,0}(t) = 0$. We will also need to define the semiquantiles, distinguished from the quantiles by a star:

$$\int_0^{\hat{\gamma}^*_{r,i}(t)} \rho_{r,t}(\mathfrak{e}_{r,t}^+ + \omega)\,\mathrm{d}\omega = \frac{i-\frac{1}{2}}{N}, \quad \int_0^{\check{\gamma}^*_{r,i}(t)} \rho_{r,t}(\mathfrak{m}_{r,t} + \omega)\,\mathrm{d}\omega = \frac{i-\frac{1}{2}}{N}, \quad 1 \le i \ll N, \tag{4-11}$$

$$\int_0^{\hat{\gamma}^*_{r,i}(t)} \rho_{r,t}(\mathfrak{e}_{r,t}^+ + \omega)\,\mathrm{d}\omega = \frac{i+\frac{1}{2}}{N}, \quad \int_0^{\check{\gamma}^*_{r,i}(t)} \rho_{r,t}(\mathfrak{m}_{r,t} + \omega)\,\mathrm{d}\omega = \frac{i+\frac{1}{2}}{N}, \quad -N \ll i \le -1. \tag{4-12}$$

Note that the definition is slightly different for positive and negative i's; in particular $\hat{\gamma}_i^* \in [\hat{\gamma}_{i-1}, \hat{\gamma}_i]$ for $i \ge 1$ and $\hat{\gamma}_i^* \in [\hat{\gamma}_i, \hat{\gamma}_{i+1}]$ for $i < 0$. The semiquantiles are not defined for $i = 0$.

Lemma 4.4. *For $1 \le |i| \ll N$, $r = \alpha, \lambda, \mu$ and $0 \le t \le t_*$ we have*

$$\begin{aligned} \hat{\gamma}_{r,i}(t) &\sim \operatorname{sgn}(i) \max\left\{ \left(\frac{|i|}{N}\right)^{\frac{3}{4}}, \left(\frac{|i|}{N}\right)^{\frac{2}{3}} (t_* - t)^{\frac{1}{6}} \right\} - \begin{cases} 0, & i > 0, \\ \Delta_{r,t}, & i < 0, \end{cases} \\ \hat{\gamma}_{r,i}(t) &= \hat{\gamma}_{\mu,i}(t)\left[1 + \mathcal{O}\left((t_* - t)^{\frac{1}{3}} + \min\left\{ \frac{\hat{\gamma}_{\mu,i}(t)^{\frac{1}{2}}}{(t_* - t)^{\frac{1}{4}}}, \hat{\gamma}_{\mu,i}(t)^{\frac{1}{3}} \right\}\right)\right], \end{aligned} \tag{4-13a}$$

while for $t_* \le t \le 2t_*$ we have

$$\begin{aligned} \check{\gamma}_{r,i}(t) &\sim \operatorname{sgn}(i) \min\left\{ \left(\frac{|i|}{N}\right)^{\frac{3}{4}}, \frac{|i|}{N} (t_* - t)^{-\frac{1}{2}} \right\}, \\ \check{\gamma}_{r,i}(t) &= \check{\gamma}_{\mu,i}(t)\left[1 + \mathcal{O}\left((t_* - t)^{\frac{1}{2}} + \min\left\{ \frac{\check{\gamma}_{\mu,i}(t)^2}{(t_* - t)^{\frac{11}{4}}}, \frac{\check{\gamma}_{\mu,i}(t)}{t_* - t}, \check{\gamma}_{\mu,i}(t)^{\frac{1}{3}} \right\}\right)\right]. \end{aligned} \tag{4-13b}$$

Moreover, the quantiles of $\rho_{\alpha,t}$ are the convex combination

$$\hat{\gamma}_{\alpha,i}(t) = \alpha \hat{\gamma}_{\lambda,i}(t) + (1-\alpha)\hat{\gamma}_{\mu,i}(t), \quad \check{\gamma}_{\alpha,i}(t) = \alpha \check{\gamma}_{\lambda,i}(t) + (1-\alpha)\check{\gamma}_{\mu,i}(t). \tag{4-13c}$$

Proof. The proof follows directly from the estimates in (4-9a) and (4-9b). The relation (4-13c) follows directly from (4-2) in the definition of the α-interpolating density. □

4B. ***Movement of edges, quantiles and minima.*** For the analysis of the Dyson Brownian motion it is necessary to have a precise understanding of the movement of the reference points $\mathfrak{e}_{r,t}^{\pm}$ and $\mathfrak{m}_{r,t}$, $r = \lambda, \mu$. For technical reasons it is slightly easier to work with an auxiliary quantity $\tilde{\mathfrak{m}}_{r,t}$ which is very close to $\mathfrak{m}_{r,t}$. According to [Erdős et al. 2018, Lemma 5.1] the minimum $\mathfrak{m}_{r,t}$ can approximately be found by solving the implicit equation

$$\tilde{\mathfrak{m}}_{r,t} = \mathfrak{c}_r - (t - t_*)\Re m_{r,t}(\tilde{\mathfrak{m}}_{r,t}), \quad \tilde{\mathfrak{m}}_{r,t} \in \mathbb{R}, \quad r = \lambda, \mu. \tag{4-14a}$$

The explicit relation (4-14a) is the main reason why it is more convenient to study the movement of $\tilde{\mathfrak{m}}_t$ rather than the one of $\mathfrak{m}_t$. We claim that $\tilde{\mathfrak{m}}_{r,t}$ is indeed a very good approximation for $\mathfrak{m}_{r,t}$ in the sense that

$$|\mathfrak{m}_{r,t}-\tilde{\mathfrak{m}}_{r,t}|\lesssim(t-t_*)^{\frac{3}{2}+\frac{1}{4}},\quad \Im m_{r,t}(\tilde{\mathfrak{m}}_{r,t})=\gamma^2(t-t_*)^{\frac{1}{2}}+\mathcal{O}(t-t_*),\quad r=\lambda,\mu. \tag{4-14b}$$

Proof of (4-14b). The first claim in (4-14b) is a direct consequence of [Erdős et al. 2018, Lemma 5.1]. For the second claim we refer to [Erdős et al. 2018, equation (89a)], which implies

$$\Im m_{r,t}(\tilde{\mathfrak{m}}_{r,t})=(t-t_*)^{\frac{1}{2}}\gamma^2[1+\mathcal{O}((t-t_*)^{\frac{1}{3}}[\Im m_{r,t}(\tilde{\mathfrak{m}}_{r,t})]^{\frac{1}{3}})]=\gamma^2(t-t_*)^{\frac{1}{2}}+\mathcal{O}(t-t_*). \qquad \square$$

For the t-derivative of (semi-)quantiles $\gamma_{r,t}$, i.e., points such that $\int_{-\infty}^{\gamma_{r,t}}\rho_{r,t}(x)\,\mathrm{d}x$ is constant in t, as well as for the minima $\tilde{\mathfrak{m}}_{r,t}$, we have the explicit relations

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma_{r,t}=-\Re m_{r,t}(\gamma_{r,t}), \tag{4-14c}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{r,t}=-\Re m_{r,t}(\tilde{\mathfrak{m}}_{r,t})+\mathcal{O}(t-t_*),\quad t_*\le t\le 2t_*. \tag{4-14d}$$

In particular, for the spectral edges it follows from (4-14c) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{e}_{r,t}^{+}=-m_{r,t}(\mathfrak{e}_{r,t}^{+}),\quad 0\le t\le t_*. \tag{4-14e}$$

Proof of (4-14c)–(4-14e). For the proof of (4-14c) we first recall that from the defining equation (4-1) of the semicircular flow it follows that the Stieltjes transform $m=m_t(\zeta)$ of ρ_t satisfies the Burgers equation

$$\dot{m}=mm'=\tfrac{1}{2}(m^2)', \tag{4-15}$$

where prime denotes the $\frac{\mathrm{d}}{\mathrm{d}\zeta}$ derivative and dot denotes the $\frac{\mathrm{d}}{\mathrm{d}t}$ derivative. Thus

$$\begin{aligned}\dot{\gamma}_{r,t}&=-\frac{1}{\rho_{r,t}(\gamma_{r,t})}\Im\int_{-\infty}^{\gamma_{r,t}}\dot{m}_{r,t}(E)\,\mathrm{d}E=-\frac{1}{2\rho_{r,t}(\gamma_{r,t})}\Im\int_{-\infty}^{\gamma_{r,t}}(m_{r,t}^2)'(E)\,\mathrm{d}E\\&=-\frac{\Im m_{r,t}^2(\gamma_{r,t})}{2\Im m_{r,t}(\gamma_{r,t})}=-\Re m_{r,t}(\gamma_{r,t})\end{aligned}$$

follows directly from differentiating $\int_{-\infty}^{\gamma_{r,t}}\rho_{r,t}(x)\,\mathrm{d}x\equiv\text{const}$.

For (4-14d) we begin by computing the integral

$$m_{r,t_*}'(\mathfrak{c}_r+\mathrm{i}\eta)=\int_{\mathbb{R}}\frac{\rho_{t_*}(\mathfrak{c}_r+x)}{(x-\mathrm{i}\eta)^2}\,\mathrm{d}x=\int_{\mathbb{R}}\frac{\sqrt{3}\gamma^{\frac{4}{3}}|x|^{\frac{1}{3}}+\mathcal{O}(|x|^{\frac{2}{3}})}{2\pi(x-\mathrm{i}\eta)^2}\,\mathrm{d}x=\frac{\gamma^{\frac{4}{3}}}{3\eta^{\frac{2}{3}}}+\mathcal{O}(\eta^{-\frac{1}{3}}), \tag{4-16}$$

so that by the definition $m_{r,t}(z)=m_{r,t_*}(z+(t-t_*)m_{r,t}(z))$ of the free semicircular flow,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}m_{r,t}(\tilde{\mathfrak{m}}_{r,t})&=m_{r,t_*}'(\tilde{\mathfrak{m}}_{r,t}+(t-t_*)m_{r,t}(\tilde{\mathfrak{m}}_{r,t}))\left[\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{r,t}+m_{r,t}(\tilde{\mathfrak{m}}_{r,t})+(t-t_*)\frac{\mathrm{d}}{\mathrm{d}t}m_{r,t}(\tilde{\mathfrak{m}}_{z,t})\right]\\&=\left(\frac{1}{3(t-t_*)}+\mathcal{O}((t-t_*)^{-\frac{1}{2}})\right)\left[\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{r,t}+m_{r,t}(\tilde{\mathfrak{m}}_{r,t})+(t-t_*)\frac{\mathrm{d}}{\mathrm{d}t}m_{r,t}(\tilde{\mathfrak{m}}_{r,t})\right]\\&=\mathrm{i}\left(\frac{1}{3(t-t_*)}+\mathcal{O}((t-t_*)^{-\frac{1}{2}})\right)\left[\Im m_{r,t}(\tilde{\mathfrak{m}}_{r,t})+(t-t_*)\frac{\mathrm{d}}{\mathrm{d}t}\Im m_{r,t}(\tilde{\mathfrak{m}}_{r,t})\right]\\&=\left(\mathrm{i}\frac{\gamma^2}{3(t-t_*)^{\frac{1}{2}}}+\frac{\mathrm{i}}{3}\frac{\mathrm{d}}{\mathrm{d}t}\Im m_{r,t}(\tilde{\mathfrak{m}}_{r,t})\right)[1+\mathcal{O}((t-t_*)^{\frac{1}{2}})].\end{aligned}$$

Here we used (4-14a), (4-14b) together with (4-16) in the second step. The third step follows from taking the t-derivative of (4-14a). The ultimate inequality is again a consequence of (4-14b). By considering real and imaginary parts separately it thus follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Im m_{r,t}(\tilde{\mathfrak{m}}_{r,t}) = \frac{\gamma^2}{2(t-t_*)^{\frac{1}{2}}}[1+\mathcal{O}((t-t_*)^{\frac{1}{2}})], \quad \frac{\mathrm{d}}{\mathrm{d}t}\Re m_{r,t}(\tilde{\mathfrak{m}}_{r,t}) = \mathcal{O}(1),$$

and therefore (4-14d) follows by differentiating (4-14a). □

4C. ***Rigidity scales.*** In this section we compute, up to leading order, the fluctuations of the eigenvalues around their classical locations, i.e., the quantiles defined in Section 4A. Indeed, the computation of the fluctuation scale for the particles $x_i(t)$, $y_i(t)$, defined in (5-5), (5-7), will be one of the fundamental inputs to prove rigidity for the interpolated process in Section 6. The fluctuation scale $\eta_{\mathrm{f}}^{\rho}(\tau)$ of any density function $\rho(\omega)$ around τ is defined via

$$\int_{\tau-\eta_{\mathrm{f}}^{\rho}(\tau)}^{\tau+\eta_{\mathrm{f}}^{\rho}(\tau)} \rho(\omega)\,\mathrm{d}\omega = \frac{1}{N}$$

for $\tau \in \operatorname{supp}\rho$ and by the value $\eta_{\mathrm{f}}(\tau) := \eta_{\mathrm{f}}(\tau')$, where $\tau' \in \operatorname{supp}\rho$ is the edge closest to τ for $\tau \notin \operatorname{supp}\rho$. If this edge is not unique, an arbitrary choice can be made between the two possibilities. From (4-13a) we immediately obtain for $0 \le t \le t_*$ and $1 \le i \le N$ that

$$\eta_{\mathrm{f}}^{\rho_{r,t}}(\mathfrak{e}_{r,t}^{+} + \hat{\gamma}_{r,\pm i}(t)) \sim \max\left\{\frac{\Delta_{r,t}^{\frac{1}{9}}}{N^{\frac{2}{3}} i^{\frac{1}{3}}}, \frac{1}{N^{\frac{3}{4}} i^{\frac{1}{4}}}\right\} \sim \max\left\{\frac{(t_*-t)^{\frac{1}{6}}}{N^{\frac{2}{3}} i^{\frac{1}{3}}}, \frac{1}{N^{\frac{3}{4}} i^{\frac{1}{4}}}\right\}, \quad r = \alpha, \lambda, \mu, \tag{4-17a}$$

while for $t_* \le t \le 2t_*,\ 1 \le |i| \ll N$ we obtain from (4-13b) that

$$\eta_{\mathrm{f}}^{\rho_{r,t}}(\mathfrak{m}_{r,t} + \check{\gamma}_{r,i}(t)) \sim \min\left\{\frac{1}{N\rho_{r,t}(\mathfrak{m}_{r,t})}, \frac{1}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}\right\} \sim \min\left\{\frac{1}{N(t-t_*)^{\frac{1}{2}}}, \frac{1}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}\right\}, \quad r = \alpha, \lambda, \mu. \tag{4-17b}$$

In the second relations we used (4-6a) and (4-6c). For reference purposes we also list for $0 < i, j \ll N$ the bounds

$$|\hat{\gamma}_{r,i}(t) - \hat{\gamma}_{r,j}(t)| \sim \max\left\{\frac{\Delta_{r,t}^{\frac{1}{9}}|i-j|}{N^{\frac{2}{3}}(i+j)^{\frac{1}{3}}}, \frac{|i-j|}{N^{\frac{3}{4}}(i+j)^{\frac{1}{4}}}\right\} \tag{4-18}$$

in the case $t \le t_*$ and

$$|\check{\gamma}_{r,i}(t) - \check{\gamma}_{r,j}(t)| \sim \min\left\{\frac{|i-j|}{\rho_{r,t}(\mathfrak{m}_{r,t})N}, \frac{|i-j|}{N^{\frac{3}{4}}(i+j)^{\frac{1}{4}}}\right\} \tag{4-19}$$

in the case $t > t_*$. Furthermore we have

$$\rho_{r,t}(\mathfrak{e}_{r,t}^{+} + \hat{\gamma}_{r,i}(t)) \sim \min\left\{\frac{i^{\frac{1}{3}}}{N^{\frac{1}{3}}(t_*-t)^{\frac{1}{6}}}, \frac{i^{\frac{1}{4}}}{N^{\frac{1}{4}}}\right\}, \tag{4-20}$$

$$\rho_{r,t}(\mathfrak{m}_{r,t} + \check{\gamma}_{r,i}(t)) \sim \max\left\{\rho_{r,t}(\mathfrak{m}_{r,t}), \frac{i^{\frac{1}{4}}}{N^{\frac{1}{4}}}\right\}. \tag{4-21}$$

4D. ***Stieltjes transform bounds.*** It follows from (4-6b) and (4-6d) that also the real parts of the Stieltjes transforms $m_{\alpha,t}, m_{\lambda,t}, m_{\mu,t}$ are close. We claim that for $r=\lambda,\alpha$ and $\nu\in[-\delta_*,\delta_*]$ and $0\le t\le t_*$ we have

$$\begin{aligned}\bigl|\Re[(m_{r,t}(\mathfrak{e}^+_{r,t}+\nu)-m_{r,t}(\mathfrak{e}^+_{r,t}))-(m_{\mu,t}(\mathfrak{e}^+_{\mu,t}+\nu)-m_{\mu,t}(\mathfrak{e}^+_{\mu,t}))]\bigr|\\ \lesssim|\nu|^{\frac13}[|\nu|^{\frac13}+(t_*-t)^{\frac13}]\,|\log|\nu||+(t_*-t)^{\frac{11}{18}}\mathbf{1}\bigl(\nu\le-\tfrac12\Delta_{\mu,t}\bigr),\end{aligned}\tag{4-22a}$$

while for $t_*\le t\le 2t_*$ we have

$$\begin{aligned}\bigl|\Re[(m_{r,t}(\mathfrak{m}_{r,t}+\nu)-m_{r,t}(\mathfrak{m}_{r,t}))-(m_{\mu,t}(\mathfrak{m}_{\mu,t}+\nu)-m_{\mu,t}(\mathfrak{m}_{\mu,t}))]\bigr|\\ \lesssim[|\nu|^{\frac13}(t-t_*)^{\frac14}+(t_*-t)^{\frac34}+|\nu|^{\frac23}]\,|\log|\nu||.\end{aligned}\tag{4-22b}$$

Proof of (4-22). We first recall from Lemma 4.3 that also the density $\rho_{\alpha,t}$ is $\frac13$-Hölder continuous, which we will use repeatedly in the following proof. We begin with the proof of (4-22a) and compute for $r=\alpha,\lambda,\mu$

$$\Re[m_{r,t}(\mathfrak{e}^+_{r,t}+\nu)-m_{r,t}(\mathfrak{e}^+_{r,t})]=\int_0^\infty\frac{\nu\rho_{r,t}(\mathfrak{e}^+_{r,t}+\omega)}{(\omega-\nu)\omega}\,\mathrm{d}\omega+\int_0^\infty\frac{\nu\rho_{r,t}(\mathfrak{e}^-_{r,t}-\omega)}{(\omega+\Delta_{r,t}+\nu)(\omega+\Delta_{r,t})}\,\mathrm{d}\omega.\tag{4-23}$$

For $\nu>0$ the first of the two terms is the more critical one. Our goal is to obtain a bound on

$$\int_0^\infty\frac{\nu}{(\omega-\nu)\omega}[\rho_{\lambda,t}(\mathfrak{e}^+_{\lambda,t}+\omega)-\rho_{\mu,t}(\mathfrak{e}^+_{\mu,t}+\omega)]\,\mathrm{d}\omega$$

by using (4-6b). Let $0<\epsilon<\frac12\nu$ be a small parameter for which we separately consider the two critical regimes $0\le\omega\le\epsilon$ and $|\nu-\omega|\le\epsilon$. We use

$$\rho_{r,t}(\mathfrak{e}^+_{r,t}+\omega)\lesssim\omega^{\frac13}\quad\text{and}\quad\rho_{r,t}(\mathfrak{e}^+_{r,t}+\omega)=\rho_{r,t}(\mathfrak{e}^+_{r,t}+\nu)+\mathcal{O}(|\omega-\nu|^{\frac13}),\quad r=\lambda,\mu,\tag{4-24}$$

from the $\frac13$-Hölder continuity of $\rho_{r,t}$ and the fact that the integral over $1/(\omega-\nu)$ from $\nu-\epsilon$ to $\nu+\epsilon$ vanishes by symmetry to estimate, for $r=\lambda,\mu$,

$$\biggl|\int_0^\epsilon\frac{\nu}{(\omega-\nu)\omega}\rho_{r,t}(\mathfrak{e}^+_{r,t}+\omega)\,\mathrm{d}\omega\biggr|\lesssim\int_0^\epsilon|\omega|^{-\frac23}\,\mathrm{d}\omega\lesssim\epsilon^{\frac13},$$

$$\biggl|\int_{\nu-\epsilon}^{\nu+\epsilon}\biggl[\frac{\rho_{r,t}(\mathfrak{e}^+_{r,t}+\omega)}{\omega-\nu}-\frac{\rho_{r,t}(\mathfrak{e}^+_{r,t}+\omega)}{\omega}\biggr]\mathrm{d}\omega\biggr|\lesssim\int_{\nu-\epsilon}^{\nu+\epsilon}|\omega-\nu|^{-\frac23}\,\mathrm{d}\omega+\epsilon\nu^{-\frac23}\lesssim\epsilon^{\frac13}+\epsilon\nu^{-\frac23}.$$

Next, we consider the remaining integration regimes where we use (4-6b) and (4-24) to estimate

$$\begin{aligned}&\biggl|\int_\epsilon^{\nu-\epsilon}\frac{\nu}{(\omega-\nu)\omega}[\rho_{r,t}(\mathfrak{e}^+_{r,t}+\omega)-\rho_{\mu,t}(\mathfrak{e}^+_{\mu,t}+\omega)]\,\mathrm{d}\omega\biggr|\\ &\qquad\lesssim\int_\epsilon^{\frac12\nu}\frac{\omega^{\frac13}(t_*-t)^{\frac13}+\omega^{\frac23}}{\omega}\,\mathrm{d}\omega+\int_{\frac12\nu}^{\nu-\epsilon}\biggl(\frac{\nu^{\frac13}(t_*-t)^{\frac13}}{\omega-\nu}+\frac{\nu^{\frac23}}{\omega-\nu}\biggr)\mathrm{d}\omega\\ &\qquad\lesssim\nu^{\frac13}((t_*-t)^{\frac13}+\nu^{\frac13})|\log\epsilon|\end{aligned}$$

and similarly

$$\left|\int_{\nu+\epsilon}^{\infty} \frac{\nu}{(\omega-\nu)\omega}[\rho_{r,t}(\mathfrak{e}_{r,t}^{+}+\omega)-\rho_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\omega)]\,\mathrm{d}\omega\right| \lesssim \nu^{\frac{1}{3}}((t_*-t)^{\frac{1}{3}}+\nu^{\frac{1}{3}})|\log\epsilon|.$$

We now consider the difference of the first terms in (4-23) for $r=\lambda,\mu$ and for $\nu<0$, where the bound is simpler because the integration regime close to ν does not have to be singled out. Using (4-6b) we find

$$\left|\int_{0}^{\infty} \frac{\nu}{(\omega-\nu)\omega}[\rho_{r,t}(\mathfrak{e}_{r,t}^{+}+\omega)-\rho_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\omega)]\,\mathrm{d}\omega\right| \lesssim |\nu|^{\frac{2}{3}}+(t_*-t)^{\frac{1}{3}}|\nu|^{\frac{1}{3}}.$$

Finally, it remains to consider the difference of the second terms in (4-23). We first treat the regime where $\nu\geq-\frac{3}{4}\Delta_{r,t}$ and split the difference into the sum of two terms

$$\begin{aligned}\left|\int_0^\infty\left(\frac{\nu\rho_{r,t}(\mathfrak{e}_{r,t}^{-}-\omega)}{(\omega+\Delta_{r,t}+\nu)(\omega+\Delta_{r,t})}-\frac{\nu\rho_{r,t}(\mathfrak{e}_{r,t}^{-}-\omega)}{(\omega+\Delta_{\mu,t}+\nu)(\omega+\Delta_{\mu,t})}\right)\mathrm{d}\omega\right| \\ \leq |\nu||\Delta_{r,t}-\Delta_{\mu,t}|\int_0^\infty \frac{\rho_{r,t}(\mathfrak{e}_{r,t}^{-}-\omega)[2\Delta_{r,t}+2\omega+|\nu|]}{(\omega+\Delta_{r,t}+\nu)^2(\omega+\Delta_{r,t})^2}\,\mathrm{d}\omega \\ \lesssim \frac{|\Delta_{r,t}-\Delta_{\mu,t}|}{\Delta_{r,t}^{\frac{2}{3}}}-\frac{|\Delta_{r,t}-\Delta_{\mu,t}|}{(\Delta_{r,t}+|\nu|)^{\frac{2}{3}}}\lesssim (t_*-t)^{\frac{1}{3}}|\nu|^{\frac{1}{3}}\end{aligned}$$

and

$$\left|\int_0^\infty\left(\frac{\nu\rho_{r,t}(\mathfrak{e}_{r,t}^{-}-\omega)}{(\omega+\Delta_{\mu,t}+\nu)(\omega+\Delta_{\mu,t})}-\frac{\nu\rho_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}-\omega)}{(\omega+\Delta_{\mu,t}+\nu)(\omega+\Delta_{\mu,t})}\right)\mathrm{d}\omega\right|\lesssim|\nu|^{\frac{2}{3}}+(t_*-t)^{\frac{1}{3}}|\nu|^{\frac{1}{3}}.$$

Here we used $\rho_{r,t}(\mathfrak{e}_{r,t}^{-}-\omega)\lesssim\omega^{\frac{1}{3}}$, as well as (4-6a) for the first and (4-6a), (4-6b) for the second computation. By collecting the various error terms and choosing $\epsilon=\nu^2$ we conclude (4-22a).

We define $\kappa:=-\nu-\Delta_{r,t}$. Then we are left with the regime $\nu<-\frac{3}{4}\Delta_{r,t}$ or equivalently $\kappa>-\frac{1}{4}\Delta_{r,t}$ and use

$$m_{r,t}(\mathfrak{e}_{r,t}^{+}+\nu)-m_{r,t}(\mathfrak{e}_{r,t}^{+})=(m_{r,t}(\mathfrak{e}_{r,t}^{-}-\kappa)-m_{r,t}(\mathfrak{e}_{r,t}^{-}))+(m_{r,t}(\mathfrak{e}_{r,t}^{-})-m_{r,t}(\mathfrak{e}_{r,t}^{+}))\,,$$

as well as

$$\begin{aligned}m_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}+\nu)-m_{\mu,t}(\mathfrak{e}_{\mu,t}^{+})&=\big(m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}-\kappa+\Delta_{\mu,t}-\Delta_{r,t})-m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}-\kappa)\big)\\&\quad+\big(m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}-\kappa)-m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-})\big)+\big(m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-})-m_{\mu,t}(\mathfrak{e}_{\mu,t}^{+})\big)\end{aligned}\tag{4-25}$$

in the left-hand side of (4-22a). Thus we have to estimate the three expressions

$$\big|\Re[(m_{r,t}(\mathfrak{e}_{r,t}^{-}-\kappa)-m_{r,t}(\mathfrak{e}_{r,t}^{-}))-(m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}-\kappa)-m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}))]\big|,\tag{4-26a}$$

$$\big|\Re[(m_{r,t}(\mathfrak{e}_{r,t}^{-})-m_{r,t}(\mathfrak{e}_{r,t}^{+}))-(m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-})-m_{\mu,t}(\mathfrak{e}_{\mu,t}^{+}))]\big|,\tag{4-26b}$$

$$\big|\Re[m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}-\kappa+\Delta_{\mu,t}-\Delta_{r,t})-m_{\mu,t}(\mathfrak{e}_{\mu,t}^{-}-\kappa)]\big|.\tag{4-26c}$$

In order to bound the first term we use that estimating (4-26a) for $\kappa\geq-\frac{3}{4}\Delta_{r,t}$ is equivalent to estimating the left-hand side of (4-22a) for $\nu\geq-\frac{3}{4}\Delta_{r,t}$, i.e., the regime we already considered above. This equivalence follows by using the reflection $A\to-A$ of the expectation (see (3-1)) that turns every left edge $\mathfrak{e}^{+}_{z,t}$ into a right edge $\mathfrak{e}^{-}_{z,t}$. In particular, by the analysis that we already performed (4-26a) is bounded by $|\kappa|^{\frac{1}{3}}[|\kappa|^{\frac{1}{3}}+(t_*-t)^{\frac{1}{3}}]|\log|\kappa||$. Since $|\kappa|\le|\nu|$, this is the desired bound.

For the second term (4-26b) we see from (4-23) that we have to estimate the difference between the expressions

$$\int_0^\infty \frac{\Delta_{r,t}\rho_{r,t}(\mathfrak{e}^{+}_{r,t}+\omega)}{\omega(\omega+\Delta_{r,t})}\,\mathrm{d}\omega+\int_0^\infty \frac{\Delta_{r,t}\rho_{r,t}(\mathfrak{e}^{-}_{r,t}-\omega)}{\omega(\omega+\Delta_{r,t})}\,\mathrm{d}\omega \tag{4-27}$$

for $r=\alpha,\lambda,\mu$. The summands in (4-27) are treated analogously, so we focus on the first summand. We split the integrand of the difference between the first summands and estimate

$$\frac{(\Delta_{r,t}-\Delta_{\mu,t})\rho_{r,t}(\mathfrak{e}^{+}_{r,t}+\omega)}{(\omega+\Delta_{r,t})(\omega+\Delta_{\mu,t})}+\frac{\Delta_{\mu,t}}{\omega(\omega+\Delta_{\mu,t})}(\rho_{r,t}(\mathfrak{e}^{+}_{r,t}+\omega)-\rho_{\mu,t}(\mathfrak{e}^{+}_{\mu,t}+\omega))\lesssim\frac{\Delta(\omega^{\frac{1}{3}}+(t_*-t)^{\frac{1}{3}})}{\omega^{\frac{2}{3}}(\omega+\Delta)},$$

where $\Delta:=\Delta_{r,t}\sim\Delta_{\mu,t}$ and we used (4-6a), (4-6b) and the first inequality of (4-24). Thus

$$\left|\int_0^\infty \frac{\Delta_{r,t}\rho_{r,t}(\mathfrak{e}^{+}_{r,t}+\omega)}{\omega(\omega+\Delta_{r,t})}\,\mathrm{d}\omega-\int_0^\infty \frac{\Delta_{\mu,t}\rho_{\mu,t}(\mathfrak{e}^{+}_{\mu,t}+\omega)}{\omega(\omega+\Delta_{\mu,t})}\,\mathrm{d}\omega\right|\lesssim\Delta^{\frac{2}{3}}+\Delta^{\frac{1}{3}}(t_*-t)^{\frac{1}{3}}.$$

Since $|\nu|\gtrsim\Delta$, this finishes the estimate on (4-26b).

For (4-26c) we use the $\frac{1}{3}$-Hölder regularity of $m_{\mu,t}$ and (4-6a) to get an upper bound

$$\Delta^{\frac{1}{3}}(t_*-t)^{\frac{1}{9}}\lesssim(t_*-t)^{\frac{11}{18}}.$$

This finishes the proof of (4-22a).

We now turn to the case of a small local minimum in (4-22b) and compute for $r=\alpha,\lambda,\mu$ and $\nu\neq0$

$$\Re[m_{r,t}(\mathfrak{m}_{r,t}+\nu)-m_{r,t}(\mathfrak{m}_{r,t})]=\int_{\mathbb{R}}\frac{\nu\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)}{(\omega-\nu)\omega}\,\mathrm{d}\omega.$$

Without loss of generality, we consider the case $\nu>0$, as $\nu<0$ is completely analogous. As before, we first pick a threshold $\frac{1}{2}\epsilon\le\nu$ and single out the integration over $[-\epsilon,\epsilon]$ and $[\nu-\epsilon,\nu+\epsilon]$. From the $\frac{1}{3}$-Hölder continuity of $\rho_{r,t}$ we have, for $r=\lambda,\mu$,

$$\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)=\rho_{r,t}(\mathfrak{m}_{r,t}+\nu)+\mathcal{O}(|\nu-\omega|^{\frac{1}{3}})$$

and therefore

$$\left|\int_{-\epsilon}^{\epsilon}\frac{\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)}{\omega-\nu}\,\mathrm{d}\omega\right|\lesssim\frac{\epsilon}{\nu},\qquad\left|\int_{-\epsilon}^{\epsilon}\frac{\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)}{\omega}\,\mathrm{d}\omega\right|\lesssim\int_{-\epsilon}^{\epsilon}|\omega|^{-\frac{2}{3}}\,\mathrm{d}\omega\lesssim\epsilon^{\frac{1}{3}}$$

and

$$\left|\int_{\nu-\epsilon}^{\nu+\epsilon}\frac{\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)}{\omega-\nu}\,\mathrm{d}\omega\right|\lesssim\int_{\nu-\epsilon}^{\nu+\epsilon}|\omega-\nu|^{-\frac{2}{3}}\,\mathrm{d}\omega\lesssim\epsilon^{\frac{1}{3}},\qquad\left|\int_{\nu-\epsilon}^{\nu+\epsilon}\frac{\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)}{\omega}\,\mathrm{d}\omega\right|\lesssim\frac{\epsilon}{\nu}.$$

We now consider the difference between $\rho_{r,t}$ and $\rho_{\mu,t}$ for which we have

$$|\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)-\rho_{\mu,t}(\mathfrak{m}_{\mu,t}+\omega)|\lesssim(t-t_*)|\omega|^{\frac{1}{3}}(t-t_*)^{\frac{1}{4}}+(t-t_*)^{\frac{3}{4}}+|\omega|^{\frac{2}{3}}$$

from (4-6d), (4-6c) and the $\frac{1}{3}$-Hölder continuity of $\rho_{r,t}$. Thus we can estimate

$$\begin{aligned}\left|\left[\int_{-\infty}^{-\epsilon}+\int_{\epsilon}^{\nu-\epsilon}+\int_{\nu+\epsilon}^{\infty}\right]\frac{\nu\big(\rho_{\lambda,t}(\mathfrak{m}_{r,t}+\omega)-\rho_{r,t}(\mathfrak{m}_{r,t}+\omega)\big)}{(\omega-\nu)\omega}\,\mathrm{d}\omega\right| \\ \lesssim\left[\int_{-\infty}^{-\epsilon}+\int_{\epsilon}^{\nu-\epsilon}+\int_{\nu+\epsilon}^{\infty}\right]\frac{\nu\big(|\omega|^{\frac{1}{3}}(t-t_*)^{\frac{1}{4}}+(t-t_*)^{\frac{3}{4}}+|\omega|^{\frac{2}{3}}\big)}{|\omega-\nu|\omega}\,\mathrm{d}\omega \\ \lesssim|\log\epsilon|\,[\nu^{\frac{1}{3}}(t-t_*)^{\frac{1}{4}}+(t-t_*)^{\frac{3}{4}}+\nu^{\frac{2}{3}}].\end{aligned}$$

We again choose $\epsilon=\nu^2$ and by collecting the various error estimates can conclude (4-22b). □

5. Index matching for two DBM

For two real symmetric matrix-valued standard (GOE) Brownian motions $\mathfrak{B}_t^{(\lambda)},\mathfrak{B}_t^{(\mu)}\in\mathbb{R}^{N\times N}$ we define the matrix flows

$$H_t^{(\lambda)}:=H^{(\lambda)}+\mathfrak{B}_t^{(\lambda)},\quad H_t^{(\mu)}:=H^{(\mu)}+\mathfrak{B}_t^{(\mu)}. \tag{5-1}$$

In particular, by (5-1) it follows that

$$H_t^{(\lambda)}\stackrel{d}{=}H^{(\lambda)}+\sqrt{t}U^{(\lambda)},\quad H_t^{(\mu)}\stackrel{d}{=}H^{(\mu)}+\sqrt{t}U^{(\mu)} \tag{5-2}$$

for any fixed $0\le t\le t_1$, where $U^{(\lambda)}$ and $U^{(\mu)}$ are GOE matrices. In (5-2) with $X\stackrel{d}{=}Y$ we denote that the two random variables X and Y are equal in distribution.

We will prove Proposition 3.1 by comparing the two Dyson Brownian motions for the eigenvalues of the matrices $H_t^{(\lambda)}$ and $H_t^{(\mu)}$ for $0\le t\le t_1$; see (5-3)–(5-4) below. To do this, we will use the coupling idea of [Bourgade and Yau 2017; Bourgade et al. 2016], where the DBMs for the eigenvalues of $H_t^{(\lambda)}$ and $H_t^{(\mu)}$ are coupled in such a way that the difference of the two DBMs obeys a discrete parabolic equation with good decay properties. In order to analyse this equation we consider a *short-range approximation* for the DBM, first introduced in [Erdős and Yau 2015]. Coupling only the short-range approximation of the DBMs leads to a parabolic equation whose heat kernel has a rapid off-diagonal decay by *finite speed of propagation* estimates. In this way the kernels of both DBMs are locally determined and thus can be directly compared by optimal rigidity since locally the two densities, hence their quantiles, are close. Technically it is much easier to work with a one-parameter interpolation between the two DBMs and consider its derivative with respect to the parameter, as introduced in [Bourgade and Yau 2017]; the proof of the finite-speed propagation for this dynamics does not require us to establish level repulsion, unlike in several previous works [Erdős and Schnelli 2017; Erdős and Yau 2015; Landon and Yau 2017]. However, it requires us to establish (almost) optimal rigidity for the interpolating dynamics as well. Note that optimal rigidity is known for $H_t^{(\lambda)}$ and $H_t^{(\mu)}$ from [Erdős et al. 2018], see Lemma 6.1, but not for the interpolation. For a complete picture, we mention that in the works [Erdős and Schnelli 2017; Erdős and Yau 2015; Landon and Yau 2017] on *bulk gap universality*, beyond heat kernel and Sobolev estimates, a version of the De Giorgi–Nash–Moser parabolic regularity estimate, which used level repulsion in a more substantial way than finite speed of propagation, was also necessary. *Fixed energy universality* in the bulk can be proven via homogenisation without De Giorgi–Nash–Moser estimates; hence level repulsion can also be avoided [Landon et al. 2019]. In a certain sense, the situation at the edge/cusp is easier than the bulk regime since relatively simple heat kernel bounds are sufficient for local relaxation to equilibrium. In another sense, due to singularities in the density, the edge regime and especially the cusp regime are more difficult.

In Section 6 we will establish rigidity for the interpolating process by DBM methods. Armed with this rigidity, in Section 7 we prove Proposition 3.1 for the small gap and the exact cusp case, i.e., $t_1 \le t_*$. Some estimates are slightly different for the small minimum case, i.e., $t_* \le t_1 \le 2t_*$; the modifications are given in Section 8. We recall that t_* is the time at which both $H_{t_*}^{(\lambda)}$ and $H_{t_*}^{(\mu)}$ have an exact cusp. Some technical details on the corresponding Sobolev inequality and heat kernel estimates, as well as finite speed of propagation and short-range approximation, are deferred to the Appendix; these are similar to the corresponding estimates for the edge case (see [Bourgade et al. 2014] and [Landon and Yau 2017], respectively).

In the rest of this section we prepare the proof of Proposition 3.1 by setting up the appropriate framework. While we are interested only in the eigenvalues near the physical cusp, the DBM is highly nonlocal, so we need to define the dynamics for all eigenvalues. In the setup of Proposition 3.1 we could easily assume that the cusps for the two matrix flows are formed at the same time and their slope parameters coincide — these could be achieved by a rescaling and a trivial time shift. However, the number of eigenvalues to the left of the cusp may macroscopically differ for the two ensembles, which would mean that the labels of the ordered eigenvalues near the cusp would not be constant along the interpolation. To resolve this discrepancy, we will pad the system with N fictitious particles in addition to the original flow of N eigenvalues, much as in [Landon et al. 2019], giving sufficient freedom to match the labels of the eigenvalues near the cusp. These artificial particles will be placed very far from the cusp regime and from each other so that their effect on the dynamics of the relevant particles is negligible.

With the notation of Section 4, we let $\rho_{\lambda,t}, \rho_{\mu,t}$ denote the (self-consistent) densities at time $0 \le t \le t_1$ of $H_t^{(\lambda)}$ and $H_t^{(\mu)}$, respectively. In particular, $\rho_{\lambda,0} = \rho_\lambda$ and $\rho_{\mu,0} = \rho_\mu$, where ρ_λ, ρ_μ are the self-consistent densities of $H^{(\lambda)}$ and $H^{(\mu)}$ and $\rho_{\lambda,t}, \rho_{\mu,t}$ are their semicircular evolutions. For each $0 \le t \le t_*$ both densities $\rho_{\lambda,t}, \rho_{\mu,t}$ have a small gap, denoted by $[\mathfrak{e}_{\lambda,t}^-, \mathfrak{e}_{\lambda,t}^+]$ and $[\mathfrak{e}_{\mu,t}^-, \mathfrak{e}_{\mu,t}^+]$, and we let

$$\Delta_{\lambda,t} := \mathfrak{e}_{\lambda,t}^+ - \mathfrak{e}_{\lambda,t}^-, \quad \Delta_{\mu,t} := \mathfrak{e}_{\mu,t}^+ - \mathfrak{e}_{\mu,t}^-$$

denote the lengths of these gaps. In the case of $t_* \le t \le 2t_*$, the densities $\rho_{\lambda,t}, \rho_{\mu,t}$ have a small minimum denoted by $\mathfrak{m}_{\lambda,t}$ and $\mathfrak{m}_{\mu,t}$ respectively. Since we always assume $0 \le t \le t_1 \ll 1$, both $H_t^{(\lambda)}$ and $H_t^{(\mu)}$ will always have exactly one physical cusp near $\mathfrak{c}_\lambda$ and $\mathfrak{c}_\mu$, respectively, using that the Stieltjes transform of the density is a Hölder continuous function of t; see [Alt et al. 2018a, Proposition 10.1].

Let i_λ and i_μ be the indices defined by

$$\int_{-\infty}^{\mathfrak{e}_{\lambda,0}^-} \rho_\lambda = \frac{i_\lambda - 1}{N}, \quad \int_{-\infty}^{\mathfrak{e}_{\mu,0}^-} \rho_\mu = \frac{i_\mu - 1}{N}.$$

By band rigidity (see Remark 2.6 in [Alt et al. 2018b]) i_λ and i_μ are integers. Note that by the explicit expression of the density in (2-3a)–(2-3b) it follows that $cN \le i_\lambda, i_\mu \le (1-c)N$ with some small $c > 0$, because the density on both sides of a physical cusp is macroscopic.

We let $\lambda_i(t)$ and $\mu_i(t)$ denote the eigenvalues of $H_t^{(\lambda)}$ and $H_t^{(\mu)}$ respectively. Let $\{B_i\}_{i\in[-N,N]\setminus\{0\}}$ be a family of independent standard (scalar) Brownian motions. It is well known [Dyson 1962] that the eigenvalues of $H_t^{(\lambda)}$ satisfy the equation for *Dyson Brownian motion*, i.e., the system of coupled SDEs

$$\mathrm{d}\lambda_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_{i-i_\lambda+1} + \frac{1}{N}\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j}\,\mathrm{d}t \tag{5-3}$$

with initial conditions $\lambda_i(0)=\lambda_i(H^{(\lambda)})$. Similarly, for the eigenvalues of $H_t^{(\mu)}$ we have

$$\mathrm{d}\mu_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_{i-i_\mu+1} + \frac{1}{N}\sum_{j\neq i}\frac{1}{\mu_i-\mu_j}\,\mathrm{d}t \tag{5-4}$$

with initial conditions $\mu_i(0)=\mu_i(H^{(\mu)})$. Note that we chose the Brownian motions for λ_i and $\mu_{i+i_\mu-i_\lambda}$ to be identical. This is the key ingredient for the coupling argument, since in this way the stochastic differentials will cancel when we take the difference of the two DBMs or we differentiate it with respect to an additional parameter.

For convenience of notation, we will shift the indices so that the same index labels the last quantile before the gap in ρ_λ and ρ_μ. This shift was already prepared by choosing the Brownian motions for μ_{i_μ} and λ_{i_λ} to be identical. We achieve this shift by adding N "ghost" particles very far away and relabelling, as in [Landon et al. 2019]. We thus embed λ_i and μ_i into the enlarged processes $\{x_i\}_{i\in[-N,N]\setminus\{0\}}$ and $\{y_i\}_{i\in[-N,N]\setminus\{0\}}$. Note that the index 0 is always omitted.

More precisely, the processes x_i are defined by the SDE (*extended Dyson Brownian motion*)

$$\mathrm{d}x_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j\neq i}\frac{1}{x_i-x_j}\,\mathrm{d}t, \quad 1\le |i|\le N, \tag{5-5}$$

with initial data

$$x_i(0) = \begin{cases} -N^{200}+iN & \text{if } -N\le i\le -i_\lambda, \\ \lambda_{i+i_\lambda}(0) & \text{if } 1-i_\lambda\le i\le -1, \\ \lambda_{i+i_\lambda-1}(0) & \text{if } 1\le i\le N+1-i_\lambda, \\ N^{200}+iN & \text{if } N+2-i_\lambda\le i\le N, \end{cases} \tag{5-6}$$

and the y_i are defined by

$$\mathrm{d}y_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j\neq i}\frac{1}{y_i-y_j}\,\mathrm{d}t, \quad 1\le |i|\le N, \tag{5-7}$$

with initial data

$$y_i(0) = \begin{cases} -N^{200}+iN & \text{if } -N\le i\le -i_\mu, \\ \mu_{i+i_\mu}(0) & \text{if } 1-i_\mu\le i\le -1, \\ \mu_{i+i_\mu-1}(0) & \text{if } 1\le i\le N+1-i_\mu, \\ N^{200}+iN & \text{if } N+2-i_\mu\le i\le N. \end{cases} \tag{5-8}$$

The summations in (5-5) and (5-7) extend to all j with $1\le |j|\le N$ except $j=i$.

The following lemma shows that the additional particles at distance N^{200} have negligible effect on the dynamics of the re-indexed eigenvalues; thus we can study the processes x_i and y_i instead of the eigenvalues λ_i, μ_i. The proof of this lemma follows by Appendix C of [Landon et al. 2019].

Lemma 5.1. *With very high probability the following estimates hold*:

$$\sup_{0\le t\le 1}\ \sup_{1\le i\le N+1-i_\lambda} |x_i(t)-\lambda_{i+i_\lambda-1}(t)|\le N^{-100},$$

$$\sup_{0\le t\le 1}\ \sup_{1-i_\lambda\le i\le N+1-i_\lambda} |x_i(t)-\lambda_{i+i_\lambda}(t)|\le N^{-100},$$

$$\sup_{0\le t\le 1}\ \sup_{1\le i\le N+1-i_\mu} |y_i(t)-\mu_{i+i_\mu-1}(t)|\le N^{-100},$$

$$\sup_{0\le t\le 1}\ \sup_{1-i_\mu\le i\le N+1-i_\mu} |y_i(t)-\mu_{i+i_\mu}(t)|\le N^{-100},$$

$$\sup_{0\le t\le 1} x_{-i_\lambda}(t)\lesssim -N^{200},\qquad \sup_{0\le t\le 1} x_{N+2-i_\lambda}(t)\gtrsim N^{200},$$

$$\sup_{0\le t\le 1} y_{-i_\mu}(t)\lesssim -N^{200},\qquad \sup_{0\le t\le 1} y_{N+2-i_\mu}(t)\gtrsim N^{200}.$$

Remark 5.2. For notational simplicity we assumed that $H^{(\lambda)}$ and $H^{(\mu)}$ have the same dimensions, but our proof works as long as the corresponding dimensions N_λ and N_μ are merely comparable, say $\frac{2}{3}N_\lambda\le N_\mu\le\frac{3}{2}N_\lambda$. The only modification is that the times in (5-1) need to be scaled differently in order to keep the strength of the stochastic differential terms in (5-3)–(5-4) identical. In particular, we rescale the time in the process (5-3) as $t'=(N_\mu/N_\lambda)t$, in such a way the N-scaling in front of the stochastic differential and in front of the potential term are exactly the same in both the processes (5-3) and (5-4); namely we may replace N with N_μ in both (5-3) and (5-4). Furthermore, the number of additional "ghost" particles in the *extended Dyson Brownian motion* (see (5-5) and (5-7)) will be different to ensure that we have the same total number of particles; i.e., the total number of x- and y-particles will be $2N:=2\max\{N_\mu,N_\lambda\}$, after the extension. Hence, assuming that $N_\mu\ge N_\lambda$, there will be $N=N_\mu$ particles added to the DBM of the eigenvalues of $H^{(\mu)}$ and $2N_\mu-N_\lambda$ particles added to the DBM of $H^{(\lambda)}$. In particular, under the assumption $N_\mu\ge N_\lambda$, we may replace (5-6) and (5-8) by

$$x_i(0)=\begin{cases}-N_\mu^{200}+iN_\mu & \text{if } -N_\mu\le i\le -i_\lambda,\\ \lambda_{i+i_\lambda}(0) & \text{if } 1-i_\lambda\le i\le -1,\\ \lambda_{i+i_\lambda-1}(0) & \text{if } 1\le i\le N_\lambda+1-i_\lambda,\\ N_\mu^{200}+iN_\mu & \text{if } N_\lambda+2-i_\lambda\le i\le N_\mu,\end{cases}\qquad y_i(0)=\begin{cases}-N_\mu^{200}+iN_\mu & \text{if } -N_\mu\le i\le -i_\mu,\\ \mu_{i+i_\mu}(0) & \text{if } 1-i_\mu\le i\le -1,\\ \mu_{i+i_\mu-1}(0) & \text{if } 1\le i\le N_\mu+1-i_\mu,\\ N_\mu^{200}+iN_\mu & \text{if } N_\mu+2-i_\mu\le i\le N_\mu.\end{cases}$$

Then, all the proofs of Sections 5 and 6 are exactly the same as in the case $N:=N_\mu=N_\lambda$, since all the analysis of the latter sections is done in a small, order-1 neighbourhood of the physical cusp. In particular, only the particles $x_i(t)$, $y_i(t)$ with $1\le|i|\le\epsilon\min\{N_\mu,N_\lambda\}$, for some small fixed $\epsilon>0$, will matter for our analysis. The far-away particles in the case will be treated exactly as in (5-9)–(5-13) replacing N by N_μ.

We now construct the analogues of the self-consistent densities $\rho_{\lambda,t}$, $\rho_{\mu,t}$ for the $x(t)$ and $y(t)$ processes, as well as for their α-interpolations. We start with $\rho_{x,t}$. Recall $\rho_{\lambda,t}$ from Section 4, and set

$$\rho_{x,t}(E) := \rho_{\lambda,t}(E) + \frac{1}{N}\sum_{i=-N}^{-i_\lambda} \psi(E - x_i(t)) + \frac{1}{N}\sum_{i=N+2-i_\lambda}^{N} \psi(E - x_i(t)), \quad E \in \mathbb{R}, \tag{5-9}$$

where ψ is a nonnegative symmetric approximate delta-function on scale N^{-1}, i.e., it is supported in an N^{-1} neighbourhood of zero, $\int \psi = 1$, $\|\psi\|_\infty \lesssim N$ and $\|\psi'\|_\infty \lesssim N^2$. Note that the total mass is $\int_{\mathbb{R}} \rho_{x,t} = 2$. For the Stieltjes transform $m_{x,t}$ of $\rho_{x,t}$, we have $\sup_{z\in\mathbb{C}^+}|m_{x,t}(z)| \le C$ since the same bound holds for $\rho_{\lambda,t}$ by the shape analysis. Note that $\rho_{\lambda,t}$ is the semicircular flow with initial condition $\rho_{\lambda,t=0} = \rho_\lambda$ by definition, but $\rho_{x,t}$ is not exactly the semicircular evolution of $\rho_{x,0}$. We will not need this information, but in fact, the effect of the far-away padding particles on the density near the cusp is very tiny.

Since $\rho_{x,t}$ coincides with $\rho_{\lambda,t}$ in a big finite interval, their edges and local minima near the cusp regime coincide; i.e., we can identify

$$\mathfrak{e}^\pm_{x,t} = \mathfrak{e}^\pm_{\lambda,t}, \quad \mathfrak{m}_{x,t} = \mathfrak{m}_{\lambda,t}.$$

The shifted quantiles and semiquantiles $\hat\gamma_{x,i}(t), \check\gamma_{x,i}(t)$ and $\hat\gamma^*_{x,i}(t), \check\gamma^*_{x,i}(t)$ of $\rho_{x,t}$ are defined by the obvious analogues of the formulas (4-10)–(4-12) except that the r subscript is replaced with x and the indices run over the entire range $1 \le |i| \le N$. As before, $\gamma_{x,0}(t) = \mathfrak{e}^+_{x,t}$. The unshifted quantiles are defined by

$$\gamma_{x,i}(t) = \hat\gamma_{x,i}(t) + \mathfrak{e}^+_{x,t}, \quad 0 \le t \le t_*, \qquad \gamma_{x,i}(t) = \check\gamma_{x,i}(t) + \mathfrak{m}_{x,t}, \quad t_* \le t \le 2t_*,$$

and similarly for the semiquantiles.

So far we explained how to construct $\rho_{x,t}$ and its quantiles from $\rho_{\lambda,t}$; in exactly the same way we obtain $\rho_{y,t}$ from $\rho_{\mu,t}$ with straightforward notation.

Now for any $\alpha \in [0,1]$ we construct the α-interpolation of $\rho_{x,t}$ and $\rho_{y,t}$, which we will denote by $\bar\rho_t$. The bar will indicate quantities related to α-interpolation that implicitly depend on α; a dependence that we often omit from the notation. The interpolating measure will be constructed via its quantiles; i.e., we define

$$\bar\gamma_i(t) := \alpha\hat\gamma_{x,i}(t) + (1-\alpha)\hat\gamma_{y,i}(t), \quad \bar\gamma^*_i(t) := \alpha\hat\gamma^*_{x,i}(t) + (1-\alpha)\hat\gamma^*_{y,i}(t), \quad 1 \le |i| \le N,\ 0 \le t \le t_*, \tag{5-10}$$

and similarly for $t_* \le t \le 2t_*$ involving $\check\gamma$'s. We also set the interpolating edges

$$\bar{\mathfrak{e}}^\pm_t = \alpha\mathfrak{e}^\pm_{x,t} + (1-\alpha)\mathfrak{e}^\pm_{y,t}. \tag{5-11}$$

Recall the parameter δ_* describing the size of a neighbourhood around the physical cusp where the shape analysis for ρ_λ and ρ_μ in Section 2 holds. Choose $i(\delta_*) \sim N$ such that $|\bar\gamma_{x,-i(\delta_*)}(t)| \le \delta_*$ and $|\bar\gamma_{x,i(\delta_*)}(t)| \le \delta_*$ hold for all $0 \le t \le 2t_*$. Then define, for any $E \in \mathbb{R}$, the function

$$\bar\rho_t(E) := \rho_{\alpha,t}(E)\cdot\mathbf{1}\big(\bar\gamma_{-i(\delta_*)}(t) + \bar{\mathfrak{e}}^+_t \le E \le \bar\gamma_{i(\delta_*)}(t) + \bar{\mathfrak{e}}^+_t\big) + \frac{1}{N}\sum_{i(\delta_*)<|i|\le N} \psi(E - \bar{\mathfrak{e}}^+_t - \bar\gamma^*_i(t)), \tag{5-12}$$

where $\rho_{\alpha,t}$ is the α-interpolation, constructed in Definition 4.1, between $\rho_{\lambda,t}(E) = \rho_{x,t}(E)$ and $\rho_{\mu,t}(E) = \rho_{y,t}(E)$ for $|E| \le \delta_*$. By this construction (using also the symmetry of ψ) we know that all shifted semiquantiles of $\bar{\rho}_t$ are exactly $\bar{\gamma}_i^*(t)$. The same holds for all shifted quantiles $\bar{\gamma}_i(t)$ at least in the interval $[-\delta_*, \delta_*]$ since here $\bar{\rho}_t \equiv \rho_{\alpha,t}$ and the latter was constructed exactly by the requirement of linearity of the quantiles (5-10); see (4-13c).

We also record $\int \bar{\rho}_t = 2$ and that for the Stieltjes transform $\bar{m}_t(z)$ of $\bar{\rho}_t$ we have

$$\max_{|\Re z - \bar{\mathfrak{e}}_t^+| \le \frac{1}{2}\delta_*} |\bar{m}_t(z)| \le C \tag{5-13}$$

for all $0 \le t \le 2t_*$. The first bound follows easily from the same boundedness of the Stieltjes transform of $\rho_{\alpha,t}$. Moreover, $\bar{m}_t(z)$ is $\frac{1}{3}$-Hölder continuous in the regime $|\Re z - \bar{\mathfrak{e}}_t^+| \le \frac{1}{2}\delta_*$ since in this regime $\bar{\rho}_t = \rho_{\alpha,t}$ and $\rho_{\alpha,t}$ is $\frac{1}{3}$-Hölder continuous by Lemma 4.3.

6. Rigidity for the short-range approximation

In this section we consider Dyson Brownian Motion (DBM), i.e., a system of $2N$ coupled stochastic differential equations for $z(t) = \{z_i(t)\}_{[-N,N]\setminus\{0\}}$ of the form

$$\mathrm{d}z_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_j \frac{1}{z_i - z_j}\,\mathrm{d}t, \quad 1 \le |i| \le N, \tag{6-1}$$

with some initial condition $z_i(t=0) = z_i(0)$, where $B(s) = (B_{-N}(s), \dots, B_{-1}(s), B_1(s), \dots, B_N(s))$ is the vector of $2N$ independent standard Brownian motions. We use the indexing convention that all indices i, j, etc., run from $-N$ to N but the zero index is excluded.

We will assume that $z_i(0)$ is an α-linear interpolation of $x_i(0), y_i(0)$ for some $\alpha \in [0,1]$:

$$z_i(0) = z_i(0,\alpha) := \alpha x_i(0) + (1-\alpha) y_i(0). \tag{6-2}$$

Throughout this section we will refer to the process defined by (6-1) using $z(t,\alpha)$ in order to underline the α-dependence of the process. Clearly for $\alpha = 0, 1$ we recover the original $y(t)$ and $x(t)$ processes, $z(t,\alpha=0) = y(t)$, $z(t,\alpha=1) = x(t)$. For these processes we have the following optimal rigidity estimate, which is an immediate consequence of [Erdős et al. 2018, Corollary 2.6] and Lemma 5.1:

Lemma 6.1. *Let $r_i(t) = x_i(t)$ or $r_i(t) = y_i(t)$ and $r = x, y$. Then, there exists a fixed small $\epsilon > 0$, depending only on the model parameters, such that for each $1 \le |i| \le \epsilon N$, we have*

$$\sup_{0 \le t \le 2t_*} |r_i(t) - \gamma_{r,i}(t)| \le N^\xi \eta_{\mathrm{f}}^{\rho_{r,t}}(\gamma_{r,i}(t)) \tag{6-3}$$

for any $\xi > 0$ with very high probability, where we recall that the behaviour of $\eta_{\mathrm{f}}^{\rho_{r,t}}(\gamma_{r,i}(t))$, with $r = x, y$, is given by (4-17a).

Note that, by (4-6a), (4-6c) and (4-17), for all $1 \le |i| \le \epsilon N$ and for all $0 \le t \le t_*$ we have

$$\eta_{\mathrm{f}}^{\rho_{r,t}}(\gamma_{r,i}(t)) \lesssim \frac{N^{\frac{1}{6}\omega_1}}{|i|^{\frac{1}{4}} N^{\frac{3}{4}}}, \tag{6-4}$$

with $r = x, y$.

In particular, we know that $z(0,\alpha)$ lie close to the quantiles (5-10) of an *α-interpolating density* $\rho_z = \bar\rho_0$; see the definition in (5-12). This means that ρ_z has a small gap $[\mathfrak{e}_z^-, \mathfrak{e}_z^+]$ of size $\Delta_z \sim t_*^{\frac{3}{2}}$ (i.e., it will develop a physical cusp in a time of order t_*) and it is an α-interpolation between $\rho_{x,0}$ and $\rho_{y,0}$. Here interpolation refers to the process introduced in Section 5 that guarantees that the corresponding quantiles are convex linear combinations of the two initial densities with weights α and $1-\alpha$, i.e.,

$$\gamma_{z,i} = \alpha\gamma_{x,i} + (1-\alpha)\gamma_{y,i}.$$

In this section we will prove rigidity results for $z(t,\alpha)$ and for its appropriate short-range approximation.

Remark 6.2. Before we go into the details, we point out that we will prove rigidity *dynamically*, i.e., using the DBM. The route chosen here is very different from the one in [Landon and Yau 2017, Section 6], where the authors prove a local law for short times in order to get rigidity for the short-range approximation of the interpolated process. While it would be possible to follow the latter strategy in the cusp regime as well, the technical difficulties are overwhelming; in fact already in the much simpler edge regime a large part of [Landon and Yau 2017] was devoted to this task. The current proof of the optimal law at the cusp regime [Erdős et al. 2018] heavily uses an effective mean-field condition (called *flatness*) that corresponds to large time in the DBM. Relaxing this condition would require adjusting not only [Erdős et al. 2018] but also the necessary deterministic analysis from [Alt et al. 2018a] to the short-time case. Similar complications would have arisen if we had followed the strategy of [Adhikari and Huang 2018; Huang and Landon 2016] where rigidity is proven by analysing the characteristics of the McKean–Vlasov equation. The route chosen here is shorter and more interesting.

Since the group velocity of the entire cusp regime is different for $\rho_{x,t}$ and $\rho_{y,t}$, the interpolated process will have an intermediate group velocity. Since we have to follow the process for time scales $t \sim N^{-\frac{1}{2}+\omega_1}$, much bigger than the relevant rigidity scale $N^{-\frac{3}{4}}$, we have to determine the group velocity quite precisely. Technically, we will encode this information by defining an appropriately shifted process $\tilde z(t,\alpha) = z(t,\alpha) - \mathrm{Shift}(t,\alpha)$. It is essential that the shift function is independent of the indices i to preserve the local statistics of the process. In the next section we explain how to choose the shift.

6A. ***Choice of the shifted process $\tilde z$.*** The remainder of Section 6 is formulated for the small gap regime, i.e., for $0 \le t \le t_*$. We will comment on the modifications in the small minimum regime in Section 8. To match the location of the gap, the natural guess would be to study the shifted process $z_i(t,\alpha) - \mathfrak{e}_{z,t}^+$, where $[\mathfrak{e}_{z,t}^-, \mathfrak{e}_{z,t}^+]$ is the gap of the semicircular evolution $\rho_{z,t}$ of ρ_z near the physical cusp, and approximate $z_i(t,\alpha) - \mathfrak{e}_{z,t}^+$ by the shifted semiquantiles $\hat\gamma_{z,i}^*(t)$ of $\rho_{z,t}$. However, the evolution of the semicircular flow $t \to \rho_{z,t}$ near the cusp is not sufficiently well understood. We circumvent this technical problem by considering the quantiles of another approximating density $\bar\rho_t$ defined by the requirement that its quantiles are exactly the α-linear combinations of the quantiles of $\rho_{x,t}$ and $\rho_{y,t}$ as described in Section 5. The necessary regularity properties of $\bar\rho_t$ follow directly from its construction. The precise description below assumes that $0 \le t \le 2t_*$; i.e., we are in the small gap situation. For $t_* \le t \le t_*$ an identical construction works but the reference point $\mathfrak{e}_{r,t}^+$ is replaced with the approximate minimum $\tilde{\mathfrak{m}}_{r,t}$ for $r = x, y$. For simplicity we present all formulas for $0 \le t \le t^*$ and we will comment on the other case in Section 8.

More concretely, for any fixed $\alpha \in [0,1]$ recall the (semi-)quantiles from (5-10). These are the (semi-)quantiles of the interpolating density $\bar{\rho} = \bar{\rho}_t$ defined in (5-12), and let its Stieltjes transform be denoted by $\bar{m} = \bar{m}_t$. The use of a bar will indicate quantities related to this interpolation; implicitly all quantities marked by a bar depend on the interpolation parameter α, and this dependence will be omitted from the notation. Notice that $\bar{\rho}_t$ has a gap $[\bar{\mathfrak{e}}_t^-, \bar{\mathfrak{e}}_t^+]$ near the cusp satisfying (5-11). Initially at $t = 0$ we have $\bar{\rho}_{t=0} = \rho_z$; in particular $\bar{\gamma}_i(t=0) = \hat{\gamma}_{z,i}(t=0)$ and $\bar{\mathfrak{e}}_0^{\pm} = \mathfrak{e}_z^{\pm}$. We will choose the shift in the definition of the $\tilde{z}_i(t,\alpha)$ process so that we can use $\bar{\gamma}_i^*(t)$ to trail it.

The semicircular flow and the α-interpolation do not commute, and hence $\bar{\gamma}_i(t)$ are not the same as the quantiles $\hat{\gamma}_{z,i}(t)$ of the semicircular evolution $\rho_{z,t}$ of the initial density ρ_z. We will, however, show that they are sufficiently close near the cusp and up to times relevant for us, modulo an irrelevant time-dependent shift. Notice that the evolution of $\hat{\gamma}_{z,i}(t)$ is hard to control since analysing

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\gamma}_{z,i}(t) = -\Re m_{z,t}(\gamma_{z,i}(t)) + \Re m_{z,t}(\mathfrak{e}_{z,t}^+)$$

would involve knowing the evolved density $\rho_{z,t}$ quite precisely in the critical cusp regime. While this necessary information is in principle accessible from the explicit expression for the semicircular flow and the precise shape analysis of ρ_z obtained from that of ρ_x and ρ_y, here we chose a different, technically lighter path by using $\bar{\gamma}_i(t)$. Note that unlike $\hat{\gamma}_{z,i}(t)$, the derivative of $\bar{\gamma}_i(t)$ involves only the Stieltjes transform of the densities $\rho_{x,t}$ and $\rho_{y,t}$, for which shape analysis is available.

However, the global group velocities of $\bar{\gamma}(t)$ and $\hat{\gamma}_z(t)$ are not the same near the cusp. We thus need to define $\tilde{z}(t,\alpha)$ not as $z(t,\alpha) - \bar{\mathfrak{e}}_t^+$ but with a modified time-dependent shift to make up for this velocity difference so that $\bar{\gamma}(t)$ indeed correctly follows $\tilde{z}(t,\alpha)$. To determine this shift, we first define the function

$$h^*(t,\alpha) := \Re[-\bar{m}_t(\bar{\mathfrak{e}}_t^+) + (1-\alpha)m_{y,t}(\mathfrak{e}_{y,t}^+) + \alpha m_{x,t}(\mathfrak{e}_{x,t}^+)], \tag{6-5}$$

where recall that $\bar{m}_t$ is the Stieltjes transform of the measure $\bar{\rho}_t$. Note that $h^*(t) = O(1)$ following from the boundedness of the Stieltjes transforms $m_{x,t}$, $m_{y,t}$ and $\bar{m}_t(\bar{\mathfrak{e}}_t^+)$. The boundedness of $m_{x,t}$ and $m_{y,t}$ follows by (4-1) and $|\bar{m}_t(\bar{e}_t^+)| \le C$ by (5-13).

We note that

$$h^*(t,\alpha=0) = m_{y,t}(\mathfrak{e}_{y,t}^+) - \bar{m}_t(\bar{\mathfrak{e}}_t^+) = m_{y,t}(\mathfrak{e}_{y,t}^+) - \bar{m}_t(\mathfrak{e}_{y,t}^+)$$

since for $\alpha = 0$ we have $\mathfrak{e}_{y,t}^+ = \bar{\mathfrak{e}}_t^+$ by construction. At $\alpha = 0$ the measure $\bar{\rho}_t$ is given exactly by the density $\rho_{y,t}$ in an $\mathcal{O}(1)$ neighbourhood of the cusp. Away from the cusp, depending on the precise construction in the analogue of (5-12), the continuous $\rho_{y,t}$ is replaced by locally smoothed out Dirac measures at the quantiles. A similar statement holds at $\alpha = 1$, i.e., for the density $\rho_{x,t}$. It is easy to see that the difference of the corresponding Stieltjes transforms evaluated at the cusp regime is of order N^{-1}, i.e.,

$$|h^*(t,\alpha=0)| + |h^*(t,\alpha=1)| = O(N^{-1}). \tag{6-6}$$

Since later in (6-110) we will need to give a very crude estimate on the α-derivative of $h^*(t,\alpha)$, but it actually blows up since $\bar{m}_t'$ is singular at the edge, we introduce a tiny regularisation of h^*; i.e., we define the function

$$h^{**}(t,\alpha) := \Re[-\bar{m}_t(\bar{\mathfrak{e}}_t^+ + \mathrm{i}N^{-100}) + (1-\alpha)m_{y,t}(\mathfrak{e}_{y,t}^+) + \alpha m_{x,t}(\mathfrak{e}_{x,t}^+)]. \tag{6-7}$$

Note that by the $\frac{1}{3}$-Hölder continuity of $\bar{m}_t$ in the cusp regime, i.e., for $z \in \mathbb{H}$ such that $|\Re z - \bar{\mathfrak{e}}_t^+| \le \frac{1}{2}\delta_*$, it follows that

$$h^{**}(t,\alpha) = h^*(t,\alpha) + \mathcal{O}(N^{-30}). \tag{6-8}$$

Then, we define

$$h(t) = h(t,\alpha) := h^{**}(t,\alpha) - \alpha h^{**}(t,1) - (1-\alpha)h^{**}(t,0) = O(1) \tag{6-9}$$

to ensure that

$$h(t,\alpha{=}0) = h(t,\alpha{=}1) = 0. \tag{6-10}$$

In particular, we have

$$h(t,\alpha) = \Re[-\bar{m}_t(\bar{\mathfrak{e}}_t^+) + (1-\alpha)m_{y,t}(\mathfrak{e}_{y,t}^+) + \alpha m_{x,t}(\mathfrak{e}_{x,t}^+)] + O(N^{-1}). \tag{6-11}$$

Define its antiderivative

$$H(t,\alpha) := \int_0^t h(s,\alpha)\,\mathrm{d}s, \qquad H(0,\alpha) = 0, \qquad \max_{0\le t\le t_*} |H(t,\alpha)| \lesssim N^{-\frac{1}{2}+\omega_1}. \tag{6-12}$$

Now we are ready to define the correctly shifted process

$$\tilde{z}_i(t) = \tilde{z}_i(t,\alpha) := z_i(t) - [\alpha\mathfrak{e}_{x,t}^+ + (1-\alpha)\mathfrak{e}_{y,t}^+] - H(t,\alpha), \tag{6-13}$$

which will be trailed by $\bar{\gamma}_i(t)$. It satisfies the shifted DBM

$$\mathrm{d}\tilde{z}_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \left[\frac{1}{N}\sum_{j\ne i}\frac{1}{\tilde{z}_i - \tilde{z}_j} + \Phi_\alpha(t)\right]\mathrm{d}t, \tag{6-14}$$

with

$$\Phi(t) := \Phi_\alpha(t) = \alpha\Re m_{x,t}(\mathfrak{e}_{x,t}^+) + (1-\alpha)\Re m_{y,t}(\mathfrak{e}_{y,t}^+) - h(t,\alpha), \tag{6-15}$$

and with initial conditions $\tilde{z}(0) := z(0) - \mathfrak{e}_z^+$ by (5-11) and $H(0,\alpha) = 0$. The shift function satisfies

$$\Phi_\alpha(t) = \Re[\bar{m}_t(\bar{\mathfrak{e}}_t^+)] + \mathcal{O}(N^{-1}). \tag{6-16}$$

Notice that for $\alpha = 0, 1$ this definition gives back the naturally shifted $x(t)$ and $y(t)$ processes since we clearly have

$$\tilde{z}(t,\alpha{=}1) = \tilde{x}(t) := x(t) - \mathfrak{e}_{x,t}^+, \quad \tilde{z}(t,\alpha{=}0) = \tilde{y}(t) := y(t) - \mathfrak{e}_{y,t}^+, \tag{6-17}$$

which are trailed by the shifted semiquantiles

$$\bar{\gamma}_i^*(t,\alpha{=}1) = \hat{\gamma}_{x,i}^*(t) := \gamma_{x,i}^*(t) - \mathfrak{e}_{x,t}^+, \quad \bar{\gamma}_i^*(t,\alpha{=}0) = \hat{\gamma}_{y,i}^*(t) := \gamma_{y,i}^*(t) - \mathfrak{e}_{y,t}^+. \tag{6-18}$$

As we explained, the time-dependent shift $H(t,\alpha)$ in (6-13) makes up for the difference between the true edge velocity of the semicircular flow (which we do not compute directly) and the naive guess which is $\frac{\mathrm{d}}{\mathrm{d}t}[\alpha\mathfrak{e}_{x,t}^+ + (1-\alpha)\mathfrak{e}_{y,t}^+]$ hinted at by the linear combination procedure. The precise expression (6-5) will come out of the proof. The key point is that this adjustment is time-dependent but global; i.e., it is independent of i since this expresses a group velocity of the entire cusp regime.

6B. ***Plan of the proof.*** In the following three subsections we prove an almost-optimal rigidity not directly for $\tilde{z}_i(t)$ but for its appropriate short-range approximation $\hat{z}_i(t)$. This will be sufficient for the proof of the universality. The proof of the rigidity will be divided into three phases, which we first explain informally, as follows.

Phase 1 (Section 6C): The main result is a rigidity for $\tilde{z}_i(t)-\bar{\gamma}_i(t)$ for $1 \le |i| \lesssim \sqrt{N}$ on scale $N^{-\frac{3}{4}+C\omega_1}$ without i-dependence in the error term. First we prove a crude rigidity on scale $N^{-\frac{1}{2}+C\omega_1}$ for all indices i. Using this rigidity, we can define a short-range approximation $\mathring{z}$ of the original dynamics $\tilde{z}$ and show that $\tilde{z}_i$ and $\mathring{z}_i$ are close by $N^{-\frac{3}{4}+C\omega_1}$ for $1 \le |i| \lesssim \sqrt{N}$. Then we analyse the short-range process $\mathring{z}$ that has a finite speed of propagation, so we can localise the dynamics. Finally, we can directly compare $\mathring{z}$ with a deterministic particle dynamics because the effect of the stochastic term $\sqrt{2/N}\,\mathrm{d}B_i$, i.e., $\sqrt{t_*/N} = N^{-\frac{3}{4}+\frac{1}{2}\omega_1} \ll N^{-\frac{3}{4}+C\omega_1}$, remains below the rigidity scale of interest in this Phase 1.

However, to understand this deterministic particle dynamics we need to compare it with the corresponding continuum evolution; this boils down to estimating the difference of a Stieltjes transform and its Riemann sum approximation at the semiquantiles. Since the Stieltjes transform is given by a singular integral, this approximation relies on quite delicate cancellations which require some strong regularity properties of the density. We can easily guarantee this regularity by considering the density $\bar{\rho}_t$ of the linear interpolation between the quantiles of $\rho_{x,t}$ and $\rho_{y,t}$.

Phase 2 (Section 6D): In this section we improve the rigidity from scale $N^{-\frac{3}{4}+C\omega_1}$ to scale $N^{-\frac{3}{4}+\frac{1}{6}\omega_1}$, for a smaller range of indices, but we can achieve this not for $\tilde{z}$ directly, but for its short-range approximation $\hat{z}$. Unlike $\mathring{z}$ in Phase 1, this time we choose a very short-scale approximation $\hat{z}$ on scale $N^{4\omega_\ell}$ with $\omega_1 \ll \omega_\ell \ll 1$. As an input, we need the rigidity of $\tilde{z}_i$ on scale $N^{-\frac{3}{4}+C\omega_1}$ for $1 \le |i| \lesssim \sqrt{N}$ obtained in Phase 1. We use heat kernel contraction for a direct comparison with the $y_i(t)$ dynamics for which we know optimal rigidity by [Erdős et al. 2018], with the precise matching of the indices (*band rigidity*). In particular, when the gap is large, this guarantees that band rigidity is transferred to the $\hat{z}$-process from the $\hat{y}$-process.

Phase 3 (Section 6E): Finally, we establish the optimal i-dependence in the rigidity estimate for $\hat{z}_i$ from Phase 2; i.e., we get a precision $N^{-\frac{3}{4}+\frac{1}{6}\omega_1}|i|^{-\frac{1}{4}}$. The main method we use in Phase 3 is the maximum principle. We compare $\hat{z}_i$ with $\hat{y}_{i-K}$, a slightly shifted element of the $\hat{y}$-process, where $K = N^{\xi}$ with some tiny ξ. This method allows us to prove the optimal i-dependent rigidity (with a factor $N^{\frac{1}{6}\omega_1}$) but only for indices $|i| \gg K$ because otherwise $\hat{z}_i$ and $\hat{y}_{i-K}$ may be on different sides of the gap for small i. For very small indices, therefore, we need to rely on band rigidity for $\hat{z}$ from Phase 2.

The optimal i-dependence allows us to replace the random particles $\hat{z}$ by appropriate quantiles with a precision so that

$$|\hat{z}_i - \hat{z}_j| \lesssim N^{\frac{1}{6}\omega_1}|\bar{\gamma}_i - \bar{\gamma}_j| \sim N^{-\frac{3}{4}+\frac{1}{6}\omega_1}\big||i|^{\frac{3}{4}} - |j|^{\frac{3}{4}}\big|.$$

Such an upper bound on $|\hat{z}_i - \hat{z}_j|$, hence a lower bound on the interaction kernel $\mathcal{B}_{ij} = |\hat{z}_i - \hat{z}_j|^{-2}$ of the differentiated DBM (see (6-106)) with the correct dependence on the indices i, j, is essential since this gives the heat kernel contraction which eventually drives the precision below the rigidity scale in order to prove universality. On a time scale $t_* = N^{-\frac{1}{2}+\omega_1}$ the $\ell^p \to \ell^\infty$ contraction of the heat kernel gains a factor $N^{-\frac{4}{15}\omega_1}$ with the convenient choice of $p=5$. Notice that $\frac{4}{15}>\frac{1}{6}$, so the contraction wins over the imprecision in the rigidity $N^{\frac{1}{6}\omega_1}$ from Phase 3, but not over $N^{C\omega_1}$ from Phase 1, showing that both Phase 2 and Phase 3 are indeed necessary.

6C. ***Phase 1*: *rigidity for $\tilde{z}$ on scale $N^{-\frac{3}{4}+C\omega_1}$*.** The main result of this section is the following:

Proposition 6.3. *Fix $\alpha\in[0,1]$. Let $\tilde{z}(t,\alpha)$ solve* (6-14) *with initial condition $\tilde{z}_i(0,\alpha)$ satisfying the crude rigidity bound for all indices*

$$\max_{1\le|i|\le N}|\tilde{z}_i(0,\alpha)-\bar{\gamma}_i^*(0)|\lesssim N^{-\frac{1}{2}+2\omega_1}. \tag{6-19}$$

We also assume that

$$\|m_{x,0}\|_\infty+\|m_{y,0}\|_\infty+|\bar{m}_t(\mathfrak{e}_t^{\pm})|\le C. \tag{6-20}$$

Then we have a weak but uniform rigidity

$$\sup_{0\le t\le t_*}\max_{1\le|i|\le N}|\tilde{z}_i(t,\alpha)-\bar{\gamma}_i^*(t)|\lesssim N^{-\frac{1}{2}+2\omega_1} \tag{6-21}$$

with very high probability. Moreover, for small $|i|$, i.e., $1\le|i|\le i_*$, with $i_*:=N^{\frac{1}{2}+C_*\omega_1}$ for some large $C_*>100$, we have a stronger rigidity:

$$\sup_{0\le t\le t_*}\max_{1\le|i|\le i_*}|\tilde{z}_i(t,\alpha)-\bar{\gamma}_i^*(t)|\lesssim\max_{1\le|i|\le 2i_*}|\tilde{z}_i(0,\alpha)-\bar{\gamma}_i^*(0)|+\frac{N^{C\omega_1}}{N^{\frac{3}{4}}} \tag{6-22}$$

with very high probability.

In our application, (6-19) is satisfied and the right-hand side of (6-22) is simply $N^{-\frac{3}{4}+C\omega_1}$ since

$$\tilde{z}_i(0,\alpha)-\bar{\gamma}_i^*(0)=\alpha(x_i(0)-\gamma_{x,i}(0))+(1-\alpha)(y_i(0)-\gamma_{y,i}(0))=O\biggl(\frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}\biggr) \tag{6-23}$$

for any $\xi>0$ with very high probability, by optimal rigidity for $x_i(0)$ and $y_i(0)$ from [Erdős et al. 2018]. Similarly, the assumption (6-20) is trivially satisfied by (5-13). However, we stated Proposition 6.3 under the slightly weaker conditions (6-19), (6-20) to highlight what is really needed for its proof.

Before starting the proof, we recall the formula

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\gamma}_{i,r}^*(t)=-\Re m_{r,t}(\gamma_{r,i}^*(t))+\Re m_{r,t}(\mathfrak{e}_{r,t}^+),\quad r=x,y, \tag{6-24}$$

on the derivative of the (shifted) semiquantiles of a density which evolves by the semicircular flow and follows directly from (4-14c) and (4-14e).

Proof of Proposition 6.3. We start with the proof of the crude rigidity (6-21), then we introduce a short-range approximation and finally, with its help, we prove the refined rigidity (6-22). The main technical input of the last step is a refined estimate on the forcing term. These four steps will be presented in the next four subsections.

6C1. *Proof of the crude rigidity.* For the proof of (6-21), using (6-24) twice in (5-10), we notice that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\gamma}_i^*(t) = \alpha[-\Re m_{x,t}(\gamma_{x,i}^*(t)) + \Re m_{x,t}(\mathfrak{e}_{x,t}^+)] + (1-\alpha)[-\Re m_{y,t}(\gamma_{y,i}^*(t)) + \Re m_{y,t}(\mathfrak{e}_{y,t}^+)] = O(1)$$

since $m_{x,t}$ and $m_{y,t}$ are bounded, recalling that the semicircular flow preserves (or reduces) the ℓ^∞ norm of the Stieltjes transform by (4-1), so $\|m_{x,t}\|_\infty \le \|m_{x,0}\|_\infty \le C$, and similarly for $m_{y,t}$. This gives

$$|\bar{\gamma}_i^*(t) - \bar{\gamma}_i^*(0)| \lesssim N^{-\frac{1}{2}+\omega_1}. \tag{6-25}$$

Thus in order to prove (6-21) it is sufficient to prove

$$\|\tilde{z}(t,\alpha) - \tilde{z}(0,\alpha)\|_\infty \le N^{-\frac{1}{2}+2\omega_1} \tag{6-26}$$

for any fixed $\alpha \in [0,1]$. To do that, we compare the dynamics of (6-14) with the dynamics of the y-semiquantiles; i.e., set

$$u_i := u_i(t,\alpha) = \tilde{z}_i(t) - \hat{\gamma}_{y,i}^*(t)$$

for all $0 \le t \le t_*$.

Compute

$$\mathrm{d}u_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + (\widetilde{\mathcal{B}}u)_i\,\mathrm{d}t + \widetilde{F}_i(t)\,\mathrm{d}t, \tag{6-27}$$

with

$$(\widetilde{\mathcal{B}}f)_i := \frac{1}{N}\sum_{j\ne i}\frac{f_j - f_i}{(\tilde{z}_i - \tilde{z}_j)(\hat{\gamma}_{y,i}^* - \hat{\gamma}_{y,j}^*)} \tag{6-28}$$

and

$$\widetilde{F}_i(t) := \frac{1}{N}\sum_{j\ne i}\frac{1}{\hat{\gamma}_{y,i}^* - \hat{\gamma}_{y,j}^*} + \Re m_{y,t}(\gamma_{y,i}^*(t)) + \alpha[\Re m_{x,t}(\mathfrak{e}_{x,t}^+) - \Re m_{y,t}(\mathfrak{e}_{y,t}^+)] - h(t).$$

The operator $\widetilde{\mathcal{B}}$ is defined on $\mathbb{C}^{2N}$ and we label the vectors $f \in \mathbb{C}^{2N}$ as

$$f = (f_{-N}, f_{-N+1}, \dots, f_{-1}, f_1, \dots, f_N);$$

i.e., we omit the $i = 0$ index. Accordingly, in the summations the $j = 0$ term is always omitted since $\tilde{z}_j$, $\hat{z}_j$ and $\hat{\gamma}_{y,j}^*$ are defined for $1 \le |j| \le N$. Furthermore in the summation of the interaction terms, the $j = i$ term is always omitted.

We now show that

$$\|\widetilde{F}(t)\|_\infty \lesssim \log N, \quad 0 \le t \le t^*. \tag{6-29}$$

By the boundedness of $m_{x,t}, m_{y,t}$ and the $\frac{1}{3}$-Hölder continuity of $\bar{m}_t$ in the cusp regime, it remains to control

$$\frac{1}{N}\sum_{j\ne i}\frac{1}{\hat{\gamma}_{y,i}^*(t) - \hat{\gamma}_{y,j}^*(t)} \lesssim \sum_{1\le|j-i|\le N}\frac{1}{|i-j|} \lesssim \log N$$

since $|\hat{\gamma}_{y,j}^* - \hat{\gamma}_{y,i}^*| \ge c|i-j|/N$ as the density $\rho_{y,t}$ is bounded.

Let $\widetilde{\mathcal{U}}(s,t)$ be the fundamental solution of the heat evolution with kernel $\widetilde{\mathcal{B}}$ from (6-28); i.e, for any $0 \le s \le t$

$$\partial_t \widetilde{\mathcal{U}}(s,t) = \widetilde{\mathcal{B}}(t)\widetilde{\mathcal{U}}(s,t), \quad \widetilde{\mathcal{U}}(s,s) = I. \tag{6-30}$$

Note that $\widetilde{\mathcal{U}}$ is a contraction on every ℓ^p space and the same is true for its adjoint $\widetilde{\mathcal{U}}^*(s,t)$. In particular, for any indices a, b and times s, t we have

$$\widetilde{\mathcal{U}}_{ab}(s,t) \le 1, \quad \widetilde{\mathcal{U}}^*_{ab}(s,t) \le 1. \tag{6-31}$$

By Duhamel's principle, the solution to the SDE (6-27) is given by

$$u(t) = \widetilde{\mathcal{U}}(0,t)u(0) + \sqrt{\frac{2}{N}} \int_0^t \widetilde{\mathcal{U}}(s,t)\,\mathrm{d}B(s) + \int_0^t \widetilde{\mathcal{U}}(s,t)\widetilde{F}(s)\,\mathrm{d}s, \tag{6-32}$$

where $B(s) = (B_{-N}(s), \dots, B_{-1}(s), B_1(s), \dots, B_N(s))$ are the $2N$ independent Brownian motions from (6-1).

For the second term in (6-32) we fix an index i and consider the martingale

$$M_t := \sqrt{\frac{2}{N}} \int_0^t \sum_j \widetilde{\mathcal{U}}_{ij}(s,t)\,\mathrm{d}B_j(s),$$

with its quadratic variation process

$$[M]_t := \frac{2}{N} \int_0^t \sum_j (\widetilde{\mathcal{U}}_{ij}(s,t))^2\,\mathrm{d}s = \frac{2}{N} \int_0^t \|\widetilde{\mathcal{U}}^*(s,t)\delta_i\|_2^2\,\mathrm{d}s \le \frac{2t}{N}.$$

By the Burkholder maximal inequality for martingales, for any $p > 1$ we have

$$\mathbb{E} \sup_{0 \le t \le T} |M_t|^{2p} \le C_p\,\mathbb{E}[M]_T^p \le C_p \frac{T^p}{N^p}.$$

By Markov inequality we obtain

$$\sup_{0 \le t \le T} |M_t| \le N^\xi \sqrt{\frac{T}{N}} \tag{6-33}$$

with probability more than $1 - N^{-D}$, for any (large) $D > 0$ and (small) $\xi > 0$.

The last term in (6-32) is estimated, using (6-29), by

$$\left| \int_0^t \widetilde{\mathcal{U}}(s,t)\widetilde{F}(s)\,\mathrm{d}s \right| \le t \max_{s \le t} \|\widetilde{F}(s)\|_\infty \lesssim t \log N. \tag{6-34}$$

This, together with (6-33) and the contraction property of $\widetilde{\mathcal{B}}$ implies from (6-32) that

$$\|u(t) - u(0)\|_\infty \lesssim N^{-\frac{3}{4}+\omega_1} + t \log N \lesssim N^{-\frac{1}{2}+2\omega_1}$$

with very high probability. Recalling the definition of u and (6-25), we get (6-26) since

$$\|\tilde{z}(t) - \tilde{z}(0)\|_\infty \le \|u(t) - u(0)\|_\infty + \|\hat{\gamma}^*_y(t) - \hat{\gamma}^*_y(0)\|_\infty \lesssim N^{-\frac{1}{2}+2\omega_1}.$$

This completes the proof of the crude rigidity bound (6-21).

6C2. *Crude short-range approximation.* Now we turn to the proof of (6-22) by introducing a short-range approximation of the dynamics (6-14). Fix an integer L. Let $\mathring{z}_i = \mathring{z}_i(t)$ solve the *L-localised short-scale DBM*

$$\mathrm{d}\mathring{z}_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j:|j-i|\le L}\frac{1}{\mathring{z}_i - \mathring{z}_j}\,\mathrm{d}t + \left[\frac{1}{N}\sum_{j:|j-i|>L}\frac{1}{\bar{\gamma}_i^* - \bar{\gamma}_j^*} + \Phi(t)\right]\mathrm{d}t \tag{6-35}$$

for each $1 \le |i| \le N$ and with initial data $\mathring{z}_i(0) := \tilde{z}_i(0)$, where we recall that Φ was defined in (6-15). Then, we have the following comparison:

Lemma 6.4. *Fix* $\alpha \in [0, 1]$. *Assume that*

$$\max_{1\le|i|\le N}|\tilde{z}_i(0,\alpha) - \bar{\gamma}_i^*(0)| \lesssim N^{-\frac{1}{2}+2\omega_1}. \tag{6-36}$$

Consider the short-scale DBM (6-35) *with a range* $L = N^{\frac{1}{2}+C_1\omega_1}$ *with a constant* $10 \le C_1 \ll C_*$; *in particular* L *is much smaller than* i_*. *Then we have a weak uniform comparison*

$$\sup_{0\le t\le t_*}\max_{1\le|i|\le N}|\mathring{z}_i(t,\alpha) - \tilde{z}_i(t,\alpha)| \lesssim N^{-\frac{1}{2}+2\omega_1}, \tag{6-37}$$

and a stronger comparison for small i

$$\sup_{0\le t\le t_*}\max_{1\le|i|\le i_*}|\mathring{z}_i(t,\alpha) - \tilde{z}_i(t,\alpha)| \lesssim N^{-\frac{3}{4}+C\omega_1}, \tag{6-38}$$

both with very high probability.

Proof. For any fixed $\alpha \in [0, 1]$ and for all $0 \le t \le t_*$, set $w := w(t,\alpha) = \mathring{z}(t,\alpha) - \tilde{z}(t,\alpha)$ and subtract (6-35) and (6-14) to get

$$\partial_t w = \mathring{\mathcal{B}}_1 w + \mathring{F},$$

where

$$(\mathring{\mathcal{B}}_1 f)_i := \frac{1}{N}\sum_{j:|j-i|\le L}\frac{f_j - f_i}{(\mathring{z}_i - \mathring{z}_j)(\tilde{z}_i - \tilde{z}_j)}, \quad \mathring{F}_i := \frac{1}{N}\sum_{j:|j-i|>L}\left[\frac{1}{\bar{\gamma}_i^* - \bar{\gamma}_j^*} - \frac{1}{\tilde{z}_i - \tilde{z}_j}\right].$$

We estimate

$$|\mathring{F}_i| \le \frac{1}{N}\sum_{j:|j-i|>L}\frac{|\tilde{z}_i - \bar{\gamma}_i^*| + |\tilde{z}_j - \bar{\gamma}_j^*|}{(\bar{\gamma}_i^* - \bar{\gamma}_j^*)(\tilde{z}_i - \tilde{z}_j)} \lesssim \frac{N^{-\frac{1}{2}+2\omega_1}}{N}\sum_{j:|j-i|>L}\frac{1}{(\bar{\gamma}_i^* - \bar{\gamma}_j^*)(\tilde{z}_i - \tilde{z}_j)},$$

where we used the crude rigidity (6-21) (applicable by (6-36)), and we chose C_1 in $L = N^{\frac{1}{2}+C_1\omega_1}$ large enough so that $|\bar{\gamma}_i^* - \bar{\gamma}_j^*|$ for any $|i-j| \ge L$ is much bigger than the rigidity scale $N^{-\frac{1}{2}+2\omega_1}$ in (6-21). This is guaranteed since

$$|\bar{\gamma}_i^* - \bar{\gamma}_j^*| = \alpha|\hat{\gamma}_{x,i}^* - \hat{\gamma}_{x,j}^*| + (1-\alpha)|\hat{\gamma}_{y,i}^* - \hat{\gamma}_{y,j}^*| \gtrsim \frac{|i-j|}{N} \gtrsim N^{-\frac{1}{2}+C_1\omega}$$

with very high probability. By this choice of L we have $|\tilde{z}_i - \tilde{z}_j| \sim |\bar{\gamma}_i^* - \bar{\gamma}_j^*|$ and therefore

$$|\mathring{F}_i| \lesssim \frac{N^{-\frac{1}{2}+2\omega_1}}{N} \sum_{j:|j-i|>L} \frac{1}{(\bar{\gamma}_i^* - \bar{\gamma}_j^*)^2} \lesssim N^{\frac{1}{2}+2\omega_1} \sum_{j:|j-i|>L} \frac{1}{|i-j|^2}$$
$$\lesssim N^{-(\frac{1}{2}C_1-2)\omega_1} \le 1 \quad \text{for all } |i| \le N. \tag{6-39}$$

Since $\mathcal{B}_1$ is positivity-preserving, its evolution is a contraction, so by a Duhamel formula, similarly to (6-32), we get

$$\|\mathring{z}(t) - \tilde{z}(t)\|_\infty = \|w(t)\|_\infty \le \|w(0)\|_\infty + t \max_{s\le t} \|\mathring{F}(s)\|_\infty \lesssim N^{-\frac{1}{2}+\omega_1}$$

with very high probability.

Next, we proceed with the proof of (6-38).

In fact, for $1 \le |i| \le 2i_*$, with i_* much bigger than L, we have a better bound:

$$|\mathring{F}_i| \lesssim \frac{N^{-\frac{1}{2}+2\omega_1}}{N} \sum_{j:|j-i|>L} \frac{1}{(\bar{\gamma}_i^* - \bar{\gamma}_j^*)^2} \lesssim \sum_{j:|j-i|>L} \frac{N^{2\omega_1}}{||i|^{\frac{3}{4}} - |j|^{\frac{3}{4}}|^2}$$
$$\lesssim N^{-\frac{1}{4}-(\frac{1}{2}C_1-2)\omega_1} \le N^{-\frac{1}{4}}, \quad |i| \le 2i_*, \tag{6-40}$$

which we can use to get the better bound (6-38). To do so, we define a continuous interpolation $v(t,\beta)$ between $\tilde{z}$ and $\mathring{z}$. More precisely, for any fixed $\beta \in [0,1]$ we set $v(t,\beta) = \{v(t,\beta)_i\}_{i=-N}^N$ as the solution to the SDE

$$\mathrm{d}v_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N} \sum_{j:|j-i|\le L} \frac{1}{v_i - v_j}\,\mathrm{d}t + \Phi_\alpha(t)\,\mathrm{d}t$$
$$+ \frac{1-\beta}{N} \sum_{j:|j-i|>L} \frac{1}{\tilde{z}_i - \tilde{z}_j}\,\mathrm{d}t + \frac{\beta}{N} \sum_{j:|j-i|>L} \frac{1}{\bar{\gamma}_i^* - \bar{\gamma}_j^*}\,\mathrm{d}t \tag{6-41}$$

with initial condition $v(t=0,\beta) = (1-\beta)\tilde{z}_i(0) + \beta\mathring{z}_i(0)$. Clearly $v(t,\beta=0) = \tilde{z}(t)$ and $v(t,\beta=1) = \mathring{z}(t)$.

Differentiating in β, for $u := u(t,\beta) = \partial_\beta v(t,\beta)$ we obtain the SDE

$$\mathrm{d}u_i = (\mathcal{B}^v u)_i\,\mathrm{d}t + \mathring{F}_i\,\mathrm{d}t, \quad \text{with } (\mathcal{B}^v f)_i := \frac{1}{N} \sum_{j:|j-i|\le L} \frac{f_j - f_i}{(v_i - v_j)^2}, \tag{6-42}$$

with initial condition $u(t=0,\beta) = \mathring{z}(0) - \tilde{z}(0) = 0$. By the contraction property of the heat evolution kernel $\mathcal{U}^v$ of $\mathcal{B}^v$, with a simple Duhamel formula, we have for any fixed β

$$\sup_{0\le t\le t_*} \|u(t,\beta)\|_\infty \le t_* \|\mathring{F}\|_\infty \le N^{-\frac{1}{2}+\frac{3}{2}\omega_1} \tag{6-43}$$

with very high probability, where we used (6-39). After integration in β we get

$$\|v(t,\beta) - \bar{\gamma}^*(t)\|_\infty \le \|v(t,0) - \bar{\gamma}^*(t)\|_\infty + \left\| \int_0^\beta u(t,\beta')\,\mathrm{d}\beta' \right\|_\infty, \quad 0 \le t \le t_*, \quad \beta \in [0,1]. \tag{6-44}$$

From (6-43) we have

$$\mathbb{E}\left\| \int_0^\beta u(t,\beta')\,\mathrm{d}\beta' \right\|_\infty^p \le \int_0^\beta \mathbb{E}\|u(t,\beta')\|^p\,\mathrm{d}\beta' \lesssim (N^{-\frac{1}{2}+\frac{3}{2}\omega_1})^p \tag{6-45}$$

for any exponent p. Hence, using a high-moment Markov inequality, we have

$$\mathbb{P}\left(\left\| \int_0^\beta u(t,\beta')\,\mathrm{d}\beta' \right\|_\infty \ge N^{-\frac{1}{2}+\frac{3}{2}\omega_1+\xi} \right) \le N^{-D} \tag{6-46}$$

for any (large) $D > 0$ and (small) $\xi > 0$ by choosing p large enough. Since $v(t,0) = \tilde{z}(t)$, for which we have rigidity in (6-21), by (6-44) and (6-46) we conclude that

$$\sup_{0\le t\le t_*} \|v(t,\beta) - \bar{\gamma}^*(t)\|_\infty \lesssim N^{-\frac{1}{2}+2\omega_1} \tag{6-47}$$

with very high probability for any $\beta \in [0,1]$.

In particular L is much larger than the rigidity scale of $v = v(t,\beta)$. This means that

$$||v_i - v_j| - |\bar{\gamma}_i^* - \bar{\gamma}_j^*|| \lesssim N^{-\frac{1}{2}+2\omega_1}$$

and

$$|\bar{\gamma}_i^* - \bar{\gamma}_j^*| \gtrsim \frac{|i-j|}{N} \ge N^{-\frac{1}{2}+C_1\omega_1} \gg N^{-\frac{1}{2}+2\omega_1}$$

whenever $|i-j| \ge L$, so we have

$$|v_i - v_j| \sim |\bar{\gamma}_i^* - \bar{\gamma}_j^*|, \quad |i-j| \ge L. \tag{6-48}$$

Since i_* is much bigger than L and L is much larger than the rigidity scale of $v_i(t,\beta)$ in the sense of (6-48), the heat evolution kernel $\mathcal{U}^v$ satisfies the following finite speed of propagation estimate (the proof is given in Appendix B):

Lemma 6.5. *With the notation above we have*

$$\sup_{0\le s\le t\le t_*} [\mathcal{U}_{pi}^v + \mathcal{U}_{ip}^v] \le N^{-D}, \quad 1 \le |i| \le i_*, \ |p| \ge 2i_*, \tag{6-49}$$

for any D if N is sufficiently large.

Using a Duhamel formula again, for any fixed β, we have

$$u_i(t) = \sum_p \mathcal{U}_{ip}^v u_p(0) + \int_0^t \sum_p \mathcal{U}_{ip}^v(s,t) \mathring{F}_p(s)\,\mathrm{d}s.$$

We can split the summation and estimate

$$|u_i(t)| \le \left[\sum_{|p|\le 2i_*} + \sum_{|p|>2i_*} \right] \mathcal{U}_{ip}^v |u_p(0)| + \int_0^t \left[\sum_{|p|\le 2i_*} + \sum_{|p|>2i_*} \right] \mathcal{U}_{ip}^v(s,t) |\mathring{F}_p(s)|\,\mathrm{d}s.$$

For $|i| \le i_*$, the terms with $|p| > 2i_*$ are negligible by (6-49) and the trivial bounds (6-39) and (6-43). For $1 \le |p| \le 2i_*$ we use the improved bound (6-40). This gives

$$|u_i(t,\beta)| \le \max_{1\le |j|\le 2i_*} |u_j(0,\beta)| + N^{-\frac{3}{4}+\omega_1} = N^{-\frac{3}{4}+\omega_1}, \quad |i| \le i_*,$$

since $u(t=0,\beta)=0$. Integrating from $\beta=0$ to $\beta=1$, and recalling that $v(\beta=0)=\tilde z$ and $v(\beta=1)=\mathring{z}$, by a high-moment Markov inequality, we conclude

$$|\tilde z_i(t) - \mathring{z}_i(t)| \lesssim N^{-\frac{3}{4}+\omega_1}, \quad 1 \le |i| \le i_*,$$

with very high probability. This yields (6-38) and completes the proof of Lemma 6.4.

We remark that it would have been sufficient to require that $|\tilde z_j(0) - \mathring{z}_j(0)| \le N^{-\frac{3}{4}+\omega_1}$ for all $1 \le |j| \le 2i_*$ instead of setting $\mathring{z}(0) := \tilde z(0)$ initially. Later in Section 6D we will use a similar finite speed of propagation mechanism to show that changing the initial condition for large indices has negligible effect. □

6C3. *Refined rigidity for small $|i|$.* Finally, in the last but main step of the proof of (6-22) in Proposition 6.3 we compare $\mathring{z}_i$ with $\bar\gamma_i^*$ for small $|i|$ with a much higher precision than the crude bound $N^{-\frac{1}{2}+C\omega_1}$ which directly follows from (6-37) and (6-21). Notice that we use the semiquantiles for comparison since $\bar\gamma_i^* \in [\bar\gamma_{i-1}, \bar\gamma_i]$ and $\bar\gamma_i^*$ is typically close to the midpoint of this interval. In particular, $\bar\rho_t(\bar\gamma_i^*(t))$ is never zero, in fact we have $\bar\rho_t(\bar\gamma_i^*(t)) \ge cN^{-\frac{1}{3}}$, because by band rigidity quantiles may fall exactly at spectral edges, but semiquantiles cannot. This lower bound makes the semiquantiles much more convenient reference points than the quantiles.

Proposition 6.6. *Fix $\alpha \in [0,1]$; then with the notation above for the localised DBM $\mathring{z}(t,\alpha)$ on the short scale $L = N^{\frac{1}{2}+C_1\omega_1}$ with $10 \le C_1 \le \frac{1}{10}C_*$, defined in* (6-35), *we have*

$$|(\mathring{z}_i(t,\alpha) - \bar\gamma_i^*(t)) - (\mathring{z}_i(0,\alpha) - \bar\gamma_i^*(0))| \le N^{-\frac{3}{4}+C\omega_1}, \quad 1 \le |i| \le i_* = N^{\frac{1}{2}+C_*\omega_1}, \tag{6-50}$$

with very high probability.

Combining (6-50) with (6-38) and noticing that

$$\mathring{z}_i(0,\alpha) - \bar\gamma_i^*(0) = \tilde z_i(0,\alpha) - \bar\gamma_i^*(0) = O\left(\frac{N^{\xi} N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}\right)$$

for any $\xi > 0$ with very high probability by (6-23), we obtain (6-22) and complete the proof of Proposition 6.3. □

Proof of Proposition 6.6. We recall from (6-24) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar\gamma_i^*(t) = \alpha[-\Re m_{x,t}(\gamma_{x,i}^*(t)) + \Re m_{x,t}(\mathfrak{e}_{x,t}^+)] + (1-\alpha)[-\Re m_{y,t}(\gamma_{y,i}^*(t)) + \Re m_{y,t}(\mathfrak{e}_{y,t}^+)]. \tag{6-51}$$

Next, we define a dynamics that interpolates between $\mathring{z}_i(t,\alpha)$ and $\bar\gamma_i^*(t)$, i.e., between (6-35) and (6-51). Let $\beta \in [0,1]$ and for any fixed β define the process $v = v(t,\beta) = \{v_i(t,\beta)\}_{i=-N}^N$ as the solution of the interpolating DBM

$$\mathrm{d}v_i = \beta\sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j:|j-i|\le L}\frac{1}{v_i-v_j}\,\mathrm{d}t + \beta\Big[\frac{1}{N}\sum_{j:|j-i|>L}\frac{1}{\bar\gamma_i^*-\bar\gamma_j^*}\,\mathrm{d}t + \Phi(t)\Big]\mathrm{d}t$$
$$+(1-\beta)\Big[\frac{\mathrm{d}}{\mathrm{d}t}\bar\gamma_i^*(t) - \frac{1}{N}\sum_{j:|j-i|\le L}\frac{1}{\bar\gamma_i^*-\bar\gamma_j^*}\Big]\mathrm{d}t, \quad 1\le|i|\le N, \tag{6-52}$$

with initial condition $v_i(0,\beta) := \beta\mathring{z}_i(0) + (1-\beta)\bar\gamma_i^*(0)$. Notice that

$$v_i(t,\beta{=}0) = \bar\gamma_i^*(t), \quad v_i(t,\beta{=}1) = \mathring{z}_i(t). \tag{6-53}$$

Here we use the same letter v as in (6-41) within the proof of Lemma 6.4, but this is now a new interpolation. Since both appearances of the letter v are used only within the proofs of separate lemmas, this should not cause any confusion. The same remark applies to the letter u below.

Let $u := u(t,\beta) = \partial_\beta v(t,\beta)$; then it satisfies the equation

$$\mathrm{d}u_i = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \sum_{j\ne i}\mathcal{B}_{ij}(u_i-u_j)\,\mathrm{d}t + F_i\,\mathrm{d}t, \quad 1\le|i|\le N, \tag{6-54}$$

with a time-dependent short-range kernel (omitting the time argument and the β-parameter)

$$\mathcal{B}_{ij}(t) = \mathcal{B}_{ij} := -\frac{1}{N}\frac{\mathbf{1}(|i-j|\le L)}{(v_i-v_j)^2} \tag{6-55}$$

and external force

$$F_i = F_i(t) := -\frac{1}{N}\sum_j \frac{1}{\bar\gamma_j^*(t)-\bar\gamma_i^*(t)}$$
$$+\alpha\Re m_{x,t}(\gamma_{x,i}^*(t)) + (1-\alpha)\Re m_{y,t}(\gamma_{y,i}^*(t)) - h(t,\alpha), \quad 1\le|i|\le N. \tag{6-56}$$

Since the density $\bar\rho$ is regular, at least near the cusp regime, we can replace the sum over j with an integral with very high precision for small i; this integral is $\Re\bar m(\bar{\mathfrak{e}}^+ + \bar\gamma_i^*)$. A simple rearrangement of various terms yields

$$F_i = \Big[\Re\bar m(\bar{\mathfrak{e}}^+ + \bar\gamma_i^*) - \frac{1}{N}\sum_j\frac{1}{\bar\gamma_{j\ne i}^* - \bar\gamma_i^*}\Big] - (1-\alpha)D_{y,i} - \alpha D_{x,i} + O(N^{-1}), \tag{6-57}$$

with

$$D_{r,i} := \Re[(\bar m(\bar{\mathfrak{e}}^+ + \bar\gamma_i^*) - \bar m(\bar{\mathfrak{e}}^+)) - (m_r(\gamma_{r,i}^*) - m_r(\mathfrak{e}_r^+))], \quad r = x,y,$$

where we used the formula for h from (6-11) and the definition of Φ from (6-15). The choice of the shift h was governed by the idea to replace the last three terms in (6-56) by $\Re\bar m(\bar{\mathfrak{e}}^+ + \bar\gamma_i^*)$. However, the shift cannot be i-dependent as it would result in an i-dependent shift in the definition of $\tilde z_i$, see (6-13), which would mean that the differences (gaps) of the processes z_i and $\tilde z_i$ are not the same. Therefore, we defined the shift $h(t)$ by the similar formula evaluated at the edge, justifying the choice (6-11). The discrepancy is expressed by $D_{x,i}$ and $D_{y,i}$ which are small. Indeed we have, for $r = x, y$ and $1 \le |i| \le 2i_*$, that

$$\begin{aligned}|D_{r,i}| &\le \big|\Re[(\bar{m}(\mathfrak{e}^+ + \hat{\gamma}^*_{r,i}) - \bar{m}(\mathfrak{e}^+)) - (m_r(\mathfrak{e}^+_r + \hat{\gamma}^*_{r,i}) - m_r(\mathfrak{e}^+_r))]\big| + |\bar{m}(\mathfrak{e}^+ + \hat{\gamma}^*_{r,i}) - \bar{m}(\mathfrak{e}^+ + \bar{\gamma}^*_i)| \\ &\lesssim |\hat{\gamma}^*_{r,i}|^{\frac{1}{3}}[|\hat{\gamma}^*_{r,i}|^{\frac{1}{3}} + N^{-\frac{1}{6}+\frac{1}{3}\omega_1}]|\log|\hat{\gamma}^*_{r,i}|| + N^{-\frac{11}{36}+\omega_1} + \frac{|\hat{\gamma}^*_{r,i} - \bar{\gamma}^*_i|}{\bar{\rho}(\bar{\gamma}^*_i)^2} \\ &\lesssim \bigg[\Big(\frac{|i|}{N}\Big)^{\frac{1}{2}} + \Big(\frac{|i|}{N}\Big)^{\frac{1}{4}} N^{-\frac{1}{6}+\frac{1}{3}\omega_1}\bigg](\log N) + N^{-\frac{11}{36}+\omega_1} + \frac{(|i|/N) + (|i|/N)^{\frac{3}{4}} N^{-\frac{1}{6}+\omega_1}}{(|i|/N)^{\frac{1}{2}}} \\ &\lesssim N^{-\frac{1}{4}+C\omega_1},\end{aligned} \tag{6-58}$$

where from the first to the second line we used (4-22a) and the bound on the derivative of $\bar{m}$; see (4-7b). In the last inequality we used (4-13a) to estimate $|\hat{\gamma}^*_{r,i}| \lesssim (|i|/N)^{\frac{3}{4}} N^{C\omega_1}$ and similarly $|\hat{\gamma}^*_{r,i} - \bar{\gamma}^*_i|$ in the regime $|i| \le i_* = N^{\frac{1}{2}+C_*\omega_1}$; furthermore we used that $\bar{\rho}(\bar{\gamma}^*_i) \ge (|i|/N)^{\frac{1}{4}}$ and also $|\bar{\gamma}^*_i| \ge c/N$, since a semiquantile is always away from the edge.

Let $\mathcal{U}(s,t)$ be the fundamental solution of the heat evolution with kernel $\mathcal{B}$ from (6-55). Similarly to (6-32), the solution to the SDE (6-54) is given by

$$u(t) = \mathcal{U}(0,t)u + \sqrt{\frac{2}{N}} \int_0^t \mathcal{U}(s,t)\,\mathrm{d}B(s) + \int_0^t \mathcal{U}(s,t)F(s)\,\mathrm{d}s. \tag{6-59}$$

The middle martingale term can be estimated as in (6-33). The last term in (6-59) is estimated by

$$\left|\int_0^t \mathcal{U}(s,t)F(s)\,\mathrm{d}s\right| \le t \max_{0\le s\le t} \|F(s)\|_\infty. \tag{6-60}$$

First we use these simple Duhamel bounds to obtain a crude rigidity bound on $v_i(t,\beta)$ by integrating the bound on u

$$\begin{aligned}|v_i(t,\beta) - v_i(t,\beta=0)| &\le \beta \max_{\beta'\in[0,\beta]} |u_i(t,\beta')| \\ &\le \max_{\beta'\in[0,1]} \|u(0,\beta')\|_\infty + N^{-\frac{1}{2}+\omega_1+\xi}, \quad 1 \le |i| \le N,\end{aligned} \tag{6-61}$$

for any $\xi > 0$ with very high probability, using (6-33), (6-59), (6-60) and that $\mathcal{U}$ is a contraction. Note that in the first inequality of (6-61) we used that it holds with very high probability by a Markov inequality as in (6-45)–(6-46). We also used the trivial bound

$$\max_{0\le s\le t_*} \|F(s)\|_\infty \lesssim \log L \sim \log N, \tag{6-62}$$

which easily follows from (6-56), (6-58) and the fact that $|\bar{\gamma}^*_j(t) - \bar{\gamma}^*_i(t)| \gtrsim |i-j|/N$.

Recalling that $v_i(t,\beta=0) = \bar{\gamma}^*_i(t)$ and $u_i(0,\beta') = \mathring{z}_i(0) - \bar{\gamma}^*_i(0)$, together with (6-37) and (6-21), by (6-61), we obtain the crude rigidity

$$|v_i(t,\beta) - \bar{\gamma}^*_i(t)| \le N^{-\frac{1}{2}+2\omega_1}, \quad 1 \le |i| \le N, \tag{6-63}$$

with very high probability.

The main technical result is a considerable improvement of the bound (6-63) at least for i near the cusp regime. This is the content of the following proposition whose proof is postponed:

Proposition 6.7. *The vector F defined in* (6-56) *satisfies the bound*

$$\max_{s\le t_*}|F_i(s)|\le N^{-\frac{1}{4}+C\omega_1},\quad 1\le|i|\le 2i_*. \tag{6-64}$$

Since i_* is much bigger than $L=N^{\frac{1}{2}+C_1\omega_1}$ with a large C_1, and we have the rigidity (6-63) on a scale much smaller than L, similarly to Lemma 6.5, we have the following finite speed of propagation result. The proof is identical to that of Lemma 6.5.

Proposition 6.8. *For the short-range dynamics $\mathcal{U}=\mathcal{U}^{\mathcal{B}}$ defined by the operator* (6-55),

$$\sup_{0\le s\le t\le t_*}[\mathcal{U}_{pi}(s,t)+\mathcal{U}_{ip}(s,t)]\le N^{-D},\quad 1\le|i|\le i_*,\ |p|\ge 2i_*, \tag{6-65}$$

for any D if N is sufficiently large. □

Armed with these two propositions, we can easily complete the proof of Proposition 6.6. For any $1\le|i|\le i_*$ we have from (6-32), using (6-31), (6-33), (6-65) and that $\mathcal{U}$ is a contraction on ℓ^∞, that

$$\begin{aligned}|u_i(t)|&\le N^{-\frac{3}{4}+\omega_1+\xi}+\sum_p\mathcal{U}_{ip}|u_p(0)|+\int_0^t\sum_p\mathcal{U}_{ip}(s,t)|F_p(s)|\,\mathrm{d}s\\&\le N^{-\frac{3}{4}+\omega_1+\xi}+\max_{|p|\le 2i_*}|u_p(0)|+t\max_{0\le s\le t_*}\max_{|p|\le 2i_*}|F_p(s)|+N^{-D}\max_{0\le s\le t}\|F(s)\|_\infty.\end{aligned} \tag{6-66}$$

The trivial bound (6-62) together with (6-64) completes the proof of (6-50) by integrating back the bound (6-66) for $u=\partial_\beta v$ in β, using a high-moment Markov inequality similar to (6-45)–(6-46), and recalling (6-53). This completes the proof of Proposition 6.6. □

6C4. *Estimate of the forcing term.*

Proof of Proposition 6.7. Within this proof we will use $\gamma_i:=\bar\gamma_i(t)$, $\gamma_i^*:=\bar\gamma_i^*(t)$, $\rho=\bar\rho_t$, $m=\bar m_t$ and $\mathfrak{e}^+=\bar{\mathfrak{e}}_t^+$ for brevity. For notational simplicity we may assume within this proof that $\mathfrak{e}^+=0$ by a simple shift. The key input is the following bound on the derivative of the density, proven in [Alt et al. 2018a] for self-consistent densities of Wigner type matrices:

$$|\rho'(x)|\le\frac{C}{\rho(x)[\rho(x)+\Delta^{\frac{1}{3}}]},\quad |x|\le\delta_*, \tag{6-67}$$

where $\Delta=\bar\Delta_t$ is the length of the unique gap in the support of $\rho=\bar\rho_t$ in a small neighbourhood of size $\delta_*\sim 1$ around $\mathfrak{e}^+=0$. If there is no such gap, then we set $\Delta=0$ in (6-67). By the definition of the interpolated density $\bar\rho_t$ in (5-12) clearly follows that it satisfies (6-67) by Lemma 4.3. Notice that (6-67) implies local Hölder continuity; i.e.,

$$|\rho(x)-\rho(y)|\le\min\{|x-y|^{\frac{1}{3}},|x-y|^{\frac{1}{2}}\Delta^{-\frac{1}{6}}\} \tag{6-68}$$

for any x,y in a small neighbourhood of the gap or the local minimum.

Throughout the entire proof we fix an i with $1 \le |i| \le 2i_*$. For simplicity, we assume $i > 0$; the case $i < 0$ is analogous. We rewrite F_i from (6-57) as

$$F_i = G_1 + G_2 + G_3 + G_4, \tag{6-69}$$

with

$$G_1 := \sum_{1\le|j-i|\le L} \int_{\gamma_{j-1}}^{\gamma_j} \left[\frac{1}{x-\gamma_i^*} - \frac{1}{\gamma_j^*-\gamma_i^*}\right]\rho(x)\,\mathrm{d}x, \quad G_2 := \int_{\gamma_{i-1}}^{\gamma_i} \frac{\rho(x)\,\mathrm{d}x}{x-\gamma_i^*},$$
$$G_3 := \sum_{|j-i|>L} \int_{\gamma_{j-1}}^{\gamma_j} \left[\frac{1}{x-\gamma_i^*} - \frac{1}{\gamma_j^*-\gamma_i^*}\right]\rho(x)\,\mathrm{d}x, \qquad G_4 := -(1-\alpha)D_{y,i} - \alpha D_{x,i} + O(N^{-1}).$$

The term G_4 was already estimated in (6-58). In the following we will show separately that $|G_a| \lesssim N^{-\frac{1}{4}}$, $a = 1, 2, 3$.

Estimate of G_3. By elementary computations, using the crude rigidity (6-21), it follows that

$$|G_3| \lesssim \frac{N^{-\frac{1}{2}+2\omega_1}}{N} \sum_{j:|j-i|>L} \frac{1}{(\gamma_i^*-\gamma_j^*)^2}.$$

Then, the estimate $|G_3| \lesssim N^{-\frac{1}{4}}$ follows using the same computations as in (6-40).

Estimate of G_2. We write

$$G_2 = \int_{\gamma_{i-1}}^{\gamma_i} \frac{\rho(x)\,\mathrm{d}x}{x-\gamma_i^*} = \int_{\gamma_{i-1}}^{\gamma_i} \frac{\rho(x)-\rho(\gamma_i^*)}{x-\gamma_i^*}\,\mathrm{d}x + \rho(\gamma_i^*) \int_{\gamma_{i-1}}^{\gamma_i} \frac{\mathrm{d}x}{x-\gamma_i^*} \tag{6-70}$$

and we will show that both summands are bounded by $CN^{-\frac{1}{4}}$. We make the convention that if γ_{i-1} is exactly at the left edge of a gap, then for the purpose of this proof we redefine it to be the right edge of the same gap and similarly, if γ_i is exactly at the right edge of the gap, then we set it to be the left edge. This is just to make sure that $[\gamma_{i-1}, \gamma_i]$ is always included in the support of ρ.

In the first integral we use (6-68) to get

$$\left|\int_{\gamma_{i-1}}^{\gamma_i} \frac{\rho(x)-\rho(\gamma_i^*)}{x-\gamma_i^*}\,\mathrm{d}x\right| \lesssim \min\{(\gamma_i-\gamma_{i-1})^{\frac{1}{3}}, (\gamma_i-\gamma_{i-1})^{\frac{1}{2}}\Delta^{-\frac{1}{6}}\} = O(N^{-\frac{1}{4}}). \tag{6-71}$$

Here we used that the local eigenvalue spacing (with the convention above) is bounded by

$$\gamma_i - \gamma_{i-1} \lesssim \max\left\{\frac{\Delta^{\frac{1}{9}}}{N^{\frac{2}{3}}}, \frac{1}{N^{\frac{3}{4}}}\right\}. \tag{6-72}$$

For the second integral in (6-70) is an explicit calculation:

$$\rho(\gamma_i^*) \int_{\gamma_{i-1}}^{\gamma_i} \frac{\mathrm{d}x}{x-\gamma_i^*} = \rho(\gamma_i^*) \log \frac{\gamma_i-\gamma_i^*}{\gamma_i^*-\gamma_{i-1}}. \tag{6-73}$$

Using the definition of the quantiles and (6-68), we have

$$\frac{1}{2N} = \int_{\gamma_{i-1}}^{\gamma_i^*} \rho(x)\,\mathrm{d}x = \rho(\gamma_i^*)(\gamma_i^*-\gamma_{i-1}) + O\big(\min\{|\gamma_i^*-\gamma_{i-1}|^{\frac{4}{3}}, |\gamma_i^*-\gamma_{i-1}|^{\frac{3}{2}}\Delta^{-\frac{1}{6}}\}\big),$$

and similarly

$$\frac{1}{2N} = \int_{\gamma_i^*}^{\gamma_i} \rho(x)\,\mathrm{d}x = \rho(\gamma_i^*)(\gamma_i - \gamma_i^*) + O\big(\min\{|\gamma_i^* - \gamma_i|^{\frac{4}{3}}, |\gamma_i^* - \gamma_i|^{\frac{3}{2}}\Delta^{-\frac{1}{6}}\}\big).$$

The error terms are comparable and they are $O(N^{-1})$ using (6-72); thus, subtracting these two equations, we have

$$|(\gamma_i - \gamma_i^*) - (\gamma_i^* - \gamma_{i-1})| \lesssim \frac{\min\{|\gamma_i^* - \gamma_i|^{\frac{4}{3}}, |\gamma_i^* - \gamma_i|^{\frac{3}{2}}\Delta^{-\frac{1}{6}}\}}{\rho(\gamma_i^*)}.$$

Expanding the logarithm in (6-73), we have

$$\left|\rho(\gamma_i^*)\int_{\gamma_{i-1}}^{\gamma_i} \frac{\mathrm{d}x}{x-\gamma_i^*}\right| \lesssim \rho(\gamma_i^*)\frac{|(\gamma_i - \gamma_i^*) - (\gamma_i^* - \gamma_{i-1})|}{\gamma_i^* - \gamma_{i-1}} \lesssim \min\{|\gamma_i^* - \gamma_i|^{\frac{1}{3}}, |\gamma_i^* - \gamma_i|^{\frac{1}{2}}\Delta^{-\frac{1}{6}}\} \lesssim N^{-\frac{1}{4}}$$

as in (6-71). This completes the estimate

$$|G_2| \lesssim N^{-\frac{1}{4}}. \tag{6-74}$$

Estimate of G_1. Fix $i > 0$ and set $n = n(i)$ as

$$n(i) := \min\big\{n \in \mathbb{N} \,:\, \min\{|\gamma_{i-n-1} - \gamma_i^*|, |\gamma_{i+n} - \gamma_i^*|\} \geq cN^{-\frac{3}{4}}\big\}, \tag{6-75}$$

with some small constant $c > 0$.

Next, we estimate $n(i)$. Notice that for $i = 1$ we have $n(i) = 0$. If $i \geq 2$, then we notice that one can choose c sufficiently small, depending only on the model parameters, such that

$$\tfrac{1}{2} \leq \frac{\rho(x)}{\rho(\gamma_i^*)} \leq 2 \quad \text{for all } x \in [\gamma_{i-n(i)-1}, \gamma_{i+n(i)}],\ i \geq 2. \tag{6-76}$$

Let

$$m(i) := \max\left\{m \in \mathbb{N} : \tfrac{1}{2} \leq \frac{\rho(x)}{\rho(\gamma_i^*)} \leq 2 \text{ for all } x \in [\gamma_{i-m-1}, \gamma_{i+m}]\right\};$$

then, in order to verify (6-76), we need to prove that $m(i) \geq n(i)$.

Then by a case-by-case calculation it follows that

$$m(i) \geq c_1|i|, \tag{6-77}$$

and thus

$$\min\{|\gamma_{i-m(i)-1} - \gamma_i^*|, |\gamma_{i+m(i)} - \gamma_i^*|\} \gtrsim \max\left\{\left(\frac{i}{N}\right)^{\frac{2}{3}}\Delta^{\frac{1}{9}}, \left(\frac{i}{N}\right)^{\frac{3}{4}}\right\} \geq c_2N^{-\frac{3}{4}} \tag{6-78}$$

with some c_1, c_2. Hence (6-76) will hold if $c \leq c_2$ is chosen in the definition (6-75). Notice that in these estimates it is important that the semiquantiles are always at a certain distance away from the quantiles.

Now we give an upper bound on $n(i)$ when γ_i^* is near a (possible small) gap as in the proof above. The local eigenvalue spacing is

$$\gamma_i - \gamma_i^* \sim \max\left\{\frac{\Delta^{\frac{1}{9}}}{N^{\frac{2}{3}}(i)^{\frac{1}{3}}}, \frac{1}{N^{\frac{3}{4}}(i)^{\frac{1}{4}}}\right\}, \tag{6-79}$$

which is bigger than $cN^{-\frac{3}{4}}$ if $i \le \Delta^{\frac{1}{3}} N^{\frac{1}{4}}$. So in this case $n(i) = 0$ and we may now assume that $i \ge \Delta^{\frac{1}{3}} N^{\frac{1}{4}}$ and still $i \ge 2$.

Consider first the so-called *cusp case* when $i \ge N\Delta^{\frac{4}{3}}$; in this case, as long as $n \le \frac{1}{2}i$, we have

$$\gamma_{i+n} - \gamma_i^* \sim \frac{n}{N^{\frac{3}{4}}(i+1)^{\frac{1}{4}}}.$$

This is bigger than $cN^{-\frac{3}{4}}$ if $n \ge i^{\frac{1}{4}}$; thus we have $n(i) \le i^{\frac{1}{4}}$ in this case.

In the opposite case, the so-called *edge case*, $i \le N\Delta^{\frac{4}{3}}$, which together with the above assumption $i \ge \Delta^{\frac{1}{3}} N^{\frac{1}{4}}$ also implies that $\Delta \ge N^{-\frac{3}{4}}$. In this case, as long as $n \le \frac{1}{2}i$, we have

$$\gamma_{i+n} - \gamma_i^* \sim \frac{n\Delta^{\frac{1}{9}}}{N^{\frac{2}{3}} i^{\frac{1}{3}}}.$$

This is bigger than $cN^{-\frac{3}{4}}$ if $n \ge \Delta^{-\frac{1}{9}} N^{-\frac{1}{12}} i^{\frac{1}{3}}$. So we have $n(i) \le \Delta^{-\frac{1}{9}} N^{-\frac{1}{12}} i^{\frac{1}{3}} \le i^{\frac{1}{3}}$ in this case.

We split the sum in the definition of G_1, see (6-69), as follows:

$$G_1 = \sum_{1\le|j-i|\le L} \int_{\gamma_{j-1}}^{\gamma_j} \frac{x-\gamma_j^*}{(\gamma_i^*-\gamma_j^*)(x-\gamma_i^*)} \rho(x)\,\mathrm{d}x = \Bigg(\sum_{n(i)<|j-i|\le L} + \sum_{1\le|j-i|\le n(i)} \Bigg) =: S_1 + S_2. \quad (6\text{-}80)$$

For the first sum we use $|x-\gamma_j^*| \le \gamma_{j+1}^* - \gamma_j^*$, $|\gamma_i^* - x| \sim |\gamma_i^* - \gamma_j^*|$. Moreover, we have

$$\rho(\gamma_i^*)(\gamma_i - \gamma_{i-1}) \sim \frac{1}{N} \quad (6\text{-}81)$$

from the definition of the semiquantiles. Thus we restore the integration in the first sum S_1 and estimate

$$|S_1| \lesssim \frac{1}{N} \Bigg[\int_{-\infty}^{\gamma_{i-n(i)-1}} + \int_{\gamma_{i+n(i)}}^{\infty} \Bigg] \frac{\mathrm{d}x}{|x-\gamma_i^*|^2} \lesssim \frac{1}{N} \Bigg[\frac{1}{|\gamma_{i-n(i)-1} - \gamma_i^*|} + \frac{1}{|\gamma_{i+n(i)} - \gamma_i^*|} \Bigg] \le CN^{-\frac{1}{4}}. \quad (6\text{-}82)$$

In the last step we used the definition of $n(i)$.

Now we consider S_2. Notice that this sum is nonempty only if $n(i) \ne 0$ In this case to estimate S_2 we have to symmetrise. Fix $1 \le n \le n(i)$, assume $i > n$ and consider together

$$\begin{aligned}
&\int_{\gamma_{i-n-1}}^{\gamma_{i-n}} \frac{x-\gamma_{i-n}^*}{(\gamma_i^*-\gamma_{i-n}^*)(x-\gamma_i^*)}\rho(x)\,\mathrm{d}x + \int_{\gamma_{i+n-1}}^{\gamma_{i+n}} \frac{x-\gamma_{i+n}^*}{(\gamma_i^*-\gamma_{i+n}^*)(x-\gamma_i^*)}\rho(x)\,\mathrm{d}x \\
&= \frac{1}{\gamma_i^*-\gamma_{i-n}^*} \int_{\gamma_{i-n-1}}^{\gamma_{i-n}} \frac{x-\gamma_{i-n}^*}{x-\gamma_i^*}\rho(x)\,\mathrm{d}x + \frac{1}{\gamma_i^*-\gamma_{i+n}^*} \int_{\gamma_{i+n-1}}^{\gamma_{i+n}} \frac{x-\gamma_{i+n}^*}{x-\gamma_i^*}\rho(x)\,\mathrm{d}x \\
&= \frac{1}{N}\Bigg[\frac{1}{\gamma_i^*-\gamma_{i-n}^*} + \frac{1}{\gamma_i^*-\gamma_{i+n}^*}\Bigg] + \Bigg[\int_{\gamma_{i-n-1}}^{\gamma_{i-n}} \frac{\rho(x)\,\mathrm{d}y}{x-\gamma_i^*} + \int_{\gamma_{i+n-1}}^{\gamma_{i+n}} \frac{\rho(x)\,\mathrm{d}x}{x-\gamma_i^*}\Bigg] =: B_1(n) + B_2(n). \quad (6\text{-}83)
\end{aligned}$$

We now use $\frac{1}{3}$-Hölder regularity,

$$\rho(x) = \rho(\gamma_i^*) + O(|x-\gamma_i^*|^{\frac{1}{3}}).$$

We thus have

$$\sum_{n\le n(i)}\int_{\gamma_{i-n-1}}^{\gamma_{i-n}}\frac{\rho(x)\,\mathrm{d}y}{x-\gamma_i^*}=\sum_{n\le n(i)}\rho(\gamma_i^*)\log\frac{\gamma_{i-n-1}-\gamma_i^*}{\gamma_{i-n}-\gamma_i^*}+O\biggl(\int_{\gamma_{i-n(i)-1}}^{\gamma_{i+n(i)}}\frac{\mathrm{d}x}{|x-\gamma_i^*|^{\frac{2}{3}}}\biggr) \tag{6-84}$$

and similarly

$$\sum_{n\le n(i)}\int_{\gamma_{i+n-1}}^{\gamma_{i+n}}\frac{\rho(x)\,\mathrm{d}y}{x-\gamma_i^*}=\sum_{n\le n(i)}\rho(\gamma_i^*)\log\frac{\gamma_{i+n-1}-\gamma_i^*}{\gamma_{i+n}-\gamma_i^*}+O\biggl(\int_{\gamma_{i-n(i)-1}}^{\gamma_{i+n(i)}}\frac{\mathrm{d}x}{|x-\gamma_i^*|^{\frac{2}{3}}}\biggr). \tag{6-85}$$

The error terms are bounded by $CN^{-\frac{1}{4}}$ using (6-75) and therefore we have

$$\begin{aligned}\sum_{n\le n(i)}B_2(n)&=\sum_{n\le n(i)}\rho(\gamma_i^*)\biggl[\log\frac{\gamma_i^*-\gamma_{i-n-1}}{\gamma_i^*-\gamma_{i-n}}-\log\frac{\gamma_{i+n}-\gamma_i^*}{\gamma_{i+n-1}-\gamma_i^*}\biggr]+O(N^{-\frac{1}{4}})\\&=\sum_{n\le n(i)}\rho(\gamma_i^*)\biggl[\log\frac{\gamma_i^*-\gamma_{i-n-1}}{\gamma_{i+n}-\gamma_i^*}+\log\frac{\gamma_{i+n-1}-\gamma_i^*}{\gamma_i^*-\gamma_{i-n}}\biggr]+O(N^{-\frac{1}{4}}).\end{aligned}$$

We now use the bound

$$|\rho(x)-\rho(\gamma_i^*)|\lesssim\frac{|x-\gamma_i^*|}{\rho(\gamma_i^*)^2+\rho(\gamma_i^*)\Delta^{\frac{1}{3}}},\quad x\in[\gamma_{i-n(i)-1},\gamma_{i+n(i)}], \tag{6-86}$$

which follows from the derivative bound (6-67) if ϵ in the definition of $i_*=\epsilon N$ is chosen sufficiently small, depending on δ since throughout the proof $1\le|i|\le 2i_*$ and $n(i)\ll i_*$.

Note that

$$\frac{n}{N}=\int_{\gamma_{i-n}}^{\gamma_i}\rho(x)\,\mathrm{d}x=\rho(\gamma_i^*)[\gamma_i-\gamma_{i-n}]+O\biggl(\frac{|\gamma_{i-n}-\gamma_i^*|^2}{\rho(\gamma_i^*)^2+\rho(\gamma_i^*)\Delta^{\frac{1}{3}}}\biggr). \tag{6-87}$$

Thus, using (6-87) also for $\gamma_{i+n}-\gamma_i$, equating the two equations and dividing by $\rho(\gamma_i^*)$, we have

$$\gamma_i-\gamma_{i-n}=\gamma_{i+n}-\gamma_i+O\biggl(\frac{|\gamma_{i-n}-\gamma_i^*|^2}{\rho(\gamma_i^*)^3+\rho(\gamma_i^*)^2\Delta^{\frac{1}{3}}}\biggr). \tag{6-88}$$

A similar relation holds for the semiquantiles,

$$\gamma_i^*-\gamma_{i-n}^*=\gamma_{i+n}^*-\gamma_i^*+O\biggl(\frac{|\gamma_{i-n}^*-\gamma_i^*|^2}{\rho(\gamma_i^*)^3+\rho(\gamma_i^*)^2\Delta^{\frac{1}{3}}}\biggr), \tag{6-89}$$

and for the mixed relations among quantiles and semiquantiles,

$$\begin{aligned}\gamma_i^*-\gamma_{i-n}&=\gamma_{i+n-1}-\gamma_i^*+O\biggl(\frac{|\gamma_{i-n}-\gamma_i^*|^2}{\rho(\gamma_i^*)^3+\rho(\gamma_i^*)^2\Delta^{\frac{1}{3}}}\biggr),\\\gamma_i^*-\gamma_{i-n-1}&=\gamma_{i+n}-\gamma_i^*+O\biggl(\frac{|\gamma_{i-n}-\gamma_i^*|^2}{\rho(\gamma_i^*)^3+\rho(\gamma_i^*)^2\Delta^{\frac{1}{3}}}\biggr).\end{aligned}$$

Thus, using $\gamma_i^* - \gamma_{i-n-1} \sim \gamma_{i+n} - \gamma_i^*$, we have

$$\rho(\gamma_i^*)\left|\log \frac{\gamma_i^* - \gamma_{i-n-1}}{\gamma_{i+n} - \gamma_i^*}\right| \lesssim \frac{\rho(\gamma_i^*)}{\gamma_{i+n} - \gamma_i^*} O\left(\frac{|\gamma_{i-n-1} - \gamma_i^*|^2}{\rho(\gamma_i^*)^3 + \rho(\gamma_i^*)^2 \Delta^{\frac{1}{3}}}\right) \lesssim \frac{|\gamma_{i-n-1} - \gamma_i^*|}{\rho(\gamma_i^*)^2 + \rho(\gamma_i^*)\Delta^{\frac{1}{3}}}. \tag{6-90}$$

Using $n \le n(i)$ and (6-75), we have $|\gamma_{i-n-1} - \gamma_i^*| \lesssim N^{-\frac{3}{4}}$. The contribution of this term to $\sum_n B_2(n)$ is thus

$$N^{-\frac{3}{4}} \sum_{n \le n(i)} \frac{1}{\rho(\gamma_i^*)^2 + \rho(\gamma_i^*)\Delta^{\frac{1}{3}}} \le \frac{n(i) N^{-\frac{3}{4}}}{\rho(\gamma_i^*)^2 + \rho(\gamma_i^*)\Delta^{\frac{1}{3}}}. \tag{6-91}$$

In the bulk regime we have $\rho(\gamma_i^*) \sim 1$ and $n(i) \sim N^{\frac{1}{4}}$, so this contribution is much smaller than $N^{-\frac{1}{4}}$.

In the cusp regime, i.e., when $\Delta \le (i/N)^{\frac{3}{4}}$, then we have $\gamma_i^* \sim (i/N)^{\frac{3}{4}}$ and $\rho(\gamma_i^*) \sim (i/N)^{\frac{1}{4}}$, thus we get

$$\text{(6-91)} \le \frac{n(i) N^{-\frac{3}{4}}}{\rho(\gamma_i^*)^2 + \rho(\gamma_i^*)\Delta^{\frac{1}{3}}} \le \frac{n(i) N^{-\frac{3}{4}}}{\rho(\gamma_i^*)^2} \lesssim N^{-\frac{1}{4}} n(i) i^{-\frac{1}{2}} \lesssim N^{-\frac{1}{4}}$$

since $n(i) \le i^{\frac{1}{4}}$.

In the edge regime, i.e., when $\Delta \ge (i/N)^{\frac{3}{4}}$, we have $\gamma_i^* \sim \Delta^{\frac{1}{9}} (i/N)^{\frac{2}{3}}$ and $\rho(\gamma_i^*) \sim \Delta^{-\frac{1}{9}} (i/N)^{\frac{1}{3}}$; thus we get

$$\text{(6-91)} \le \frac{n(i) N^{-\frac{3}{4}}}{\rho(\gamma_i^*)^2 + \rho(\gamma_i^*)\Delta^{\frac{1}{3}}} \le \frac{n(i) N^{-\frac{3}{4}}}{\rho(\gamma_i^*)\Delta^{\frac{1}{3}}} \lesssim \frac{n(i) N^{-\frac{5}{12}}}{\Delta^{\frac{2}{9}} i^{\frac{1}{3}}} \le \frac{N^{-\frac{5}{12}}}{\Delta^{\frac{2}{9}}} \le N^{-\frac{1}{4}}$$

since $n(i) \le i^{\frac{1}{3}}$ and $\Delta \ge N^{-\frac{3}{4}}$. This completes the proof of $\sum_n B_2(n) \lesssim N^{-\frac{1}{4}}$.

Finally the $\sum_n B_1(n)$ term from (6-83) is estimated as follows by using (6-89):

$$\begin{aligned}\sum_n \frac{1}{N}\left[\frac{1}{\gamma_i^* - \gamma_{i-n-1}^*} + \frac{1}{\gamma_i^* - \gamma_{i+n-1}^*}\right] &= \sum_n \frac{1}{N} \frac{1}{(\gamma_i^* - \gamma_{i-n}^*)^2} O\left(\frac{(\gamma_i - \gamma_{i-n-1})^2}{\rho(\gamma_i^*)^2 [\rho(\gamma_i^*) + \Delta^{\frac{1}{3}}]}\right) \\ &\lesssim \frac{n(i)}{N\rho(\gamma_i^*)^2 [\rho(\gamma_i^*) + \Delta^{\frac{1}{3}}]}.\end{aligned} \tag{6-92}$$

In the bulk regime this is trivially bounded by $CN^{-\frac{3}{4}}$. In the cusp regime, $\Delta \le (i/N)^{\frac{3}{4}}$, and we have

$$\frac{n(i)}{N\rho(\gamma_i^*)^2 [\rho(\gamma_i^*) + \Delta^{\frac{1}{3}}]} \le \frac{n(i)}{N\rho(\gamma_i^*)^3} \lesssim \frac{n(i)}{N^{\frac{1}{4}} i^{\frac{3}{4}}} \lesssim N^{-\frac{1}{4}}$$

since $n(i) \le i^{\frac{1}{4}}$.

Finally, in the edge regime, $\Delta \ge (i/N)^{\frac{3}{4}}$, we just use

$$\frac{n(i)}{N\rho(\gamma_i^*)^2 [\rho(\gamma_i^*) + \Delta^{\frac{1}{3}}]} \le \frac{n(i)}{N\rho(\gamma_i^*)^2 \Delta^{\frac{1}{3}}} \lesssim \frac{n(i)}{N^{\frac{1}{4}} i^{\frac{3}{4}}} \lesssim N^{-\frac{1}{4}}$$

since $n(i) \le i^{\frac{1}{3}}$. This gives $\sum_n B_1(n) \lesssim N^{-\frac{1}{4}}$. Together with the estimate on $\sum_n B_2(n)$ we get $|S_2| \lesssim N^{-\frac{1}{4}}$; see (6-80) and (6-83). This completes the estimate of G_1 in (6-69), which, together with (6-74) and (6-58), finishes the proof of Proposition 6.7. □

6D. ***Phase 2: rigidity of $\hat z$ on scale $N^{-\frac{3}{4}+\frac{1}{6}\omega_1}$, without i dependence.*** For any fixed $\alpha \in [0,1]$ recall the definition of the shifted process $\tilde z(t,\alpha)$ (6-14) and the shifted α-interpolating semiquantiles $\bar\gamma_i^*(t)$ from (5-10) that trail $\tilde z$. Furthermore, for all $0 \le t \le t_*$ we consider the interpolated density $\bar\rho_t$ with a small gap $[\bar{\mathfrak{e}}_t^-, \bar{\mathfrak{e}}_t^+]$, and its Stieltjes transform $\bar m_t$. In particular,

$$\bar{\mathfrak{e}}_t^\pm = \alpha \mathfrak{e}_{x,t}^\pm + (1-\alpha)\mathfrak{e}_{y,t}^\pm.$$

We recall that by Proposition 6.3 and (6-23) we have

$$\sup_{0\le t\le t_*} \max_{1\le |i|\le i_*} |\tilde z_i(t,\alpha) - \bar\gamma_i^*(t)| \le N^{-\frac{3}{4}+C\omega_1} \tag{6-93}$$

holds with very high probability for some $i_* = N^{\frac{1}{2}+C_*\omega_1}$.

In this section we improve the rigidity (6-93) from scale $N^{-\frac{3}{4}+C\omega_1}$ to the almost-optimal but still i-independent rigidity of order $N^{-\frac{3}{4}+\frac{1}{6}\omega_1+\xi}$ but only for a new short-range approximation $\hat z_i(t,\alpha)$ of $\tilde z_i(t,\alpha)$. The range of this new approximation $\ell^4 = N^{4\omega_\ell}$ with some $\omega_\ell \ll 1$ is much shorter than that of $\mathring{z}_i(t,\alpha)$ in Section 6C. Furthermore, the result will hold only for $1 \le |i| \le N^{4\omega_\ell+\delta_1}$ for some small $\delta_1 > 0$. The rigorous statement is in Proposition 6.10 below, after we give the definition of the short-range approximation.

Short-range approximation on fine scale. Adapting the idea of [Landon and Yau 2017] to the cusp regime, we now introduce a new short-range approximation process $\hat z(t,\alpha)$ for the solution to (6-14). The short-range approximation in this section will always be denoted by hat, $\hat z$, to distinguish it from the other short-range approximation, $\mathring{z}$, used in Section 6C; see (6-35). Not only is the length scale shorter for $\hat z$, but the definition of $\hat z$ is more subtle than in (6-35)

The new short-scale approximation is characterised by two exponents ω_ℓ and ω_A. In particular, we will always assume that $\omega_1 \ll \omega_\ell \ll \omega_A \ll 1$, where recall that $t_* \sim N^{-\frac{1}{2}+\omega_1}$ is defined in such a way that $\bar\rho_{t_*}$ has an exact cusp. The key quantity is $\ell := N^{\omega_\ell}$, which determines the scale on which the interaction term in (6-14) will be cut off and replaced by its mean-field value. This scale is not constant; it increases away from the cusp at a certain rate. The cutoff will be effective only near the cusp; for indices beyond $\frac{1}{2}i_*$, with $i_* = N^{\frac{1}{2}+C_*\omega_1}$, no cutoff is made. Finally, the intermediate scale N^{ω_A} is used for a technical reason: closer to the cusp, for indices less than N^{ω_A}, we always use the density $\rho_{y,t}$ of the reference process $y(t)$ to define the mean field approximation of the cutoff long-range terms. Beyond this scale we use the actual density $\bar\rho_t$. In this way we can exploit the closeness of the density $\bar\rho_t$ to the reference density $\rho_{y,t}$ near the cusp and simplify the estimate. This choice will guarantee that the error term ζ_0 in (6-105) below is nonzero only for $|i| > N^{\omega_A}$.

Now we define the $\hat z$-process precisely. Let

$$\mathcal{A} := \{(i,j) : |i-j| \le \ell(10\ell^3 + |i|^{\frac{3}{4}} + |j|^{\frac{3}{4}})\} \cup \{(i,j) : |i|, |j| > \tfrac{1}{2}i_*\}. \tag{6-94}$$

One can easily check that for each i with $1 \le |i| \le \frac{1}{2}i_*$ the set $\{j : (i,j) \in \mathcal{A}\}$ is an interval of the nonzero integers and that $(i,j) \in \mathcal{A}$ if and only if $(j,i) \in \mathcal{A}$. For each such fixed i we denote the smallest and the biggest j such that $(i, j) \in \mathcal{A}$ by $j_-(i)$ and $j_+(i)$, respectively. We will use the notation

$$\sum_j^{\mathcal{A},(i)} := \sum_{\substack{j:(i,j)\in\mathcal{A}\\ i\neq j}}, \qquad \sum_j^{\mathcal{A}^c,(i)} := \sum_{j:(i,j)\notin\mathcal{A}}.$$

Assuming for simplicity that i_* is divisible by 4, we introduce the intervals

$$\mathcal{J}_z(t) := [\bar{\gamma}_{-\frac{3}{4}i_*}(t), \bar{\gamma}_{\frac{3}{4}i_*}(t)], \tag{6-95}$$

and for each $0 < |i| \le \frac{1}{2}i_*$ we define

$$\mathcal{I}_{z,i}(t) := [\bar{\gamma}_{j_-(i)}(t), \bar{\gamma}_{j_+(i)}(t)]. \tag{6-96}$$

For a fixed $\alpha \in [0, 1]$ and $N \ge |i| > \frac{1}{2}i_*$ we let

$$\mathrm{d}\hat{z}_i(t,\alpha) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \left[\frac{1}{N}\sum_j^{\mathcal{A},(i)} \frac{1}{\hat{z}_i(t,\alpha)-\hat{z}_j(t,\alpha)} + \frac{1}{N}\sum_j^{\mathcal{A}^c,(i)} \frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)} + \Phi_\alpha(t)\right]\mathrm{d}t, \tag{6-97}$$

for $0 < |i| \le N^{\omega_A}$

$$\mathrm{d}\hat{z}_i(t,\alpha) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \left[\frac{1}{N}\sum_j^{\mathcal{A},(i)} \frac{1}{\hat{z}_i(t,\alpha)-\hat{z}_j(t,\alpha)} + \int_{\mathcal{I}_{y,i}(t)^c} \frac{\rho_{y,t}(E+\mathfrak{e}_{y,t}^+)}{\hat{z}_i(t,\alpha)-E}\,\mathrm{d}E + \Re[m_{y,t}(\mathfrak{e}_{y,t}^+)]\right]\mathrm{d}t, \tag{6-98}$$

and for $N^{\omega_A} < |i| \le \frac{1}{2}i_*$

$$\begin{aligned}\mathrm{d}\hat{z}_i(t,\alpha) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \Bigg[\frac{1}{N}\sum_j^{\mathcal{A},(i)} \frac{1}{\hat{z}_i(t,\alpha)-\hat{z}_j(t)} + \int_{\mathcal{I}_{z,i}(t)^c\cap\mathcal{J}_z(t)} \frac{\bar{\rho}_t(E+\bar{\mathfrak{e}}_t^+)}{\hat{z}_i(t,\alpha)-E}\,\mathrm{d}E \\ + \frac{1}{N}\sum_{|j|\ge\frac{3}{4}i_*} \frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)} + \Phi_\alpha(t)\Bigg]\mathrm{d}t,\end{aligned} \tag{6-99}$$

with initial data

$$\hat{z}_i(0,\alpha) := \tilde{z}_i(0,\alpha), \tag{6-100}$$

where we recall that $\tilde{z}_i(0,\alpha) = \alpha\tilde{x}_i(0) + (1-\alpha)\tilde{y}_i(0)$ for any $\alpha \in [0, 1]$. In particular, $\hat{z}(t, 1) = \hat{x}(t)$ and $\hat{z}(t, 0) = \hat{y}(t)$, which are the short-range approximations of the $\tilde{x}(t) := x(t) - \mathfrak{e}_{x,t}^+$ and $\tilde{y}(t) := x(t) - \mathfrak{e}_{y,t}^+$ processes.

Using the rigidity estimates in (6-21) and (6-93) we will prove the following lemma in Appendix C.

Lemma 6.9. *Assuming that the rigidity estimates* (6-21) *and* (6-93) *hold, for any fixed* $\alpha \in [0, 1]$ *we have*

$$\sup_{1\le|i|\le N}\ \sup_{0\le t\le t_*} |\hat{z}_i(t,\alpha) - \tilde{z}_i(t,\alpha)| \le \frac{N^{C\omega_1}}{N^{\frac{3}{4}}} \tag{6-101}$$

with very high probability.

In particular, since (6-21) and (6-93) have already been proven, we conclude from (6-93) and (6-101) that

$$\sup_{0\le t\le t_*} |\hat{z}_i(t,\alpha)-\bar{\gamma}_i(t)| \le \frac{N^{C\omega_1}}{N^{\frac{3}{4}}}, \quad 1\le |i|\le i_*, \tag{6-102}$$

for any fixed $\alpha\in[0,1]$.

Now we state the improved rigidity for $\hat{z}$, the main result of Section 6D:

Proposition 6.10. *Fix any $\alpha\in[0,1]$. There exists a constant $C>0$ such that if $0<\delta_1<C\omega_\ell$ then*

$$\sup_{0\le t\le t_*} |\hat{z}_i(t,\alpha)-\bar{\gamma}_i(t)| \lesssim \frac{N^{\xi}N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}}, \quad 1\le |i|\le N^{4\omega_\ell+\delta_1}, \tag{6-103}$$

for any $\xi>0$ with very high probability.

Proof. Recall that initially $\tilde{z}_i(0,\alpha)$ is a linear interpolation between $\tilde{x}_i(0)$ and $\tilde{y}_i(0)$ and thus for $\tilde{z}_i(0,\alpha)$ optimal rigidity (6-23) holds. We define the derivative process

$$w_i(t,\alpha) := \partial_\alpha \hat{z}_i(t,\alpha). \tag{6-104}$$

In particular, we find that $w=w(t,\alpha)$ is the solution of

$$\partial_t w = \mathcal{L}w + \zeta^{(0)}, \quad \mathcal{L} := \mathcal{B}+\mathcal{V}, \tag{6-105}$$

with initial data

$$w_i(0,\alpha) = \hat{x}_i(0)-\hat{y}_i(0).$$

Here, for any $1\le |i|\le N$, the (short-range) operator $\mathcal{B}$ is defined on any vector $f\in\mathbb{C}^{2N}$ as

$$(\mathcal{B}f)_i := \sum_j^{\mathcal{A},(i)} \mathcal{B}_{ij}(f_i-f_j), \quad \mathcal{B}_{ij} := -\frac{1}{N}\frac{1}{(\hat{z}_i(t,\alpha)-\hat{z}_j(t,\alpha))^2}. \tag{6-106}$$

Moreover, $\mathcal{V}$ is a multiplication operator, i.e., $(\mathcal{V}f)_i=\mathcal{V}_i f_i$, where $\mathcal{V}_i$ is defined in different regimes of i as follows:

$$\begin{aligned} \mathcal{V}_i &:= -\int_{\mathcal{I}_{y,i}(t)^c} \frac{\rho_{y,t}(E+\mathfrak{e}^+_{y,t})}{(\hat{z}_i(t,\alpha)-E)^2}\,\mathrm{d}E, && 1\le |i|\le N^{\omega_A}, \\ \mathcal{V}_i &:= -\int_{\mathcal{I}_{z,i}(t)^c\cap\mathcal{J}_z(t)} \frac{\bar{\rho}_t(E+\bar{\mathfrak{e}}^+_t)}{(\hat{z}_i(t,\alpha)-E)^2}\,\mathrm{d}E, && N^{\omega_A}<|i|\le \tfrac{1}{2}i_*, \end{aligned} \tag{6-107}$$

and $\mathcal{V}_i=0$ for $|i|>\frac{1}{2}i_*$. The error term $\zeta_i^{(0)}=\zeta_i^{(0)}(t)$ in (6-105) is defined as follows: for $|i|>\frac{1}{2}i_*$ we have

$$\zeta_i^{(0)} := \frac{1}{N}\sum_j^{\mathcal{A}^c,(i)} \frac{\partial_\alpha\tilde{z}_j(t,\alpha)-\partial_\alpha\tilde{z}_i(t,\alpha)}{(\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha))^2} + \partial_\alpha\Phi_\alpha(t) =: Z_1 + \partial_\alpha\Phi_\alpha(t), \tag{6-108}$$

for $N^{\omega_A} < |i| \le \frac{1}{2} i_*$ we have

$$\zeta_i^{(0)} := \frac{1}{N} \sum_{|j| \ge \frac{3}{4} i_*} \frac{\partial_\alpha \tilde{z}_j(t,\alpha) - \partial_\alpha \tilde{z}_i(t,\alpha)}{(\tilde{z}_i(t,\alpha) - \tilde{z}_j(t,\alpha))^2} + \int_{\mathcal{I}_{z,i}(t)^c \cap \mathcal{J}_z(t)} \frac{\partial_\alpha [_t(E + \bar{\mathfrak{e}}_t^+)]}{\hat{z}_i(t,\alpha) - E} \, \mathrm{d}E$$
$$+ \left(\partial_\alpha \int_{\mathcal{I}_{z,i}(t)^c \cap \mathcal{J}_z(t)} \right) \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}_t^+)}{\hat{z}_i(t,\alpha) - E} \, \mathrm{d}E + \partial_\alpha \Phi_\alpha(t)$$
$$=: Z_2 + Z_3 + Z_4 + \partial_\alpha \Phi_\alpha(t), \tag{6-109}$$

and finally for $1 \le |i| \le N^{\omega_A}$ we have $\zeta_i^{(0)} = 0$. We recall that $\mathcal{I}_{z,i}(t)$ and $\mathcal{J}_z(t)$ in (6-109) are defined by (6-96) and (6-95) respectively. Next, we prove that the error term $\zeta^{(0)}$ in (6-105) is bounded by some large power of N.

Lemma 6.11. *There exists a large constant $C > 0$ such that*

$$\sup_{0 \le t \le t_*} \max_{1 \le |i| \le N} |\zeta_i^{(0)}(t)| \le N^C. \tag{6-110}$$

Proof of Lemma 6.11. By (6-15), it follows that

$$\partial_\alpha \Phi_\alpha(t) = \partial_\alpha \Re[\bar{m}_t(\bar{\mathfrak{e}}_t^+ + \mathrm{i}N^{-100})] + h^{**}(t,1) - h^{**}(t,0),$$

with $h^{**}(t,\alpha)$ defined by (6-7). Since the two h^{**} terms are small by (6-6), for each fixed t, we have

$$|\partial_\alpha \Phi_\alpha(t)| \lesssim \left|\partial_\alpha \int_{\mathbb{R}} \frac{\bar{\rho}_t(\bar{\mathfrak{e}}_t^+ + E)}{E - \mathrm{i}N^{-100}} \, \mathrm{d}E\right| + N^{-1} = U_1 + U_2 + N^{-1}, \tag{6-111}$$

where

$$U_1 := \left| \partial_\alpha \int_{\bar{\gamma}_{-i(\delta_*)}}^{\bar{\gamma}_{i(\delta_*)}} \frac{\bar{\rho}_t(\bar{\mathfrak{e}}_t^+ + E)}{E - \mathrm{i}N^{-100}} \, \mathrm{d}E \right| = \left| \partial_\alpha \int_{I_*} \frac{\bar{\rho}_t(\bar{\mathfrak{e}}_t^+ + \varphi_{\alpha,t}(s))}{\varphi_{\alpha,t}(s) - \mathrm{i}N^{-100}} \varphi_{\alpha,t}'(s) \, \mathrm{d}s \right|,$$
$$U_2 := \left| \frac{1}{N} \sum_{i_*(\delta) < |i| \le N} \partial_\alpha \int_{\mathbb{R}} \frac{\psi(E - \bar{\gamma}_i^*(t))}{E - \mathrm{i}N^{-100}} \, \mathrm{d}E \right|,$$

using the notation $\bar{\gamma}_{i(\delta_*)} = \bar{\gamma}_{i(\delta_*)}(t)$ and the definition of $\bar{\rho}_t$ from (5-12). In U_1 we changed variables, i.e., $E = \varphi_{\alpha,t}(s)$, using that $s \to \varphi_{\alpha,t}(s)$ is strictly increasing. In particular, in order to compute the limits of integration we used that $\varphi_{\alpha,t}(i/N) = \bar{\gamma}_i(t)$ by (4-2) and defined the α-independent interval $I_* := [-i(\delta_*)/N, i(\delta_*)/N]$. Furthermore, in U_1 we denoted by prime the s-derivative.

For U_1 we have that (omitting the t-dependence, $\bar{\rho} = \bar{\rho}_t$, etc.)

$$U_1 \lesssim \left| \int_{I_*} \frac{\partial_\alpha [\bar{\rho}(\bar{\mathfrak{e}}^+ + \varphi_\alpha(s))]}{\varphi_\alpha(s) - \mathrm{i}N^{-100}} \varphi_\alpha'(s) \, \mathrm{d}s \right| + \left| \int_{I_*} \frac{\bar{\rho}(\bar{\mathfrak{e}}^+ + \varphi_\alpha(s))}{(\varphi_\alpha(s) - \mathrm{i}N^{-100})^2} (\varphi_\alpha'(s))^2 \, \mathrm{d}s \right|$$
$$+ \left| \int_{I_*} \frac{\bar{\rho}(\bar{\mathfrak{e}}^+ + \varphi_\alpha(s))}{\varphi_\alpha(s) - \mathrm{i}N^{-100}} \partial_\alpha \varphi_\alpha'(s) \, \mathrm{d}s \right|. \tag{6-112}$$

For $s \in I_*$, by the definition of $\varphi_\alpha(s)$ and (4-4) it follows that

$$1 = n_\alpha'(\varphi_\alpha(s)) \varphi_\alpha'(s) = \rho_\alpha(\varphi_\alpha(s)) \varphi_\alpha(s),$$

and hence

$$\varphi_\alpha'(s) = \frac{1}{\rho_\alpha(\varphi_\alpha(s))} \lesssim s^{-\frac{1}{4}}, \tag{6-113}$$

where in the last inequality we used that $\rho_\alpha(\omega) \sim \min\{\omega^{\frac{1}{3}}, \omega^{\frac{1}{2}}\Delta^{-\frac{1}{6}}\}$ and $\varphi_\alpha(s) \sim \max\{s^{\frac{3}{4}}, s^{\frac{2}{3}}\Delta^{\frac{1}{9}}\}$ by (4-9a).

In the first integral in (6-112) we use that

$$\bar{\rho}(\bar{\mathfrak{e}}^+ + \varphi_\alpha(s)) = \rho_\alpha(\bar{\mathfrak{e}}^+ + \varphi_\alpha(s)), \quad s \in I_*,$$

by (5-12) and that $\partial_\alpha[\rho_\alpha(\bar{\mathfrak{e}}^+ + \varphi_\alpha(s))]$ is bounded by the explicit relation in (4-10). For the other two integrals in (6-112) we use that $\bar{\rho}$ is bounded on the integration domain and that $(\varphi_\alpha'(s))^2 \lesssim s^{-\frac{1}{2}}$ from (6-113); hence it is integrable. In the third integral we also observe that

$$\partial_\alpha \varphi_\alpha(s) = \varphi_\lambda(s) - \varphi_\mu(s)$$

by (4-2); thus $|\partial_\alpha \varphi_\alpha'(s)| \lesssim s^{-\frac{1}{4}}$ similarly to (6-113). Using $|\varphi_\alpha(s) - \mathrm{i}N^{-100}| \gtrsim N^{-100}$, we conclude that

$$U_1 \lesssim N^{200}.$$

Next, we proceed with the estimate for U_2.

Notice $|\partial_\alpha \psi(E - \bar{\gamma}_i^*(t))| \le \|\psi'\|_\infty |\hat{\gamma}_{x,i}(t) - \hat{\gamma}_{y,i}(t)|$ by (5-10). Furthermore, since $|E - \mathrm{i}N^{-100}| \gtrsim \delta_*$ on the domain of integration of U_2, we conclude that

$$U_2 \lesssim N^{200} \|\psi'\|_\infty,$$

and therefore from (6-111) we have

$$|\partial_\alpha \Phi_\alpha(t)| \lesssim N^{202}, \tag{6-114}$$

since $\|\psi'\|_\infty \lesssim N^2$ by the choice of ψ; see below (5-9).

Similarly, we conclude that

$$|Z_3| \lesssim N^{200} \|\psi'\|_\infty. \tag{6-115}$$

To estimate Z_2, by (6-14), it follows that

$$\mathrm{d}(\partial_\alpha \tilde{z}_i) = \left[\frac{1}{N}\sum_j \frac{\partial_\alpha \tilde{z}_j - \partial_\alpha \tilde{z}_i}{(\tilde{z}_i - \tilde{z}_j)^2} + \partial_\alpha \Phi_\alpha(t)\right]\mathrm{d}t,$$

with initial data

$$\partial_\alpha \tilde{z}_i(0, \alpha) = \tilde{x}_i(0) - \tilde{y}_i(0),$$

for all $1 \le |i| \le N$. Since $|\partial_\alpha \tilde{z}_i(0, \alpha)| \lesssim N^{200}$ for all $1 \le |i| \le N$, by Duhamel's principle and contraction, it follows that

$$|\partial_\alpha \tilde{z}_i(t, \alpha)| \lesssim N^{200} + t_* \max_{0 \le \tau \le t_*} |\partial_\alpha \Phi_\alpha(\tau)| \lesssim N^{202} \tag{6-116}$$

for all $0 \le t \le t_*$. In particular, by (6-116) it follows that

$$|Z_2| \lesssim N^{202}\sqrt{N} \tag{6-117}$$

since for all j in the summation in Z_2 we have $|i - j| \gtrsim i_* \sim N^{\frac{1}{2}}$ and thus $|\tilde{z}_i - \tilde{z}_j| \gtrsim |i - j|/N \gtrsim N^{-\frac{1}{2}}$.

Finally, we estimate Z_4 using the fact that the endpoints of $\mathcal{I}_{z,i}(t)^c \cap \mathcal{J}_z(t)$ are quantiles $\bar{\gamma}_i(t)$ whose α-derivatives are bounded by (5-10). Hence

$$\left|Z_4\right| \lesssim \left|\frac{\bar{\rho}_t(\bar{\gamma}_{j_+} + \bar{\mathfrak{e}}_t^+)}{\hat{z}_i - \bar{\gamma}_{j_+}}\right| + \left|\frac{\bar{\rho}_t(\bar{\gamma}_{j_-} + \bar{\mathfrak{e}}_t^+)}{\hat{z}_i - \bar{\gamma}_{j_-}}\right| + \left|\frac{\bar{\rho}_t(\bar{\gamma}_{\frac{3}{4}i_*} + \bar{\mathfrak{e}}_t^+)}{\hat{z}_i - \bar{\gamma}_{\frac{3}{4}i_*}}\right| \lesssim N \tag{6-118}$$

by rigidity. Combining (6-114)–(6-118) we conclude (6-110), completing the proof of Lemma 6.11. □

Continuing the analysis of (6-105), for any fixed α let us define $w^\# = w^\#(t,\alpha)$ as the solution of

$$\partial_t w^\# = \mathcal{L} w^\#, \tag{6-119}$$

with cutoff initial data

$$w_i^\#(0,\alpha) = \mathbf{1}_{\{|i| \le N^{4\omega_\ell+\delta}\}} w_i(0,\alpha),$$

with some $0 < \delta < C\omega_\ell$, where $C > 10$ a constant such that $(4+C)\omega_\ell < \omega_A$.

By the rigidity (6-102), the finite-speed estimate (B-34), with $\delta' := \delta$, for the propagator $\mathcal{U}^{\mathcal{L}}$ of $\mathcal{L}$ holds. Let $0 < \delta_1 < \frac{1}{2}\delta$; then, using Duhamel's principle, that the error term $\zeta_i^{(0)}$ is bounded by (6-110) and that $\zeta_i^{(0)} = 0$ for any $1 \le |i| \le N^{\omega_A}$, it easily follows that

$$\sup_{0\le t\le t_*} \max_{|i|\le N^{4\omega_\ell+\delta_1}} |w_i^\#(t,\alpha) - w_i(t,\alpha)| \le N^{-100} \tag{6-120}$$

for any $\alpha \in [0,1]$. In other words, the initial conditions far away do not influence the w-dynamics; hence they can be set zero.

Next, we use the heat kernel contraction for the equation in (6-119). By the optimal rigidity of $\hat{x}_i(0)$ and $\hat{y}_i(0)$, since $w_i^\#(0,\alpha)$ is nonzero only for $1 \le |i| \le N^{4\omega_\ell+\delta}$, it follows that

$$\max_{1\le|i|\le N} |w_i^\#(0,\alpha)| \le \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}}, \tag{6-121}$$

and so, by heat kernel contraction and Duhamel's principle,

$$\sup_{0\le t\le t_*} \max_{1\le|i|\le N} |w_i^\#(t,\alpha)| \le \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}}. \tag{6-122}$$

Next, we recall that $\hat{z}(t,\alpha=0) = \hat{y}(t)$.

Combining (6-120) and (6-122), integrating $w_i(t,\alpha')$ over $\alpha' \in [0,\alpha]$, and by a high-moment Markov inequality as in (6-45)–(6-46), we conclude that

$$\sup_{0\le t\le t_*} |\hat{z}_i(t,\alpha) - \hat{y}_i(t)| \le \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}}, \quad 1 \le |i| \le N^{4\omega_\ell+\delta_1},$$

for any fixed $\alpha \in [0,1]$ with very high probability for any $\xi > 0$. Since

$$|\hat{z}_i(t,\alpha) - \bar{\gamma}_i(t)| \le |\hat{y}_i(t) - \hat{\gamma}_{y,i}(t)| + |\bar{\gamma}_i(t) - \hat{\gamma}_{y,i}(t)| + \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}}$$

for all $1 \le |i| \le N^{4\omega_\ell+\delta_1}$ and $\alpha \in [0,1]$, by (4-18) and the optimal rigidity of $\hat{y}_i(t)$, see (6-3), we conclude that

$$\sup_{0\le t\le t_*} |\hat{z}_i(t,\alpha)-\bar{\gamma}_i(t)| \le \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}}, \quad 1 \le |i| \le N^{4\omega_\ell+\delta_1}, \tag{6-123}$$

for any fixed $\alpha \in [0,1]$, for any $\xi > 0$ with very high probability. This concludes the proof of (6-103). □

6E. *Phase 3*: *rigidity for $\hat{z}$ with the correct i-dependence.* In this subsection we will prove almost-optimal i-dependent rigidity for the short-range approximation $\hat{z}_i(t,\alpha)$ (see (6-97)–(6-100)) for $1 \le |i| \le N^{4\omega_\ell+\delta_1}$.

Proposition 6.12. *Let δ_1 be defined in Proposition 6.10; then, for any fixed $\alpha \in [0,1]$, we have*

$$\sup_{0\le t\le t_*} |\hat{z}_i(t,\alpha)-\bar{\gamma}_i(t)| \lesssim \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}, \quad 1 \le |i| \le N^{4\omega_\ell+\delta_1}, \tag{6-124}$$

for any $\xi > 0$ with very high probability.

Proof. Define

$$K := \lceil N^\xi \rceil;$$

then (6-103) (with $\xi \to \frac{1}{2}\xi$) implies (6-124) for all $1 \le |i| \le 2K$. Next, we prove (6-124) for all $2K \le |i| \le N^{4\omega_\ell+\delta_1}$ by coupling $\tilde{x}_i(t)$ with $\tilde{y}_{\langle i-K\rangle}(t)$, where we make the following notational convention:

$$\langle i-K\rangle := \begin{cases} i-K & \text{if } i \in [K+1,N]\cup[-N,-1], \\ i-K-1 & \text{if } i \in [1,K]. \end{cases} \tag{6-125}$$

This slight complication is due to our indexing convention that excludes $i = 0$.

In order to couple the Brownian motion of $\tilde{x}_i(t)$ with the one of $\tilde{y}_{\langle i-K\rangle}(t)$, we construct a new process $\tilde{z}^*(t,\alpha)$ satisfying

$$\mathrm{d}\tilde{z}_i^*(t,\alpha) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_{\langle i-K\rangle} + \left[\frac{1}{N}\sum_{j\ne i}\frac{1}{\tilde{z}_i^*(t,\alpha)-\tilde{z}_j^*(t,\alpha)} + \Phi_\alpha(t)\right]\mathrm{d}t, \quad 1 \le |i| \le N, \tag{6-126}$$

with initial data

$$\tilde{z}_i^*(0,\alpha) = \alpha\tilde{x}_i(0) + (1-\alpha)\tilde{y}_{\langle i-K\rangle}(0), \tag{6-127}$$

for any $\alpha \in [0,1]$. Notice that the only difference with respect to $\tilde{z}_i(t,\alpha)$ from (6-14) is a shift in the index of the Brownian motion; i.e., $\tilde{z}$ and $\tilde{z}^*$ (almost) coincide in distribution, but their coupling to the y-process is different. The slight discrepancy comes from the effect of the few extreme indices. Indeed, to make the definition (6-126) unambiguous even for extreme indices, $i \in [-N,-N+K-1]$, additionally we need to define independent Brownian motions B_j and initial padding particles $\tilde{y}_j(0) = -jN^{300}$ for $j = -N-1,\dots,-N-K$. Similarly to Lemma 5.1, the effect of these very distant additional particles is negligible on the dynamics of the particles for $1 \le |i| \le \epsilon N$ for some small ϵ.

Next, we define the process $\hat{z}^*(t,\alpha)$ as the short-range approximation of $\tilde{z}^*(t,\alpha)$, given by (6-97)–(6-99) but with B_i replaced with $B_{\langle i-K\rangle}$ and we use initial data $\hat{z}^*(0,\alpha) = \tilde{z}^*(0,\alpha)$. In particular,

$$\hat{z}_i^*(t,1) = \hat{x}_i(t) + O(N^{-100}), \quad \hat{z}_i^*(t,0) = \hat{y}_{\langle i-K\rangle}(t) + O(N^{-100}), \quad 1 \le |i| \le \epsilon N, \tag{6-128}$$

the discrepancy again coming from the negligible effect of the additional K distant particles on the particles near the cusp regime.

Let $w_i^*(t,\alpha) := \partial_\alpha \hat{z}_i^*(t,\alpha)$; i.e., $w^* = w^*(t,\alpha)$ is a solution of

$$\partial_t w^* = \mathcal{B}w^* + \mathcal{V}w^* + \zeta^{(0)},$$

with initial data

$$w_i^*(0,\alpha) = \hat{x}_i^*(0) - \hat{y}_{\langle i-K\rangle}(0).$$

The operators $\mathcal{B}$, $\mathcal{L}$ and the error term $\zeta^{(0)}$ are defined as in (6-106)–(6-109) with all $\tilde{z}$ and $\hat{z}$ replaced by $\tilde{z}^*$ and $\hat{z}^*$, respectively.

We now define $(w^*)^\#$ as the solution of

$$\partial_t (w^*)^\# = \mathcal{L}(w^*)^\#, \tag{6-129}$$

with cutoff initial data

$$(w_i^*)^\#(0,\alpha) = \mathbf{1}_{\{|i|\le N^{4\omega_\ell+\delta}\}} w_i^*(0,\alpha),$$

with $0 < \delta < C\omega_\ell$ with $C > 10$ such that $(4+C)\omega_\ell < \omega_A$.

We claim that

$$(w_i^*)^\#(0,\alpha) \ge 0, \quad 1 \le |i| \le N. \tag{6-130}$$

We need to check it for $1 \le |i| \le N^{4\omega_\ell+\delta}$, otherwise $(w_i^*)^\#(0,\alpha) = 0$ by the cutoff. In the regime $1 \le |i| \le N^{4\omega_\ell+\delta}$ we use the optimal rigidity (Lemma 6.1 with $\xi \to \frac{1}{10}\xi$) for $\hat{x}_i^*(0)$ and $\hat{y}_{\langle i-K\rangle}(0)$, which yields

$$\begin{aligned}(w_i^*)^\#(0,\alpha) &= \hat{x}_i^*(0) - \hat{y}_{\langle i-K\rangle}(0)\\ &\ge -N^{\frac{1}{10}\xi}\eta_f(\gamma_{x,i}^*(0)) + \hat{\gamma}_{x,i}(0) - \hat{\gamma}_{y,\langle i-K\rangle}(0) - N^{\frac{1}{10}\xi}\eta_f(\gamma_{y,\langle i-K\rangle}^*(0)).\end{aligned} \tag{6-131}$$

We now check that $\hat{\gamma}_{x,i}(0) - \hat{\gamma}_{y,\langle i-K\rangle}(0)$ is sufficiently positive to compensate for the $N^{\frac{1}{10}\xi}\eta_f$ error terms. Indeed, by (4-13a) and (4-18), for all $|i| \ge 2K$ we have

$$\hat{\gamma}_{x,i}(t) - \hat{\gamma}_{y,\langle i-K\rangle}(t) \gtrsim K\eta_f(\gamma_{x,i}^*(t)) \gg N^{\frac{1}{10}\xi}\eta_f(\gamma_{x,i}^*(t))$$

and

$$\eta_f(\gamma_{y,\langle i-K\rangle}^*(t)) \sim \eta_f(\gamma_{x,i}^*(t)).$$

This shows (6-130) in the $2K \le |i| \le N^{4\omega_\ell+\delta}$ regime. If $K \le |i| \le 2K$ or $-K \le i \le -1$ we have that $(w_i^*)^\#(0,\alpha) \ge 0$ since

$$\hat{\gamma}_{x,i}(0) - \hat{\gamma}_{y,\langle i-K\rangle}(0) \gtrsim \max\left\{\frac{K^{\frac{3}{4}}}{N^{\frac{3}{4}}}, (t_* - t)^{\frac{1}{6}}\frac{K^{\frac{2}{3}}}{N^{\frac{2}{3}}}\right\} \gtrsim K\max\{\eta_f(\gamma_{x,i}^*(0)), \eta_f(\gamma_{y,\langle i-K\rangle}^*(0))\},$$

so $\hat{\gamma}_{x,i}(0) - \hat{\gamma}_{y,\langle i-K\rangle}(0)$ beats the error terms $N^{\frac{1}{10}\xi}\eta_f$ as well. Finally, if $1 \le i \le K-1$, the bound in (6-131) is easy since $\hat{\gamma}_{x,i}(0)$ and $\hat{\gamma}_{y,\langle i-K\rangle}(0)$ have opposite signs; i.e., they are in two different sides of the small gap and one of them is at least of order $(K/N)^{\frac{3}{4}}$, beating $N^{\frac{1}{10}\xi}\eta_f$. This proves (6-130). Hence, by the maximum principle we conclude that

$$(w_i^*)^\#(t,\alpha) \ge 0, \quad 0 \le t \le t_*, \ \alpha \in [0,1]. \tag{6-132}$$

Let $\delta_1 < \frac{1}{2}\delta$ be defined in Proposition 6.10. The rigidity estimate in (6-102) holds for $\hat{z}^*$ as well, since $\hat{z}$ and $\hat{z}^*$ have the same distribution. Furthermore, by (6-102) the propagator $\mathcal{U}$ of $\mathcal{L} := \mathcal{B} + \mathcal{V}$ satisfies the finite-speed estimate in Lemma B.3. Then, using Duhamel's principle and (6-110), we obtain

$$\sup_{0\le t\le t_*} \max_{1\le |i|\le N^{4\omega_\ell+\delta_1}} |(w_i^*)^\#(t,\alpha) - w_i^*(t,\alpha)| \le N^{-100} \tag{6-133}$$

for any $\alpha \in [0,1]$ with very high probability.

By (6-133), integrating $w_i^*(t,\alpha')$ over $\alpha' \in [0,\alpha]$, we conclude that

$$\hat{z}_i^*(t,\alpha) - \hat{y}_{\langle i-K\rangle}(t) \ge -N^{-100}, \quad 1 \le |i| \le N,^{4\omega_\ell+\delta_1}, \tag{6-134}$$

for all $\alpha \in [0,1]$ and $0 \le t \le t_*$ with very high probability. Note that in order to prove (6-134) with very high probability we used a Markov inequality as in (6-45)–(6-46). Hence,

$$\begin{aligned} \hat{z}_i^*(t,\alpha) - \bar{\gamma}_i(t) &\ge [\hat{y}_{\langle i-K\rangle}(t) - \hat{\gamma}_{y,\langle i-K\rangle}(t)] + [\hat{\gamma}_{y,\langle i-K\rangle}(t) - \hat{\gamma}_{y,i}(t)] + [\hat{\gamma}_{y,i}(t) - \bar{\gamma}_i(t)] - N^{-100} \\ &\gtrsim -K\big(\eta_f(\gamma^*_{y,\langle i-K\rangle}(t)) + \eta_f(\gamma^*_{y,i}(t))\big) - \gamma_i^*(t)t_*^{\frac{1}{3}} \\ &\ge -2K\big(\eta_f(\gamma^*_{y,\langle i-K\rangle}(t)) + \eta_f(\gamma^*_{y,i}(t))\big) \end{aligned} \tag{6-135}$$

for all $1 \le |i| \le N^{4\omega_\ell+\delta_1}$, where we used the optimal rigidity (6-3) and (4-18) in going to the second line. In particular, since for $|i| \ge 2K$ we have $\eta_f(\gamma^*_{y,i}(t)) \sim \eta_f(\gamma^*_{y,i-K}(t))$, we conclude that

$$\hat{z}_i^*(t,\alpha) - \bar{\gamma}_i(t) \ge -\frac{CKN^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}, \quad 2K \le |i| \le N^{4\omega_\ell+\delta_1}, \tag{6-136}$$

for all $0 \le t \le t_*$ and for any $\alpha \in [0,1]$. This implies the lower bound in (6-124).

In order to prove the upper bound in (6-124) we consider a very similar process $\tilde{z}_i^*(t,\alpha)$ (we continue to denote it by star) where the index shift in y is in the other direction. More precisely, it is defined as a solution of

$$\mathrm{d}\tilde{z}_i^*(t,\alpha) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_{\langle i+K\rangle} + \left[\frac{1}{N}\sum_{j\ne i}\frac{1}{\tilde{z}_i^*(t,\alpha) - \tilde{z}_j^*(t,\alpha)} + \Phi_\alpha(t)\right]\mathrm{d}t,$$

with initial data

$$\tilde{z}_i(0,\alpha) = \alpha\tilde{y}_{\langle i+K\rangle}(0) + (1-\alpha)\tilde{x}_i(0),$$

for any $\alpha \in [0,1]$. Here $\langle i+K\rangle$ is defined analogously to (6-125). Then, by similar computations, we conclude that

$$\hat{z}_i^*(t,\alpha) - \bar{\gamma}_i(t) \le \frac{KN^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}, \quad 2K \le |i| \le N^{4\omega_\ell+\delta_1}, \tag{6-137}$$

for all $0 \le t \le t_*$ and for any $\alpha \in [0,1]$. Combining (6-136) and (6-137) we conclude (6-124) and complete the proof of Proposition 6.12. ☐

7. Proof of Proposition 3.1: Dyson Brownian motion near the cusp

In this section $t_1 \leq t_*$, indicating that we are before the cusp formation; we recall that t_1 is defined as

$$t_1 := \frac{N^{\omega_1}}{N^{\frac{1}{2}}}$$

for a small fixed $\omega_1 > 0$ and t_* is the time of the formation of the exact cusp. The main result of this section is the following proposition, from which we can quickly prove Proposition 3.1 for $t_1 \leq t_*$. If $t_1 > t_*$ we conclude Proposition 3.1 using the analogous Proposition 8.1 instead of Proposition 7.1 exactly in the same way.

Proposition 7.1. *For $t_1 \leq t_*$, with very high probability we have*

$$|(\lambda_j(t_1) - \mathfrak{e}^+_{\lambda,t_1}) - (\mu_{j+i_\mu-i_\lambda}(t_1) - \mathfrak{e}^+_{\mu,t_1})| \leq N^{-\frac{3}{4}-c\omega_1} \tag{7-1}$$

for some small constant $c > 0$ and for any j such that $|j - i_\lambda| \leq N^{\omega_1}$.

Note that if $t_1 = t_*$ then $\mathfrak{e}^+_{r,t_*} = \mathfrak{e}^-_{r,t_*} = \mathfrak{c}_r$ for $r = \lambda, \mu$, with $\mathfrak{c}_r$ being the exact cusp point of the scDOSs ρ_{r,t_*}. The proof of Proposition 7.1 will be given at the end of the section after several auxiliary lemmas.

Proof of Proposition 3.1. Firstly, we recall the definition of the physical cusp

$$\mathfrak{b}_{r,t_1} := \begin{cases} \frac{1}{2}(\mathfrak{e}^+_{r,t_1} + \mathfrak{e}^-_{r,t_1}) & \text{if } t_1 < t_*, \\ \mathfrak{c}_r & \text{if } t_1 = t_*, \\ \mathfrak{m}_{r,t_1} & \text{if } t_1 > t_* \end{cases}$$

of ρ_{r,t_1} as in (2-5), for $r = \lambda, \mu$. Then, using the change of variables $\boldsymbol{x} = N^{\frac{3}{4}}(\boldsymbol{x}' - \mathfrak{b}_{r,t_1})$ for $r = \lambda, \mu$, and the definition of correlation function, for each Lipschitz continuous and compactly supported test function F, we have

$$\begin{aligned}
&\int_{\mathbb{R}^k} F(\boldsymbol{x})\left[N^{\frac{1}{4}k} p^{(N,\lambda)}_{k,t_1}\left(\mathfrak{b}_{\lambda,t_1} + \frac{\boldsymbol{x}}{N^{\frac{3}{4}}}\right) - N^{\frac{1}{4}k} p^{(N,\mu)}_{k,t_1}\left(\mathfrak{b}_{\mu,t_1} + \frac{\boldsymbol{x}}{N^{\frac{3}{4}}}\right)\right] \mathrm{d}\boldsymbol{x} \\
&= N^k \binom{N}{k}^{-1} \sum_{\{i_1,\dots,i_k\}\subset[N]} \left[\mathbb{E}_{H^{(\lambda)}_{t_1}} F(N^{\frac{3}{4}}(\lambda_{i_1} - \mathfrak{b}_{\lambda,t_1}), \dots, N^{\frac{3}{4}}(\lambda_{i_k} - \mathfrak{b}_{\lambda,t_1})) - \mathbb{E}_{H^{(\mu)}_{t_1}} F(\lambda \to \mu)\right],
\end{aligned} \tag{7-2}$$

where $\lambda_1, \dots, \lambda_N$ and $\mu_1, \dots, \mu_N$ are the eigenvalues, labelled in increasing order, of $H^{(\lambda)}_{t_1}$ and $H^{(\mu)}_{t_1}$ respectively. In $\mathbb{E}_{H^{(\mu)}_{t_1}} F(\lambda \to \mu)$ we also replace $\mathfrak{b}_{\lambda,t_1}$ by $\mathfrak{b}_{\mu,t_1}$.

In order to apply Proposition 7.1 we split the sum in the right-hand side of (7-2) into two sums,

$$\sum_{\substack{\{i_1,\dots,i_k\}\subset[N] \\ |i_1-i_\lambda|,\dots,|i_k-i_\lambda|<N^\epsilon}} \quad \text{and its complement} \quad \sum{}', \tag{7-3}$$

where ϵ is a positive exponent with $\epsilon \ll \omega_1$.

We start with the estimate for the second sum of (7-3). In particular, we will estimate only the term $\mathbb{E}_{H^{(\lambda)}_{t_1}}(\cdot)$; the estimate for $\mathbb{E}_{H^{(\mu)}_{t_1}}(\cdot)$ will follow in an analogous way.

Since the test function F is compactly supported in some set $\Omega \subset \mathbb{R}^k$ and in $\sum'$ there is at least one index i_l such that $|i_l - i_\lambda| \geq N^\epsilon$, we have

$$\sum{}' \mathbb{E}_{H_{t_1}^{(\lambda)}} F(N^{\frac{3}{4}}(\lambda_{i_1} - \mathfrak{b}_{\lambda,t_1}), \dots, N^{\frac{3}{4}}(\lambda_{i_k} - \mathfrak{b}_{\lambda,t_1})) \lesssim N^{k-1} \|F\|_\infty \sum_{i_l : |i_l - i_\lambda| \geq N^\epsilon} \mathbb{P}_{H_{t_1}^{(\lambda)}}\big(|\lambda_{i_l} - \mathfrak{b}_{\lambda,t_1}| \lesssim C_\Omega N^{-\frac{3}{4}}\big), \quad (7\text{-}4)$$

where C_Ω is the diameter of Ω. Let

$$\gamma_{\lambda,i} = \hat{\gamma}_{\lambda,i} + \mathfrak{e}^+_{\lambda,t_1}$$

be the classical eigenvalue locations of $\rho_\lambda(t_1)$ defined by (4-10) for all $1 - i_\lambda \leq i \leq N + 1 - i_\lambda$. Then, by the rigidity estimate from [Erdős et al. 2018, Corollary 2.6], we have

$$\mathbb{P}_{H_{t_1}^{(\lambda)}}\big(|\lambda_{i_l} - \mathfrak{b}_{\lambda,t_1}| \lesssim C_\Omega N^{-\frac{3}{4}}, |i_l - i_\lambda| \geq N^\epsilon\big) \leq N^{-D} \quad (7\text{-}5)$$

for each $D > 0$ if N is large enough, depending on C_Ω. Indeed, by rigidity it follows that

$$|\lambda_{i_l} - \mathfrak{b}_{\lambda,t_1}| \geq |\gamma_{\lambda,i_l} - \gamma_{\lambda,i_\lambda}| - |\lambda_{i_l} - \gamma_{\lambda,i_l}| - |\mathfrak{b}_{\lambda,t_1} - \gamma_{\lambda,i_\lambda}| \gtrsim \frac{N^{c\epsilon}}{N^{\frac{3}{4}}} - \frac{N^{c\xi}}{N^{\frac{3}{4}}} \gtrsim \frac{N^{c\epsilon}}{N^{\frac{3}{4}}} \quad (7\text{-}6)$$

with very high probability if $N^\epsilon \leq |i_l - i_\lambda| \leq \tilde{c}N$ for some $0 < \tilde{c} < 1$. In (7-6) we used the rigidity from [Erdős et al. 2018, Corollary 2.6] in the form

$$|\lambda_i - \gamma_{\lambda,i}| \leq \frac{N^\xi}{N^{\frac{3}{4}}},$$

for any $\xi > 0$, with very high probability. Note that (7-5) and (7-6) hold for any $\epsilon \gtrsim \xi$. If $|i_l - i_\lambda| \geq \tilde{c}N$, then $|\gamma_{i_l} - \gamma_{i_\lambda}| \sim 1$ and the bound in (7-6) clearly holds. A similar estimate holds for $H_{t_1}^{(\mu)}$; hence, choosing $D > k + 1$ we conclude that the second sum in (7-3) is negligible.

Next, we consider the first sum in (7-3). For $t_1 \leq t_*$ we have, by (4-6a), that

$$|(\mathfrak{e}^+_{\lambda,t_1} - \mathfrak{b}_{\lambda,t_1}) - (\mathfrak{e}^+_{\mu,t_1} - \mathfrak{b}_{\mu,t_1})| = \tfrac{1}{2}|\Delta_{\lambda,t_1} - \Delta_{\mu,t_1}| \lesssim \Delta_{\mu,t_1}(t_* - t_1)^{\frac{1}{3}} \leq N^{-\frac{3}{4}-\frac{1}{6}+C\omega_1}.$$

Hence, by (7-1), using that

$$|F(\boldsymbol{x}) - F(\boldsymbol{x}')| \lesssim \|F\|_{C^1} \|\boldsymbol{x} - \boldsymbol{x}'\|,$$

we conclude that

$$\sum_{\substack{\{i_1,\dots,i_k\} \subset [N] \\ |i_1 - i_\lambda|,\dots,|i_k - i_\lambda| \leq N^\epsilon}} [\mathbb{E}_{H_{t_1}^{(\lambda)}} F(N^{\frac{3}{4}}(\lambda_{i_1} - \mathfrak{b}_{\lambda,t_1}), \dots, N^{\frac{3}{4}}(\lambda_{i_k} - \mathfrak{b}_{\lambda,t_1})) - \mathbb{E}_{H_{t_1}^{(\mu)}} F(\lambda \to \mu)] \leq C_k \|F\|_{C^1} \frac{N^{k\epsilon}}{N^{c\omega_1}} \quad (7\text{-}7)$$

for some $c > 0$. Then, using that

$$\frac{N^k (N-k)!}{N!} = 1 + \mathcal{O}_k(N^{-1}),$$

we easily conclude the proof of Proposition 3.1. □

7A. ***Interpolation.*** In order to prove Proposition 7.1 we recall a few concepts introduced previously. In Section 5 we introduced the padding particles $x_i(t)$, $y_i(t)$, for $1 \le |i| \le N$, which are good approximations of the eigenvalues $\lambda_j(t)$, $\mu_j(t)$ respectively, for $1 \le j \le N$, in the sense of Lemma 5.1. They satisfy a Dyson Brownian motion equation (5-5), (5-7) mimicking the DBM of genuine eigenvalue processes (5-3), (5-4). It is more convenient to consider shifted processes where the edge motion is subtracted.

More precisely, for $r = x, y$ and $r(t) = x(t), y(t)$, we defined

$$\tilde{r}_i(t) := r_i(t) - \mathfrak{e}^+_{r,t}, \quad 1 \le |i| \le N,$$

for all $0 \le t \le t_*$. In particular, $\tilde{r}(t)$ is a solution of

$$\mathrm{d}\tilde{r}_i(t) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \left(\frac{1}{N}\sum_{j\ne i}\frac{1}{\tilde{r}_i(t)-\tilde{r}_j(t)} + \Re[m_{r,t}(\mathfrak{e}^+_{r,t})]\right)\mathrm{d}t, \tag{7-8}$$

with initial data

$$\tilde{r}_i(0) = r_i(0) - \mathfrak{e}^+_{r,0}, \tag{7-9}$$

for all $1 \le |i| \le N$.

Next, following a similar idea of [Landon and Yau 2017], we also introduced in (6-14) an interpolation process between $\tilde{x}(t)$ and $\tilde{y}(t)$. For any $\alpha \in [0, 1]$ we defined the process $\tilde{z}(t, \alpha)$ as the solution of

$$\mathrm{d}\tilde{z}_i(t,\alpha) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \left(\frac{1}{N}\sum_{j\ne i}\frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)} + \Phi_\alpha(t)\right)\mathrm{d}t, \tag{7-10}$$

with initial data

$$\tilde{z}_i(0,\alpha) = \alpha\tilde{x}_i(0) + (1-\alpha)\tilde{y}_i(0),$$

for each $1 \le |i| \le N$. Recall that $\Phi_\alpha(t)$ was defined in (6-15) and it is such that $\Phi_0(t) = \Re[m_{y,t}(\mathfrak{e}^+_{y,t})]$ and $\Phi_1(t) = \Re[m_{x,t}(\mathfrak{e}^+_{x,t})]$. Note that $\tilde{z}_i(t,1) = \tilde{x}_i(t)$ and $\tilde{z}_i(t,0) = \tilde{y}_i(t)$ for all $1 \le |i| \le N$ and $0 \le t \le t_*$.

We recall the definition of the interpolated quantiles from (5-10) of Section 5:

$$\bar{\gamma}_i(t) := \alpha\hat{\gamma}_{x,i}(t) + (1-\alpha)\hat{\gamma}_{y,i}(t), \quad \alpha \in [0, 1], \tag{7-11}$$

where $\hat{\gamma}_{x,i}$ and $\hat{\gamma}_{y,i}$ are the shifted quantiles of $\rho_{x,t}$ and $\rho_{y,t}$ respectively, defined in Section 5. In particular,

$$\bar{\mathfrak{e}}^\pm_t = \alpha\mathfrak{e}^\pm_{x,t} + (1-\alpha)\mathfrak{e}^\pm_{y,t}, \quad \alpha \in [0, 1].$$

We denoted the interpolated density, whose quantiles are the $\bar{\gamma}_i(t)$, by $\bar{\rho}_t$ (5-12), and its Stieltjes transform by $\bar{m}_t$.

Let $\hat{z}(t,\alpha)$ be the short-range approximation of $\tilde{z}(t,\alpha)$ defined by (6-97)–(6-99), with exponents $\omega_1 \ll \omega_\ell \ll \omega_A \ll 1$ and with initial data $\hat{z}(0,\alpha) = \tilde{z}(0,\alpha)$ and $i_* = N^{\frac{1}{2}+C_*\omega_1}$ for some large constant $C_* > 0$. In particular, $\hat{x}(t) = \hat{z}(t,1)$ and $\hat{y}(t) = \hat{z}(t,0)$. Assuming optimal rigidity in (6-3) for $\tilde{r}_i(t) = \tilde{x}_i(t), \tilde{y}_i(t)$, the following lemma shows that the process $\tilde{r}$ and its short-range approximation $\hat{r} = \hat{x}, \hat{y}$ stay very close to each other; i.e., $|\hat{r}_i - \tilde{r}_i| \le N^{-\frac{3}{4}-c}$ for some small $c > 0$. This is the analogue of Lemma 3.7 in [Landon and Yau 2017] and its proof, given in Appendix C, follows similar lines. It assumes the optimal rigidity, see (7-12) below, which is ensured by [Erdős et al. 2018, Corollary 2.6]; see Lemma 6.1.

Lemma 7.2. *Let* $i_* = N^{\frac{1}{2}+C_*\omega_1}$. *Assume that* $\tilde{z}(t,0)$ *and* $\tilde{z}(t,1)$ *satisfy the optimal rigidity*

$$\sup_{0\le t\le t_1} |\tilde{z}_i(t,\alpha)-\hat{\gamma}_{r,i}(t)| \le N^{\xi}\eta_{\mathrm{f}}^{\rho_{r,t}}(e_{r,t}^{+}+\hat{\gamma}_{r,\pm i}(t)), \quad 1\le |i|\le i_*,\ \alpha=0,1, \tag{7-12}$$

with $r = x, y$, *for any* $\xi > 0$, *with very high probability. Then, for* $\alpha = 0$ *or* $\alpha = 1$ *we have*

$$\sup_{1\le |i|\le N}\sup_{0\le t\le t_1} |\tilde{z}_i(t,\alpha)-\hat{z}_i(t,\alpha)| \lesssim \frac{N^{\frac{1}{6}\omega_1}N^{\xi}}{N^{\frac{3}{4}}}\left(\frac{N^{\omega_1}}{N^{3\omega_\ell}}+\frac{N^{\omega_1}}{N^{\frac{1}{8}}}+\frac{N^{C\omega_1}N^{\frac{1}{2}\omega_A}}{N^{\frac{1}{6}}}+\frac{N^{\frac{1}{2}\omega_A}N^{C\omega_1}}{N^{\frac{1}{4}}}+\frac{N^{C\omega_1}}{N^{\frac{1}{18}}}\right) \tag{7-13}$$

for any $\xi > 0$, *with very high probability.*

In particular, (7-13) implies that there exists a small fixed universal constant $c > 0$ such that

$$\sup_{1\le |i|\le N}\sup_{0\le t\le t_1} |\tilde{z}_i(t,\alpha)-\hat{z}_i(t,\alpha)| \lesssim N^{-\frac{3}{4}-c}, \quad \alpha=0,1, \tag{7-14}$$

with very high probability.

Remark 7.3. Note the denominator in the first error term in (7-13): the factor $N^{3\omega_\ell}$ is better than $N^{2\omega_\ell}$ in Lemma 3.7 in [Landon and Yau 2017]; this is because of the natural cusp scaling. The fact that this power is at least $N^{(1+\epsilon)\omega_\ell}$ was essential in that paper since this allowed them to transfer the optimal rigidity from $\tilde{z}$ to the $\hat{z}$-process for all $\alpha \in [0,1]$. Optimal rigidity for $\hat{z}$ is essential (i) for the heat kernel bound for the propagator of $\mathcal{L}$, see (6-105)–(6-106), and (ii) for a good ℓ^p-norm for the initial condition in (7-25). With our approach, however, this power in (7-13) is not critical since we have already obtained an even better, i-dependent rigidity for the $\hat{z}$-process for any α by using the maximum principle; see Proposition 6.12. We still need (7-13) for the x- and y-processes (i.e., only for $\alpha = 0, 1$), but only with a precision below the rigidity scale; therefore the denominator in the first term has only to beat $N^{\frac{7}{6}\omega_1+\xi}$.

7B. ***Differentiation.*** Next, we consider the α-derivative of the process $\hat{z}(t,\alpha)$. Let

$$u_i(t) = u_i(t,\alpha) := \partial_\alpha \hat{z}_i(t,\alpha), \quad 1\le |i|\le N;$$

then u is a solution of the equation

$$\partial_t u = \mathcal{L}u + \zeta^{(0)}, \tag{7-15}$$

where $\zeta^{(0)}$, defined by (6-108)–(6-109), is an error term that is nonzero only for $|i| > N^{\omega_A}$ and such that $|\zeta_i^{(0)}| \lesssim N^C$ for some large constant $C > 0$ with very high probability, by (6-110), and the operator $\mathcal{L} = \mathcal{B}+\mathcal{V}$ acting on $\mathbb{R}^{2N}$ is defined by (6-106)–(6-107).

In the following by $\mathcal{U}^{\mathcal{L}}$ we denote the semigroup associated to (7-15); i.e., by Duhamel's principle

$$u(t) = \mathcal{U}^{\mathcal{L}}(0,t)u(0) + \int_0^t \mathcal{U}^{\mathcal{L}}(s,t)\zeta^{(0)}(s)\,\mathrm{d}s$$

and $\mathcal{U}^{\mathcal{L}}(s,s) = \mathrm{Id}$ for all $0 \le s \le t$. Furthermore, for each a, b such that $|a|, |b| \le N$, by $\mathcal{U}^{\mathcal{L}}_{ab}$ we denote the entries of $\mathcal{U}^{\mathcal{L}}$, which can be seen as the solution of (7-15) with initial condition $u_a(0) = \delta_{ab}$.

By Proposition 6.3 and Lemma C.1, for any fixed $\alpha \in [0,1]$, it follows that

$$\sup_{0 \le t \le t_*} |\hat{z}_i(t,\alpha) - \bar{\gamma}_i(t)| \lesssim \frac{N^{C\omega_1}}{N^{\frac{1}{2}}}, \quad 1 \le |i| \le N, \tag{7-16}$$

$$\sup_{0 \le t \le t_*} |\hat{z}_i(t,\alpha) - \bar{\gamma}_i(t)| \lesssim \frac{N^{C\omega_1}}{N^{\frac{3}{4}}}, \quad 1 \le |i| \le i_*, \tag{7-17}$$

with very high probability. Then, using (7-17), as a consequence of Lemma B.3 we have the following:

Lemma 7.4. *There exists a constant $C > 0$ such that for any $0 < \delta < C\omega_\ell$, if $1 \le |a| \le \frac{1}{2}N^{4\omega_\ell+\delta}$ and $|b| \ge N^{4\omega_\ell+\delta}$, then*

$$\sup_{0 \le s \le t \le t_*} \mathcal{U}^{\mathcal{L}}_{ab}(s,t) + \mathcal{U}^{\mathcal{L}}_{ba}(s,t) \le N^{-D} \tag{7-18}$$

for any $D > 0$ with very high probability.

Furthermore, by Proposition 6.12, for any fixed $\alpha \in [0,1]$, we have

$$\sup_{0 \le t \le t_*} |\hat{z}_i(t,\alpha) - \bar{\gamma}_i(t)| \lesssim \frac{N^{\xi} N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}} |i|^{\frac{1}{4}}}, \quad 1 \le |i| \le N^{4\omega_\ell+\delta_1}, \tag{7-19}$$

for some small fixed $\delta_1 > 0$ and for any $\xi > 0$ with very high probability.

Next, we introduce the ℓ^p norms

$$\|u\|_p := \Bigl(\sum_i |u_i|^p\Bigr)^{\frac{1}{p}}, \quad \|u\|_\infty := \max_i |u_i|.$$

Following a similar scheme to [Bourgade et al. 2014; Erdős and Yau 2015] with some minor modifications we will prove the following Sobolev-type inequalities in Appendix D.

Lemma 7.5. *For any small $\eta > 0$ there exists $c_\eta > 0$ such that*

$$\sum_{i \ne j \in \mathbb{Z}_+} \frac{(u_i - u_j)^2}{|i^{\frac{3}{4}} - j^{\frac{3}{4}}|^{2-\eta}} \ge c_\eta \Bigl(\sum_{i \ge 1} |u_i|^p\Bigr)^{\frac{2}{p}}, \quad \sum_{i \ne j \in \mathbb{Z}_-} \frac{(u_i - u_j)^2}{||i|^{\frac{3}{4}} - |j|^{\frac{3}{4}}|^{2-\eta}} \ge c_\eta \Bigl(\sum_{i \le -1} |u_i|^p\Bigr)^{\frac{2}{p}} \tag{7-20}$$

hold, with $p = 8/(2+3\eta)$, for any function $\|u\|_p < \infty$.

Using the Sobolev inequality in (7-20) and the finite-speed estimate of Lemma 7.4, in Appendix E we prove the energy estimates for the heat kernel in Lemma 7.6 via a Nash-type argument.

Lemma 7.6. *Assume (7-16), (7-17) and (7-19). Let $0 < \delta_4 < \frac{1}{10}\delta_1$ and $w_0 \in \mathbb{R}^{2N}$ such that $|(w_0)_i| \le N^{-100}\|w_0\|_1$ for $|i| \ge \ell^4 N^{\delta_4}$. Then, for any small $\eta > 0$ there exists a constant $C > 0$ independent of η and a constant c_η such that for all $0 \le s \le t \le t_*$*

$$\|\mathcal{U}^{\mathcal{L}}(s,t) w_0\|_2 \le \left(\frac{N^{C\eta + \frac{1}{3}\omega_1}}{c_\eta N^{\frac{1}{2}}(t-s)}\right)^{1-3\eta} \|w_0\|_1 \tag{7-21}$$

and

$$\|\mathcal{U}^{\mathcal{L}}(0,t)w_0\|_\infty \le \left(\frac{N^{C\eta+\frac{1}{3}\omega_1}}{c_\eta N^{\frac{1}{2}}t}\right)^{\frac{2}{p}(1-3\eta)}\|w_0\|_p \tag{7-22}$$

for each $p\ge 1$.

Let $0<\delta_v<\frac{1}{2}\delta_4$. Define $v_i=v_i(t,\alpha)$ to be the solution of

$$\partial_t v=\mathcal{L}v,\quad v_i(0,\alpha)=u_i(0,\alpha)\mathbf{1}_{\{|i|\le N^{4\omega_\ell+\delta_v}\}}. \tag{7-23}$$

Then, by Lemma 7.4 the next result follows.

Lemma 7.7. *Let u be the solution of the equation in* (7-15) *and v defined by* (7-23); *then we have*

$$\sup_{0\le t\le t_1}\sup_{|i|\le \ell^4}|u_i(t)-v_i(t)|\le N^{-100} \tag{7-24}$$

with very high probability.

Proof. By (7-15) and (7-23) have

$$u_i(t)-v_i(t)=\sum_{j=-N}^{N}\mathcal{U}^{\mathcal{L}}_{ij}(0,t)u_j(0)-\sum_{j=-N^{4\omega_\ell+\delta_v}}^{N^{4\omega_\ell+\delta_v}}\mathcal{U}^{\mathcal{L}}_{ij}(0,t)u_j(0)+\int_0^t\sum_{|j|\ge N^{\omega_A}}\mathcal{U}^{\mathcal{L}}_{ij}(s,t)\zeta_j^{(0)}(s)\,\mathrm{d}s.$$

Then, using that $\zeta_i^{(0)}=0$ for $1\le|i|\le N^{\omega_A}$ and (6-110), the bound in (7-24) follows by Lemma 7.4. □

Proof of Proposition 7.1. We consider only the $j=i_\lambda$ case. By Lemma 5.1 and (7-14) we have

$$\begin{aligned}|(\lambda_{i_\lambda}(t_1)-\mathfrak{e}^+_{\lambda,t_1})-(\mu_{i_\mu}(t_1)-\mathfrak{e}^+_{\mu,t_1})| &\le |\tilde{x}_1(t_1)-\hat{x}_1(t_1)|+|\hat{x}_1(t_1)-\hat{y}_1(t_1)|+|\hat{y}_1(t_1)-\tilde{y}_1(t_1)|\\ &\le |\hat{x}_1(t_1)-\hat{y}_1(t_1)|+N^{-\frac{3}{4}-c}\end{aligned}$$

with very high probability.

Since $\hat{z}_i(t_1,1)=\hat{x}_i(t_1)$ and $\hat{z}_i(t_1,0)=\hat{y}_i(t_1)$ for all $1\le|i|\le N$, by the definition of $u_i(t,\alpha)$, it follows that

$$\hat{x}_1(t_1)-\hat{y}_1(t_1)=\int_0^1 u_1(t_1,\alpha)\,\mathrm{d}\alpha.$$

Furthermore, by a high-moment Markov inequality as in (6-45)–(6-46) and Lemma 7.7, we get

$$\int_0^1|u_1(t_1,\alpha)|\,\mathrm{d}\alpha\lesssim N^{-100}+\int_0^1|v_1(t_1,\alpha)|\,\mathrm{d}\alpha.$$

Since $v_i(0)=u_i(0)\mathbf{1}_{\{|i|\le N^{4\omega_\ell+\delta_v}\}}$ and, by (4-18) and (6-3), for $1\le|i|\le N^{4\omega_\ell+\delta_v}$ we have

$$\begin{aligned}|u_i(0)| &\lesssim |\hat{x}_i(0)-\hat{\gamma}_{x,i}(0)|+|\hat{y}_i(0)-\hat{\gamma}_{y,i}(0)|+|\hat{\gamma}_{x,i}(0)-\hat{\gamma}_{y,i}(0)|\\ &\lesssim \frac{N^{\frac{1}{6}\omega_1}}{|i|^{\frac{1}{4}}N^{\frac{3}{4}}}+\frac{|i|^{\frac{3}{4}}N^{\frac{1}{2}\omega_1}}{N^{\frac{11}{12}}}\lesssim\frac{N^{\frac{1}{6}\omega_1}}{|i|^{\frac{1}{4}}N^{\frac{3}{4}}},\end{aligned}$$

we conclude that

$$\|v(0)\|_5 \lesssim \frac{N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}} \tag{7-25}$$

with very high probability. Hence, recalling that $t_1 = N^{-\frac{1}{2}+\omega_1}$, by (7-22) and Markov's inequality again, we get

$$\begin{aligned}\int_0^1 |v_1(t_1,\alpha)|\,\mathrm{d}\alpha \le \sup_{\alpha\in[0,1]} \|v(t_1,\alpha)\|_\infty &\le \left(\frac{N^{C\eta+\frac{1}{3}\omega_1}}{N^{\frac{1}{2}}t_1}\right)^{\frac{2}{5}(1-3\eta)} \|v(0)\|_5 \\ &\lesssim \frac{N^{\frac{1}{6}\omega_1+\frac{1}{5}\eta(2C+3\omega_1-6\eta C)}}{N^{\frac{3}{4}}N^{\frac{4}{15}\omega_1}} = \frac{1}{N^{\frac{3}{4}}N^{\frac{1}{20}\omega_1}}\end{aligned} \tag{7-26}$$

with very high probability, for η small enough, say $\eta \le \omega_1(8C+12\omega_1)^{-1}$. Notice that the constant in front of the ω_1 in the exponents plays a crucial role: eventually the constant $\left(1-\frac{1}{3}\right)\frac{2}{5} = \frac{4}{15}$ from the Nash estimate beats the constant $\frac{1}{6}$ from (7-25). This completes the proof of Proposition 7.1. □

8. Case of $t \ge t_*$: small minimum

In this section we consider the case when the densities $\rho_{x,t}$, $\rho_{y,t}$, hence their interpolation $\bar{\rho}_t$ as well, have a small minimum, i.e., $t_* \le t \le 2t_*$. We deal with the small minimum case in this separate section mainly for notational reasons: for $t_* \le t \le 2t_*$ the processes $x(t)$ and $y(t)$, and consequently the associated quantiles and densities, are shifted by $\tilde{\mathfrak{m}}_{r,t}$, for $r = x, y$, instead of $\mathfrak{e}^+_{r,t}$. We recall that $\tilde{\mathfrak{m}}_{r,t}$, defined in (4-14a), denotes a close approximation of the actual local minimum $\mathfrak{m}_{r,t}$ near the physical cusp. We chose to shift $x(t)$ and $y(t)$ by the tilde approximation of the minimum instead of the minimum itself for technical reasons, namely because the t-derivative of $\tilde{\mathfrak{m}}_{r,t}$, $r = x, y$, satisfies the convenient relation in (4-14d).

As we explained at the beginning of Section 7, in order to prove universality, i.e., Proposition 3.1 at time $t_1 \ge t_*$, it is enough to prove the following:

Proposition 8.1. *For $t_1 \ge t_*$, we have, with very high probability, that*

$$|(\lambda_j(t_1) - \mathfrak{m}_{\lambda,t_1}) - (\mu_{j+i_\mu-i_\lambda}(t_1) - \mathfrak{m}_{\mu,t_1})| \le N^{-\frac{3}{4}-c} \tag{8-1}$$

for some small constant $c > 0$ and for any j such that $|j - i_\lambda| \le N^{\omega_1}$. Here $\mathfrak{m}_{\lambda,t_1}$ and $\mathfrak{m}_{\mu,t_1}$ are the local minimums of ρ_{λ,t_1} and ρ_{μ,t_1}, respectively.

We introduce the shifted process $\tilde{r}_i(t) = \tilde{x}_i(t), \tilde{y}_i(t)$ for $t \ge t_*$. Let us define

$$\tilde{r}_i(t) := r_i(t) - \tilde{\mathfrak{m}}_{r,t}, \quad 1 \le |i| \le N, \tag{8-2}$$

for $r = x, y$; hence, by (4-14d), the shifted points satisfy the DBM

$$\mathrm{d}\tilde{r}_i(t) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \frac{1}{N}\sum_{j\ne i}\frac{1}{\tilde{r}_i(t)-\tilde{r}_j(t)}\,\mathrm{d}t - \left(\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{r,t}\right)\mathrm{d}t. \tag{8-3}$$

Furthermore we recall that by $\hat{\gamma}_{r,i}(t)$ we denote the quantiles of $\rho_{r,t}$, with $r = x, y$, for all $t_* \le t \le 2t_*$; i.e.,

$$\hat{\gamma}_{r,i} = \gamma_{r,i} - \tilde{\mathfrak{m}}_{r,t}, \quad 1 \le |i| \le N.$$

By the rigidity estimate of [Erdős et al. 2018, Corollary 2.6], using Lemma 5.1 and the fluctuation scale estimate in (4-17a) the proof of the following lemma is immediate.

Lemma 8.2. *Let $\tilde{r}(t) = \tilde{x}(t), \tilde{y}(t)$. There exists a fixed small $\epsilon > 0$ such that for each $1 \le |i| \le \epsilon N$ we have*

$$\sup_{t_* \le t \le t_1} |\tilde{r}_i(t) - \hat{\gamma}_{r,i}(t)| \le N^{\xi} \eta_{\mathrm{f}}^{\rho_{r,t}}(\gamma_{r,i}(t)) \tag{8-4}$$

for any $\xi > 0$ with very high probability, where we recall that the behaviour of $\eta_{\mathrm{f}}^{\rho_{r,t}}(\mathfrak{e}_{r,t}^+ + \hat{\gamma}_{r,\pm i}(t))$, with $r = x, y$, is given by (4-17b).

In order to prove Proposition 8.1, by Lemma 5.1 and (4-14b), it is enough to prove the following:

Proposition 8.3. *For $t_1 \ge t_*$ we have, with very high probability, that*

$$|(x_i(t_1) - \tilde{\mathfrak{m}}_{x,t_1}) - (y_i(t_1) - \tilde{\mathfrak{m}}_{y,t_1})| \le N^{-\frac{3}{4}-c} \tag{8-5}$$

for some small constant $c > 0$ and for any $1 \le |i| \le N^{\omega_1}$.

The remaining part of this section is devoted to the proof of Proposition 8.3. We start with some preparatory lemmas. We recall the definition of the interpolated quantiles given in Section 5,

$$\bar{\gamma}_i(t) := \alpha \hat{\gamma}_{x,i}(t) + (1-\alpha)\hat{\gamma}_{y,i}(t) \tag{8-6}$$

for all $\alpha \in [0, 1]$ and $t_* \le t \le 2t_*$, as well as

$$\bar{\mathfrak{m}}_t := \alpha \tilde{\mathfrak{m}}_{x,t} + (1-\alpha)\tilde{\mathfrak{m}}_{y,t}$$

for all $\alpha \in [0, 1]$ and $t_* \le t \le 2t_*$. Furthermore by $\bar{\rho}_t$ from (5-12) we denote the interpolated density between $\rho_{x,t}$ and $\rho_{y,t}$ and by $\bar{m}_t$ its Stieltjes transform.

We now define the process $\tilde{z}_i(t, \alpha)$ whose initial data are given by the linear interpolation of $\tilde{x}(0)$ and $\tilde{y}(0)$. Analogously to the small gap case, we define the function $\Psi_\alpha(t)$, for $t_* \le t \le 2t_*$, that represents the correct shift of the process $\tilde{z}(t, \alpha)$, in order to compensate the discrepancy of our choice of the interpolation for $\bar{\rho}_t$ with respect to the semicircular flow evolution of the density $\bar{\rho}_0$.

Analogously to the edge case, see (6-5)–(6-11), we define $h(t, \alpha)$ with the following properties:

$$h(t,\alpha) = \alpha \Re[m_{x,t}(\tilde{\mathfrak{m}}_{x,t})] + (1-\alpha)\Re[m_{y,t}(\tilde{\mathfrak{m}}_{y,t})] - \Re[\bar{m}_t(\bar{\mathfrak{m}}_t + \mathrm{i}N^{-100})] + \mathcal{O}(N^{-1}) \tag{8-7}$$

and $h(t, 0) = h(t, 1) = 0$. Then, similarly to the edge case, we define

$$\Psi_\alpha(t) := -\alpha \frac{\mathrm{d}}{\mathrm{d}t}[m_{x,t}(\tilde{\mathfrak{m}}_{x,t})] - (1-\alpha)\frac{\mathrm{d}}{\mathrm{d}t}[m_{y,t}(\tilde{\mathfrak{m}}_{y,t})] - h(t,\alpha). \tag{8-8}$$

In particular, by our definition of $h(t, \alpha)$ in (8-7) it follows that $\Psi_0(t) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{y,t}$, $\Psi_1(t) = \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{x,t}$ and

$$\Psi_\alpha(t) = \Re[\bar{m}_t(\bar{\mathfrak{m}}_t)] + \mathcal{O}(N^{-\frac{1}{2}+\omega_1}). \tag{8-9}$$

Note that the error in (8-9) is somewhat weaker than in the analogous equation (6-16) due to the additional error in (4-14d) compared with (4-14e).

More precisely, the process $\tilde{z}(t,\alpha)$ is defined by

$$\mathrm{d}\tilde{z}_i(t,\alpha)=\sqrt{\frac{2}{N}}\,\mathrm{d}B_i+\biggl[\frac{1}{N}\sum_{j\neq i}\frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)}+\Psi_\alpha(t)\biggr]\mathrm{d}t, \tag{8-10}$$

with initial data

$$\tilde{z}_i(t_*,\alpha):=\alpha\tilde{x}_i(t_*)+(1-\alpha)\tilde{y}_i(t_*), \tag{8-11}$$

for all $1\le|i|\le N$ and for all $\alpha\in[0,1]$.

We recall that $\omega_1\ll\omega_\ell\ll\omega_A\ll 1$ and that $i_*=N^{\frac{1}{2}}+C_*\omega_1$ with some large constant C_*.

Next, we define the analogues of $\mathcal{J}_z(t)$ and $\mathcal{I}_{z,i}(t)$ for the small minimum by (6-95) and (6-96) using the definition in (8-6) for the quantiles. Then, for each $t_*\le t\le t_1$, we define the short-range approximation $\hat{z}_i(t,\alpha)$ of $\tilde{z}(t,\alpha)$ by the following SDE:

For $|i|>\frac{1}{2}i_*$ we let

$$\mathrm{d}\hat{z}_i(t,\alpha)=\sqrt{\frac{2}{N}}\,\mathrm{d}B_i+\biggl[\frac{1}{N}\sum_j^{\mathcal{A},(i)}\frac{1}{\hat{z}_i(t,\alpha)-\hat{z}_j(t,\alpha)}+\frac{1}{N}\sum_j^{\mathcal{A}^c,(i)}\frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)}+\Psi_\alpha(t)\biggr]\mathrm{d}t, \tag{8-12}$$

for $|i|\le N^{\omega_A}$

$$\mathrm{d}\hat{z}_i(t,\alpha)=\sqrt{\frac{2}{N}}\,\mathrm{d}B_i+\biggl[\frac{1}{N}\sum_j^{\mathcal{A},(i)}\frac{1}{\hat{z}_i(t,\alpha)-\hat{z}_j(t,\alpha)}+\int_{\mathcal{I}_{y,i}(t)^c}\frac{\rho_{y,t}(E+\tilde{\mathfrak{m}}_{y,t}^{+})}{\hat{z}_i(t,\alpha)-E}\,\mathrm{d}E\biggr]\mathrm{d}t-\Bigl(\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{r,t}\Bigr)\mathrm{d}t, \tag{8-13}$$

and for $N^{\omega_A}<|i|\le\frac{1}{2}i_*$

$$\begin{aligned}\mathrm{d}\hat{z}_i(t,\alpha)=\sqrt{\frac{2}{N}}\,\mathrm{d}B_i+\biggl[\frac{1}{N}\sum_j^{\mathcal{A},(i)}\frac{1}{\hat{z}_i(t,\alpha)-\hat{z}_j(t,\alpha)}+\int_{\mathcal{I}_{z,i}(t)^c\cap\mathcal{J}_z(t)}\frac{\bar{\rho}_t(E+\bar{\mathfrak{m}}_t^{+})}{\hat{z}_i(t,\alpha)-E}\,\mathrm{d}E\\ +\sum_{|j|\ge\frac{3}{4}i_*}\frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)}+\Psi_\alpha(t)\biggr]\mathrm{d}t,\end{aligned} \tag{8-14}$$

with initial data

$$\hat{z}_i(t_*,\alpha):=\tilde{z}_i(t_*,\alpha). \tag{8-15}$$

Next, by Lemma C.2, by the optimal rigidity in (8-4) for $\tilde{x}(t)$ and $\tilde{y}(t)$, the next lemma follows immediately.

Lemma 8.4. *For $\alpha=0$ and $\alpha=1$, with very high probability, we have*

$$\sup_{1\le|i|\le N}\sup_{t_*\le t\le t_1}|\tilde{z}_i(t,\alpha)-\hat{z}_i(t,\alpha)|\lesssim\frac{N^\xi}{N^{\frac{3}{4}}}\biggl(\frac{N^{\omega_1}}{N^{3\omega_\ell}}+\frac{N^{C\omega_1}}{N^{\frac{1}{24}}}\biggr) \tag{8-16}$$

for any $\xi>0$ and $C>0$ a large universal constant.

In order to proceed with the heat-kernel estimates we need an optimal i-dependent rigidity for $\hat{z}_i(t,\alpha)$ for $1 \le |i| \le N^{4\omega_\ell+\delta}$ for some $0 < \delta < C\omega_\ell$. In particular, analogously to Proposition 6.12 we have:

Proposition 8.5. *Fix any* $\alpha \in [0,1]$. *There exists a small fixed* $0 < \delta_1 < C\omega_\ell$, *for some constant* $C > 0$, *such that*

$$\sup_{t_* \le t \le 2t_*} |\hat{z}_i(t,\alpha) - \bar{\gamma}_i(t)| \lesssim \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}, \quad 1 \le |i| \le N^{4\omega_\ell+\delta_1}, \tag{8-17}$$

for any $\xi > 0$ *with very high probability.*

Proof. We can adapt the arguments in Section 6 to the case of the small minimum, $t \ge t_*$, in a straightforward way. In Section 6, as the main input, we used the precise estimates on the density $\rho_{r,t}$ (4-6b), (4-20), on the quantiles $\hat{\gamma}_{r,i}(t)$ (4-13a), on the quantile gaps (4-18), on the fluctuation scale (4-17a) and on the Stieltjes transform (4-22a), all formulated for the small gap case, $0 \le t \le t_*$. In the small minimum case, $t \ge t_*$, the corresponding estimates are all available in Section 4; see (4-6d), (4-21), (4-13b), (4-19), (4-17b) and (4-22b), respectively. In fact, the semicircular flow is more regular after the cusp formation; see, e.g., the better (larger) exponent in the $(t - t_*)$ error terms when comparing (4-6b) with (4-6d). This makes handling the small minimum case easier. The most critical part in Section 6 is the estimate of the forcing term (Proposition 6.7), where the derivative of the density (4-7a) was heavily used. The main mechanism of this proof is the delicate cancellation between the contributions to S_2 from the intervals $[\gamma_{i-n-1}, \gamma_{i-n}]$ and $[\gamma_{i+n-1}, \gamma_{i+n}]$; see (6-83). This cancellation takes place away from the edge. The proof is divided into two cases: the so-called "edge regime", where the gap length Δ is relatively large, and the "cusp regime", where Δ is small or zero. The adaptation of this argument to the small minimum case, $t \ge t_*$, will be identical to the proof for the small gap case in the cusp regime. In this regime the derivative bound (4-7a) is used only in the form $|\rho'| \lesssim \rho^{-2}$, which is available in the small minimum case, $t \ge t_*$, as well; see (4-8a). This proves Proposition 6.7 for $t \ge t_*$. The rest of the argument is identical to the proof in the small minimum case up to obvious notational changes; the details are left to the reader. □

Let us define $u_i(t,\alpha) := \partial_\alpha \hat{z}_i(t,\alpha)$ for $t_* \le t \le 2t_*$. In particular, u is a solution of the equation

$$\partial_t u = \mathcal{L}u + \zeta^{(0)} \tag{8-18}$$

with initial condition $u(t_*,\alpha) = \tilde{x}(t_*) - \tilde{y}(t_*)$ from (8-11). The error term $\zeta^{(0)}$ is defined analogously to (6-108)–(6-109) but replacing Φ_α and $\bar{\mathfrak{e}}_t^+$ with Ψ_α and $\tilde{\mathfrak{m}}_t$, respectively. Note that this error term is nonzero only for $|i| \ge N^{\omega_A}$ and for any i we have $|\zeta_i^{(0)}| \le N^C$ with very high probability, for some large $C > 0$. Furthermore, $\mathcal{L} = \mathcal{B} + \mathcal{V}$ is defined as in (6-106)–(6-107) replacing $\mathfrak{e}_{y,t}^+$ and $\bar{\mathfrak{e}}_t^+$ by $\tilde{\mathfrak{m}}_{y,t}$ and $\bar{\mathfrak{m}}_t$, respectively. In the following by $\mathcal{U}^{\mathcal{L}}$ we denote the propagator of the operator $\mathcal{L}$.

Let $0 < \delta_v < \frac{1}{2}\delta_4$, with δ_4 defined in Lemma 7.6. Define $v_i = v_i(t,\alpha)$ to be the solution of

$$\partial_t v = \mathcal{L}v, \quad v_i(t_*,\alpha) = u_i(t_*,\alpha)\mathbf{1}_{\{|i| \le N^{4\omega_\ell+\delta_v}\}}. \tag{8-19}$$

By the finite speed of propagation estimate in Lemma B.3, similarly to the proof of Lemma 7.7, we immediately obtain the following:

Lemma 8.6. *Let u be the solution of* (8-18) *and v defined by* (8-19); *then we have*

$$\sup_{t_* \le t \le 2t_*} \sup_{1 \le |i| \le \ell^4} |u_i(t) - v_i(t)| \le N^{-100} \tag{8-20}$$

with very high probability.

Collecting all the previous lemmas we conclude this section with the proof of Proposition 8.3.

Proof of Proposition 8.3. We consider only the $i = 1$ case. By Lemmas 5.1 and 8.4 we have

$$\begin{aligned} |(x_1(t_1) - \tilde{\mathfrak{m}}_{x,t_1}) - (y_1(t_1) - \tilde{\mathfrak{m}}_{y,t_1})| &\le |\tilde{x}_1(t_1) - \hat{x}_1(t_1)| + |\hat{x}_1(t_1) - \hat{y}_1(t_1)| + |\hat{y}_1(t_1) - \tilde{y}_1(t_1)| \\ &\le |\hat{x}_1(t_1) - \hat{y}_1(t_1)| + \frac{1}{N^{\frac{3}{4}+c}} \end{aligned}$$

with very high probability. Since $u(t,\alpha) = \partial_\alpha \hat{z}(t,\alpha)$, using $\hat{x}_1(t_1) - \hat{y}_1(t_1) = \int_0^1 u(t_1,\alpha)\,\mathrm{d}\alpha$ and Lemma 8.6 it will be sufficient to estimate $\int_0^1 |v_1(t_1,\alpha)|\,\mathrm{d}\alpha$. By rigidity from (8-4), we have

$$|v_i(t_*,\alpha)| = |u_i(t_*,\alpha)| = |\tilde{y}_i(t_*) - \tilde{x}_i(t_*)| \lesssim \frac{N^\xi}{N^{\frac{3}{4}}|i|^{\frac{1}{4}}}$$

for any $1 \le |i| \le N^{4\omega_\ell + \delta_v}$; hence

$$\|v(t_*,\alpha)\|_5 \lesssim \frac{N^\xi}{N^{\frac{3}{4}}}$$

for any $\xi > 0$ with very high probability.

Finally, using the heat kernel estimate in (7-22) for $\mathcal{U}^{\mathcal{L}}(0,t)$ for $t_* \le t \le 2t_*$, we conclude, after a Markov inequality as in (6-45)–(6-46),

$$\int_0^1 |v_1(t_1,\alpha)|\,\mathrm{d}\alpha \lesssim \frac{N^\xi}{N^{\frac{3}{4}} N^{\frac{4}{15}\omega_1}} \tag{8-21}$$

with very high probability. □

Appendix A: Proof of Theorem 2.4

We now briefly outline the changes required for the proof of Theorem 2.4 compared to the proof of Theorem 2.2. We first note that for $0 \le \tau_1 \le \cdots \le \tau_k \lesssim N^{-\frac{1}{2}}$ in distribution $(H^{(\tau_1)}, \dots, H^{(\tau_k)})$ agrees with

$$(H + \sqrt{\tau_1}U_1, H + \sqrt{\tau_1}U_1 + \sqrt{\tau_2 - \tau_1}U_2, \dots, H + \sqrt{\tau_1}U_1 + \cdots + \sqrt{\tau_k - \tau_{k-1}}U_k), \tag{A-1}$$

where $U_1, \dots, U_k$ are independent GOE matrices. Next, we claim and prove later by Green's function comparison that the time-dependent k-point correlation function of (A-1) asymptotically agrees with the one of

$$(\widetilde{H}_t + \sqrt{\tau_1}U_1, \widetilde{H}_t + \sqrt{\tau_1}U_1 + \sqrt{\tau_2 - \tau_1}U_2, \dots, \widetilde{H}_t + \sqrt{\tau_1}U_1 + \cdots + \sqrt{\tau_k - \tau_{k-1}}U_k), \tag{A-2}$$

and thereby also with the one of

$$(H_t + \sqrt{ct}U + \sqrt{\tau_1}U_1, H_t + \sqrt{ct}U + \sqrt{\tau_1}U_1 + \sqrt{\tau_2 - \tau_1}U_2, \dots, \\ H_t + \sqrt{ct}U + \sqrt{\tau_1}U_1 + \cdots + \sqrt{\tau_k - \tau_{k-1}}U_k) \quad \text{(A-3)}$$

for any fixed $t \le N^{-\frac{1}{4}-\epsilon}$, where $\widetilde{H}_t$ and H_t are constructed as in Section 3 (see (3-3)). Finally, we notice that the joint eigenvalue distribution of the matrices in (A-3) is precisely given by the joint distribution of

$$(\lambda_i(ct + \tau_1), \dots, \lambda_i(ct + \tau_k),\ i \in [N])$$

where $\lambda_i(s)$ are the eigenvalues evolved according to the DBM

$$\mathrm{d}\lambda_i(s) = \sqrt{\frac{2}{N}}\,\mathrm{d}B_i + \sum_{j \ne i} \frac{1}{\lambda_i(s) - \lambda_j(s)}\,\mathrm{d}s, \quad \lambda_i(0) = \lambda_i(H_t). \quad \text{(A-4)}$$

The high probability control on the eigenvalues evolved according to (A-4) in Propositions 7.1 and 8.1 allows us to simultaneously compare eigenvalues at different times with those of the Gaussian reference ensemble automatically.

In order to establish Theorem 2.4 it thus only remains to argue that the k-point functions of (A-1) and (A-2) are asymptotically equal. For the sake of this argument we consider only the randomness in H and the condition on the randomness in $U_1, \dots, U_k$. Then the OU-flow

$$\mathrm{d}\widetilde{H}_s' = -\tfrac{1}{2}\big(\widetilde{H}_s' - A - \sqrt{\tau_1}U_1 - \cdots - \sqrt{\tau_l - \tau_{l-1}}U_l\big)\,\mathrm{d}s + \Sigma^{\frac{1}{2}}[\mathrm{d}\mathfrak{B}_s],$$

with initial conditions

$$\widetilde{H}_0' = H + \sqrt{\tau_1}U_1 + \cdots + \sqrt{\tau_l - \tau_{l-1}}U_l,$$

for fixed $U_1, \dots, U_l$ is given by

$$\widetilde{H}_s' = \widetilde{H}_s + \sqrt{\tau_1}U_1 + \cdots + \sqrt{\tau_l - \tau_{l-1}}U_l;$$

i.e., we view $\sqrt{\tau_1}U_1 + \cdots + \sqrt{\tau_l - \tau_{l-1}}U_l$ as an additional expectation matrix. Thus we can appeal to the standard Green's function comparison technique already used in Section 3 to compare the k-point functions of (A-1) and (A-2). Here we can follow the standard resolvent expansion argument from [Erdős et al. 2018, equation (116)] and note that the proof therein verbatim also allows us to compare products of traces of resolvents with differing expectations. Finally we then take the $\mathbb{E}_{U_1} \dots \mathbb{E}_{U_k}$ expectation to conclude that not only the conditioned k-point functions of (A-1) and (A-2) asymptotically agree, but also the k-point functions themselves.

Appendix B: Finite speed of propagation estimate

In this section we prove a finite speed of propagation estimate for the time evolution of the α-derivative of the short-range dynamics defined in (6-97)–(6-99). It is an adjustment to the analogous proof of Lemma 4.1 in [Landon and Yau 2017]. For concreteness, we present the proof for the propagator $\mathcal{U}^{\mathcal{L}}$, where $\mathcal{L} = \mathcal{B} + \mathcal{V}$ is defined in (6-105)–(6-107). The point is that once the dynamics is localised, i.e., the range of the interaction term $\mathcal{B}$ is restricted to a local scale $|i-j| \le |j_+(i)-j_-(i)|$, with $|j_+(i)-j_-(i)| \gtrsim N^{4\omega_\ell} =: L$, and the time is also restricted, $0 \le t \le 2t_* \lesssim N^{-\frac{1}{2}+\omega_1}$, the propagation cannot go beyond a scale that is much bigger than the interaction scale. This mechanism is very general and will also be used in a slightly different (simpler) setup of Lemma 6.5 and Proposition 6.8 where the interaction scale is much bigger $L \sim \sqrt{N}$. We will give the necessary changes for the proof of Lemma 6.5 and Proposition 6.8 at the end of this section.

Lemma B.1. *Let* $\hat{z}(t) = \hat{z}(t,\alpha)$ *be the solution to the short-range dynamics* (6-97)–(6-99) *with* $i_* = N^{\frac{1}{2}+C_*\omega_1}$, *exponents* $\omega_1 \ll \omega_\ell \ll \omega_A \ll 1$ *and propagator* $\mathcal{L} = \mathcal{B} + \mathcal{V}$ *from* (6-105)–(6-107). *Let us assume that*

$$\sup_{0 \le t \le t_*} |\hat{z}_i(t) - \bar{\gamma}_i(t)| \le \frac{N^{C\omega_1}}{N^{\frac{3}{4}}}, \quad 1 \le |i| \le i_*, \tag{B-1}$$

where $\bar{\gamma}_i(t)$ *are the quantiles from* (5-10). *Then, there exists a constant* $C' > 0$ *such that for any* $0 < \delta < C'\omega_\ell$, $|a| \ge LN^\delta$ *and* $|b| \le \frac{3}{4}LN^\delta$, *for any fixed* $0 \le s \le t_*$, *we have*

$$\sup_{s \le t \le t_*} \mathcal{U}^{\mathcal{L}}_{ab}(s,t) + \mathcal{U}^{\mathcal{L}}_{ba}(s,t) \le N^{-D} \tag{B-2}$$

for any $D > 0$, *with very high probability. The same result holds for the short-range dynamics after the cusp defined in* (8-18) *for* $t_* \le s \le 2t_*$.

Proof of Lemma B.1. For concreteness we assume that $0 \le s \le t \le t_*$, i.e., we are in the small gap regime. For $t_* \le s \le t \le 2t_*$ the proof is analogous using the definition (8-6) for the $\bar{\gamma}_i(t)$, the definition of the short-range approximation in (8-12)–(8-15) for the $\hat{z}_i(t,\alpha)$ and replacing $\bar{\mathfrak{e}}^+_t$ by $\bar{\mathfrak{m}}_t$. With these adjustments the proof follows in the same way except for (B-25) below, where we have to use the estimates in (4-22b) instead of (4-22a).

First we consider the $s = 0$ case; then in Lemma B.3 below we extend the proof for all $0 \le s \le t$. Let $\psi(x)$ be an even 1-Lipschitz real function; i.e., $\psi(x) = \psi(-x)$, $\|\psi'\|_\infty \le 1$ such that

$$\psi(x) = |x| \quad \text{for } |x| \le \frac{L^{\frac{3}{4}}N^{\frac{3}{4}\delta}}{N^{\frac{3}{4}}}, \qquad \psi'(x) = 0 \quad \text{for } |x| \ge 2\frac{L^{\frac{3}{4}}N^{\frac{3}{4}\delta}}{N^{\frac{3}{4}}}. \tag{B-3}$$

and

$$\|\psi''\|_\infty \lesssim \frac{N^{\frac{3}{4}}}{L^{\frac{3}{4}}N^{\frac{3}{4}\delta}}. \tag{B-4}$$

We consider a solution of the equation

$$\partial_t f = \mathcal{L}f, \quad 0 \le t \le t_*,$$

with some discrete Dirac delta initial condition $f_i(0) = \delta_{ip_*}$ at p_* for any $|p_*| \ge N^{4\omega_\ell}N^\delta$. For concreteness, assume $p_* > 0$ and set $p := N^{4\omega_\ell}N^\delta$. Define

$$\phi_i = \phi_i(t,\alpha) := e^{\frac{1}{2}\nu\psi(\hat{z}_i(t,\alpha)-\bar{\gamma}_p(t))}, \quad m_i = m_i(t,\alpha) := f_i(t,\alpha)\phi_i(t,\alpha), \quad \nu = \frac{N^{\frac{3}{4}}}{L^{\frac{3}{4}}N^{\delta'}}, \tag{B-5}$$

with some $\delta' \geq \frac{1}{2}\delta$ to be chosen later. Let $\hat{z}_i = \hat{z}_i(t,\alpha)$ and set

$$F(t) := \sum_i f_i^2 e^{\nu\psi(\hat{z}_i - \bar{\gamma}_p(t))} = \sum_i m_i^2. \tag{B-6}$$

Since

$$2\sum_i f_i(\mathcal{B}f)_i\phi_i^2 = \sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij}(m_i - m_j)^2 - \sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij} m_i m_j \left(\frac{\phi_i}{\phi_j} + \frac{\phi_j}{\phi_i} - 2\right),$$

using Itô's formula, we conclude that

$$\mathrm{d}F = \sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij}(m_i - m_j)^2\,\mathrm{d}t + 2\sum_i \mathcal{V}_i m_i^2\,\mathrm{d}t \tag{B-7}$$

$$- \sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij} m_i m_j \left(\frac{\phi_i}{\phi_j} + \frac{\phi_j}{\phi_i} - 2\right)\mathrm{d}t \tag{B-8}$$

$$+ \sum_i \nu m_i^2 \psi'(\hat{z}_i - \bar{\gamma}_p)\,\mathrm{d}(\hat{z}_i - \bar{\gamma}_p) \tag{B-9}$$

$$+ \sum_i m_i^2 \left(\frac{\nu^2}{N}\psi'(\hat{z}_i - \bar{\gamma}_p)^2 + \frac{\nu}{N}\psi''(\hat{z}_i - \bar{\gamma}_p)\right)\mathrm{d}t. \tag{B-10}$$

Let $\tau_1 \leq t_*$ be the first time such that $F \geq 5$ and let τ_2 be the stopping time such that the estimate (B-1) holds with $t \leq \tau_2$ instead of $t \leq t_*$; the condition (B-1) then says that $\tau_2 = t_*$ with very high probability. Define $\tau := \tau_1 \wedge \tau_2 \wedge t_*$; our goal is to show that $\tau = t_*$. In the following we assume $t \leq \tau$.

Now we estimate the terms in (B-7)–(B-10) one by one. We start with (B-8). Note that the rigidity scale $N^{-\frac{3}{4}+C\omega_1}$ in (B-1) is much smaller than $N^{-\frac{3}{4}(1-\delta)+3\omega_\ell}$, the range of the support of ψ', which, in turn, is comparable with $|\bar{\gamma}_i - \bar{\gamma}_p| \gtrsim (p/N)^{\frac{3}{4}}$ for any $i \geq 2p = 2LN^\delta$. Therefore $\psi'(\hat{z}_i - \bar{\gamma}_p) = 0$ unless $|i| \lesssim LN^\delta$. Moreover, if $|i| \lesssim LN^\delta$ and $(i,j) \in \mathcal{A}$, then $|j| \lesssim LN^\delta$. Hence, the nonzero terms in the sum in (B-8) have $|i|, |j| \lesssim N^{4\omega_\ell+\delta}$. By (B-1) and $C\omega_1 \ll \omega_\ell$, for such terms we have

$$|\hat{z}_i - \hat{z}_j| \lesssim \frac{|i-j|}{N^{\frac{3}{4}}\min\{|i|,|j|\}^{\frac{1}{4}}} + \frac{N^{C\omega_1}}{N^{\frac{3}{4}}} \lesssim \frac{L^{\frac{3}{4}}N^{\frac{1}{2}\delta}}{N^{\frac{3}{4}}}. \tag{B-11}$$

Note that $\nu|\hat{z}_i - \hat{z}_j| \lesssim 1$. Therefore, by Taylor expanding in the exponent, we have

$$\left|\frac{\phi_i}{\phi_j} + \frac{\phi_j}{\phi_i} - 2\right| = \left(e^{\frac{1}{2}\nu(\psi(\hat{z}_j - \bar{\gamma}_p) - \psi(\hat{z}_i - \bar{\gamma}_p))} - e^{\frac{1}{2}\nu(\psi(\hat{z}_i - \bar{\gamma}_p) - \psi(\hat{z}_j - \bar{\gamma}_p))}\right)^2 \lesssim \nu^2|\psi(\hat{z}_i - \bar{\gamma}_p) - \psi(\hat{z}_j - \bar{\gamma}_p)|^2,$$

and thus

$$\left|\mathcal{B}_{ij}\left(\frac{\phi_i}{\phi_j} + \frac{\phi_j}{\phi_i} - 2\right)\right| \lesssim \nu^2 \frac{|\psi(\hat{z}_i - \bar{\gamma}_p) - \psi(\hat{z}_j - \bar{\gamma}_p)|^2}{N(\hat{z}_i - \hat{z}_j)^2} \lesssim \frac{\nu^2}{N}, \tag{B-12}$$

where in the last inequality we used that ψ is Lipschitz continuous. Hence we conclude the estimate of (B-8) as

$$\left|\sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij} m_i m_j \left(\frac{\phi_i}{\phi_j} + \frac{\phi_j}{\phi_i} - 2\right)\right| \lesssim \frac{\nu^2}{N}\sum_i m_i^2 \sum_j^{\mathcal{A},(i)} \mathbf{1}_{\{\phi_j \neq \phi_i\}} \lesssim \frac{\nu^2 L N^{\frac{3}{4}\delta}}{N} F(t), \tag{B-13}$$

since the number of j's in the summation is at most

$$|j_+(i) - j_-(i)| \le \ell^4 + \ell |i|^{\frac{3}{4}} \le L N^{\frac{3}{4}\delta}. \tag{B-14}$$

By (B-4) and since $|\psi'(x)| \le 1$, (B-10) is bounded as follows:

$$\Big| \sum_i m_i^2 \Big(\frac{\nu^2}{N} \psi'(\hat{z}_i - \bar{\gamma}_p)^2 + \frac{\nu}{N} \psi''(\hat{z}_i - \bar{\gamma}_p) \Big) \Big| \lesssim \Big(\frac{\nu^2}{N} + \frac{\nu}{N^{\frac{1}{4}} L^{\frac{3}{4}} N^{\frac{3}{4}\delta}} \Big) F(t). \tag{B-15}$$

The next step is to get a bound for (B-9). Since $\psi'(\hat{z}_i - \bar{\gamma}_p) = 0$ unless $|i| \lesssim N^{4\omega_\ell + \delta} \ll N^{\omega_A}$, choosing $C > 0$ such that $(4 + C)\omega_\ell < \omega_A$ and using (6-98) we get

$$\mathrm{d}(\hat{z}_i(t) - \bar{\gamma}_p(t)) = \sqrt{\frac{2}{N}} \, \mathrm{d}B_i + \frac{1}{N} \sum_j^{\mathcal{A},(i)} \frac{1}{\hat{z}_i(t) - \hat{z}_j(t)} + Q_i(t), \tag{B-16}$$

with

$$Q_i(t) := \int_{\mathcal{I}_{y,i}(t)^c} \frac{\rho_{y,t}(E + \mathfrak{e}_{y,t}^+)}{\hat{z}_i(t) - E} \, \mathrm{d}E + \Re[m_{y,t}(\mathfrak{e}_{y,t}^+)] + \alpha \big(\Re[m_{x,t}(\hat{\gamma}_{x,p}(t) + \mathfrak{e}_{x,t}^+) - m_{x,t}(\mathfrak{e}_{x,t}^+)] \big) \\ + (1 - \alpha) \big(\Re[m_{y,t}(\hat{\gamma}_{y,p}(t) + \mathfrak{e}_{y,t}^+) - m_{y,t}(\mathfrak{e}_{y,t}^+)] \big). \tag{B-17}$$

We insert (B-16) into (B-9) and estimate all three terms separately in the regime $|i| \lesssim L N^\delta$. For the stochastic differential, by the definition of $\tau \le t_*$ and the Burkholder–Davis–Gundy inequality we have

$$\sup_{0 \le t \le \tau} \int_0^t \sqrt{\frac{2}{N}} \nu \sum_i m_i^2 \psi'(\hat{z}_i - \bar{\gamma}_p) \, \mathrm{d}B_i \le N^{\epsilon'} \frac{\nu}{\sqrt{N}} \sqrt{t_*} \sup_{0 \le t \le \tau} F(t) \lesssim \nu N^{\epsilon'} N^{-\frac{3}{4} + \frac{1}{2}\omega_1} \tag{B-18}$$

for any $\epsilon' > 0$, with very high probability. In (B-18) we used that $\tau \le t_* \sim N^{-\frac{1}{2} + \omega_1}$, and that, by the definition of τ, $F(t)$ is bounded for all $0 \le t \le \tau$.

The contribution of the second term in (B-16) to (B-9) is written, after symmetrisation, as

$$\frac{\nu}{N} \sum_{(i,j) \in \mathcal{A}} \frac{\psi'(\hat{z}_i - \bar{\gamma}_p) m_i^2}{\hat{z}_j - \hat{z}_i} \\ = \frac{\nu}{2N} \sum_{(i,j) \in \mathcal{A}} \frac{\psi'(\hat{z}_i - \bar{\gamma}_p)(m_i^2 - m_j^2)}{\hat{z}_j - \hat{z}_i} + \frac{\nu}{2N} \sum_{(i,j) \in \mathcal{A}} m_i^2 \frac{\psi'(\hat{z}_i - \bar{\gamma}_p) - \psi'(\hat{z}_j - \bar{\gamma}_p)}{\hat{z}_j - \hat{z}_i}. \tag{B-19}$$

Using (B-4) and (B-14), the second sum in (B-19) is bounded by

$$\Big| \frac{\nu}{2N} \sum_{(i,j) \in \mathcal{A}} m_i^2 \frac{\psi'(\hat{z}_i - \bar{\gamma}_p) - \psi'(\hat{z}_j - \bar{\gamma}_p)}{\hat{z}_j - \hat{z}_i} \Big| \lesssim \frac{\nu}{N^{\frac{1}{4}} L^{\frac{3}{4}} N^{\frac{3}{4}\delta}} \sum_i m_i^2 \sum_j^{\mathcal{A},(i)} \mathbf{1}_{\{\psi'(\hat{z}_i - \bar{\gamma}_p) \ne \psi'(\hat{z}_j - \bar{\gamma}_p)\}} \\ \lesssim \frac{\nu L^{\frac{1}{4}}}{N^{\frac{1}{4}}} F. \tag{B-20}$$

Using $m_i^2 - m_j^2 = (m_i - m_j)(m_i + m_j)$ and the Schwarz inequality, the first sum in (B-19) is bounded as follows:

$$\frac{\nu}{2N} \sum_{(i,j)\in\mathcal{A}} \frac{\psi'(\hat{z}_i - \bar{\gamma}_p)(m_i^2 - m_j^2)}{\hat{z}_j - \hat{z}_i} \le -\frac{1}{100} \sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij}(m_i - m_j)^2 + \frac{C\nu^2}{2N} \sum_{(i,j)\in\mathcal{A}} \psi'(\hat{z}_i - \bar{\gamma}_p)^2 (m_i^2 + m_j^2). \quad \text{(B-21)}$$

The second sum in (B-21), using (B-14), is bounded by

$$\frac{C\nu^2}{2N} \sum_{(i,j)\in\mathcal{A}} \psi'(\hat{z}_i - \bar{\gamma}_p)(m_i^2 + m_j^2) \le \frac{C\nu^2 L N^{\frac{3}{4}\delta}}{2N} F; \quad \text{(B-22)}$$

hence we conclude that

$$\frac{\nu}{N} \sum_{(i,j)\in\mathcal{A}} \frac{\psi'(\hat{z}_i - \bar{\gamma}_p) m_i^2}{\hat{z}_j - \hat{z}_i} \le -\frac{1}{100} \sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij}(m_i - m_j)^2 + C\left(\frac{\nu L^{\frac{1}{4}}}{N^{\frac{1}{4}}} + \frac{\nu^2 L N^{\frac{3}{4}\delta}}{N}\right) F. \quad \text{(B-23)}$$

Note that the first term on the right-hand side of (B-23) can be incorporated in the first, dissipative term in (B-7).

To conclude the estimate of (B-9) we write the third term in (B-16) as

$$\begin{aligned} Q_i(t) = {} & \left(\int_{\mathcal{I}_{y,i}(t)^c} \frac{\rho_{y,t}(E + \mathfrak{e}_{y,t}^+)}{\hat{z}_i(t) - E} \, \mathrm{d}E + \Re[m_{y,t}(\bar{\gamma}_p(t) + \mathfrak{e}_{y,t}^+)] \right) \\ & + \alpha\big(\Re[m_{x,t}(\hat{\gamma}_{x,p}(t) + \mathfrak{e}_{x,t}^+) - m_{x,t}(\mathfrak{e}_{x,t}^+)] - \Re[m_{y,t}(\hat{\gamma}_{x,p}(t) + \mathfrak{e}_{y,t}^+) - m_{y,t}(\mathfrak{e}_{y,t}^+)]\big) \\ & + \alpha\big(\Re[m_{y,t}(\hat{\gamma}_{x,p}(t) + \mathfrak{e}_{y,t}^+)] - \Re[m_{y,t}(\bar{\gamma}_p(t) + \mathfrak{e}_{y,t}^+)]\big) \\ & + (1-\alpha)\big(\Re[m_{y,t}(\hat{\gamma}_{y,p}(t) + \mathfrak{e}_{y,t}^+)] - \Re[m_{y,t}(\bar{\gamma}_p(t) + \mathfrak{e}_{y,t}^+)\big) \\ =: {} & A_1 + A_2 + A_3 + A_4. \end{aligned} \quad \text{(B-24)}$$

Similarly to the estimates in (6-58), for A_2 we use (4-22a), while for A_3, A_4 we use (4-7b); then we use the asymptotic behaviour of $\hat{\gamma}_p, \bar{\gamma}_p$ by (4-13a) and $p = LN^\delta$ to conclude that

$$|A_2| + |A_3| + |A_4| \lesssim \frac{L^{\frac{1}{4}} N^{\frac{1}{4}\delta} N^{C\omega_1} \log N}{N^{\frac{1}{4}} N^{\frac{1}{6}}}. \quad \text{(B-25)}$$

We write the A_1-term as

$$A_1 = \int_{\mathcal{I}_{y,i}(t)^c} \frac{\bar{\gamma}_p(t) - \hat{z}_i(t)}{(\hat{z}_i(t) - E)(\bar{\gamma}_p(t) - E)} \rho_{y,t}(E + \mathfrak{e}_{y,t}^+) \, \mathrm{d}E + \int_{\mathcal{I}_{y,i}(t)} \frac{\rho_{y,t}(E + \mathfrak{e}_{y,t}^+)}{\bar{\gamma}_p(t) - E} \, \mathrm{d}E. \quad \text{(B-26)}$$

Since $i \le Cp$, we have $\rho_{y,t}(E + \mathfrak{e}_{y,t}^+) \le \rho_{y,t}(\bar{\gamma}_{Cp}(t) + \mathfrak{e}_{y,t}^+) \lesssim L^{\frac{1}{4}} N^{-\frac{1}{4} + \frac{1}{4}\delta}$ for any $E \in \mathcal{I}_{y,i}(t)$; the second term in (B-26) is bounded by $L^{\frac{1}{4}} N^{-\frac{1}{4} + \frac{1}{4}\delta} \log N$. In the first term in (B-26) we use that

$$|\hat{z}_i(t) - E| \ge |\bar{\gamma}_i(t) - E| - |\hat{z}_i(t) - \bar{\gamma}_i(t)| \gtrsim \bar{\gamma}_p(t)$$

for $E \notin \mathcal{I}_{y,i}(t)$, by rigidity (B-1) and by the fact that in the $i \le Cp$ regime $|\bar{\gamma}_i(t) - \bar{\gamma}_{i\pm j_\pm(i)}(t)| \gtrsim \bar{\gamma}_p(t) \gg N^{-\frac{3}{4}+C\omega_1}$ since $\omega_1 \ll \omega_\ell$ and $= LN^\delta = N^{4\omega_\ell+\omega_1}$.

We thus conclude that the first term in (B-26) is bounded by

$$|\hat{z}_i(t) - \bar{\gamma}_p(t)| \frac{\Im[m_{y,t}(\mathfrak{e}_{y,t}^+ + \mathrm{i}\bar{\gamma}_p(t))]}{\bar{\gamma}_p(t)} \lesssim \bar{\gamma}_p^{\frac{1}{3}} \lesssim L^{\frac{1}{4}} N^{-\frac{1}{4}+\frac{1}{4}\delta},$$

where we used again the rigidity (B-1). In summary, we have

$$|A_1| \lesssim L^{\frac{1}{4}} N^{-\frac{1}{4}+\frac{1}{4}\delta} \log N. \tag{B-27}$$

In particular (B-24)–(B-27) imply that

$$Q := \sup_{0 \le t \le t_*} \sup_{|i| \lesssim LN^\delta} |Q_i(t)| \lesssim L^{\frac{1}{4}} N^{-\frac{1}{4}+\frac{1}{4}\delta} \log N. \tag{B-28}$$

Collecting all the previous estimates using the choice of ν from (B-5) with $\delta' \ge \frac{1}{2}\delta$ and that F is bounded up to $t \le \tau$, we integrate (B-7)–(B-10) from 0 up to time $0 \le t \le t_*$ and conclude that

$$\begin{aligned} \sup_{0 \le t \le \tau} F(t) - F(0) &\lesssim \left(\frac{\nu^2 L N^{\frac{3}{4}\delta+\omega_1}}{N^{\frac{3}{2}}} + \frac{\nu L^{\frac{1}{4}} N^{\omega_1}}{N^{\frac{3}{4}}} + \frac{\nu Q N^{\omega_1}}{N^{\frac{1}{2}}} \right) \\ &\lesssim \frac{N^{\frac{3}{4}\delta+\omega_1}}{L^{\frac{1}{2}} N^{2\delta'}} + \frac{N^{\omega_1}}{L^{\frac{1}{2}} N^{\delta'}} + \frac{N^{\omega_1+\frac{1}{4}\delta}}{L^{\frac{1}{2}} N^{\delta'}} \log N \le 1 \end{aligned} \tag{B-29}$$

for large N and with very high probability, where we used the choice of ν (B-5) and that $\omega_1 \ll \omega_\ell$ in the last line. Since $F(0) = 1$, we get that $\tau = t_*$ with very high probability, and so

$$\sup_{0 \le t \le t_*} F(t) \le 5 \tag{B-30}$$

with very high probability.

Furthermore, since $p = LN^\delta$, if $i \le \frac{3}{4} LN^\delta$, choosing $\delta' = \frac{3}{4}\delta - \epsilon_1$, with $\epsilon_1 \le \frac{1}{4}\delta$, then by Proposition 6.3 we have

$$\nu \psi(\hat{z}_i(t) - \bar{\gamma}_p) = \nu |\hat{z}_i(t) - \bar{\gamma}_p| \gtrsim \nu \frac{|i-p|}{N^{\frac{3}{4}} |p|^{\frac{1}{4}}} \gtrsim \frac{N^{\frac{3}{4}\delta}}{N^{\delta'}} = N^{\epsilon_1}$$

with very high probability.

Note that (B-30) implies

$$f_i(t) \le 5 e^{-\frac{1}{2}\nu\psi(\hat{z}_i(t) - \bar{\gamma}_p)}.$$

Therefore, if $i \le \frac{3}{4} LN^\delta$ and $p_* \ge p$, then for each fixed $0 \le t \le t_*$ we have

$$\mathcal{U}^{\mathcal{L}}_{ip_*}(0, t) \le N^{-D} \tag{B-31}$$

for any $D > 0$ with very high probability. A similar estimate holds if i and p_* are negative or have opposite sign. This proves the estimate on the first term in (B-2) for any fixed s. The estimate for $\mathcal{U}^{\mathcal{L}}_{p_* i}(s, t)$ is analogous with initial condition $f = \delta_i$. This proves Lemma B.1. □

Next, we enhance this result to a bound uniform in $0 \le s \le t_*$. We first have:

Lemma B.2. *Let u be a solution of*

$$\partial_t u = \mathcal{L} u \tag{B-32}$$

with nonnegative initial condition $u_i(0) \ge 0$. Then, for each $0 \le t \le t_*$ we have

$$\frac{1}{2} \sum_i u_i(0) \le \sum_i u_i(t) \le \sum_i u_i(0) \tag{B-33}$$

with very high probability.

Proof. Since $\mathcal{U}^{\mathcal{L}}$ is a contraction semigroup the upper bound in (B-33) is trivial. Notice that $\partial_t \sum_i u_i = \sum_i \mathcal{V}_i u_i$. Thus the lower bound will follow once we prove $-\mathcal{V}_i \lesssim N^{\frac{1}{2}} L^{-\frac{1}{2}}$ with very high probability since $t_* N^{\frac{1}{2}} L^{-\frac{1}{2}}$ is much smaller than 1 by $\omega_1 \ll \omega_\ell$.

The estimate $-\mathcal{V}_i \lesssim N^{\frac{1}{2}} L^{-\frac{1}{2}}$ proceeds similarly to (B-26). Indeed, for $1 \le |i| < N^{\omega_A}$ we use $\rho_{y,t}(E + \mathfrak{e}_{y,t}^+) \lesssim |E|^{\frac{1}{3}}$ and that $|\hat{z}_i(t) - E| \sim |\bar{\gamma}_i(t) - E|$ by rigidity (B-1) and by the fact that

$$|j_+(i) - i|,\ |j_-(i) - i| \gtrsim N^{4\omega_\ell} + N^{\omega_\ell} |i|^{\frac{3}{4}}$$

is much bigger than the rigidity scale. Therefore, we have

$$\begin{aligned} -\mathcal{V}_i &= \int_{\mathcal{I}_{y,i}(t)^c} \frac{\rho_{r,t}(E + \mathfrak{e}_{r,t}^+)}{(\hat{z}_i(t) - E)^2} \, \mathrm{d}E \\ &\lesssim \int_{\mathcal{I}_{y,i}(t)^c} \frac{1}{|E - \bar{\gamma}_i(t)|^{\frac{5}{3}}} \, \mathrm{d}E + \int_{\mathcal{I}_{y,i}(t)^c} \frac{|\bar{\gamma}_i|^{\frac{1}{3}}}{(E - \bar{\gamma}_i(t))^2} \, \mathrm{d}E \lesssim \frac{N^{\frac{1}{2}}}{N^{2\omega_\ell}} = \frac{N^{\frac{1}{2}}}{L^{\frac{1}{2}}}. \end{aligned}$$

The estimate of $-\mathcal{V}_i$ for $N^{\omega_A} < |i| \le \frac{1}{2} i_*$ is similar. This concludes the proof of Lemma B.2. □

Finally we prove the following version of Lemma B.1 that is uniform in s:

Lemma B.3. *Under the same hypotheses of Lemma B.1, for any $\delta' > 0$ such that $\delta' < C' \omega_\ell$, with $C' > 0$ the constant defined in Lemma B.1, $|a| \le \frac{1}{2} L N^{\delta'}$ and $|b| \ge L N^{\delta'}$ we have*

$$\sup_{0 \le s \le t \le t_*} \mathcal{U}_{ab}^{\mathcal{L}}(s,t) + \mathcal{U}_{ba}^{\mathcal{L}}(s,t) \le N^{-D} \tag{B-34}$$

with very high probability. The same result holds for $t_* \le s \le t \le 2t_*$ as well.

Proof. By the semigroup property for any $0 \le s \le t \le t_*$ and any j we have

$$\mathcal{U}_{aj}^{\mathcal{L}}(0,t) \ge \mathcal{U}_{ab}^{\mathcal{L}}(s,t) \mathcal{U}_{bj}^{\mathcal{L}}(0,s). \tag{B-35}$$

Furthermore, by Lemma B.2 for the dual dynamics we have

$$\frac{1}{2} \sum_j u_j(0) \le \sum_j u_j(s) = \sum_i \sum_j (\mathcal{U}_{ji}^{\mathcal{L}}(0,s))^T u_i(0),$$

and so, by choosing $u(0) = \delta_b$ we conclude that

$$\sum_j \mathcal{U}^{\mathcal{L}}_{bj}(0,s) \geq \tfrac{1}{2} \quad \text{for all } 0 \leq s \leq t_*.$$

From the last inequality and since $\sup_{s\leq t_*} \mathcal{U}^{\mathcal{L}}_{bj}(0,s) \leq N^{-100}$ with very high probability for any $|j| \leq \frac{3}{4}LN^{\delta'}$ by Lemma B.1, it follows that there exists an $j_* = j_*(s)$, maybe depending on s, with $|j_*(s)| \geq \frac{3}{4}LN^{\delta'}$, such that $\mathcal{U}^{\mathcal{L}}_{bj_*(s)}(0,s) \geq 1/(4N)$. Furthermore, by the finite-speed propagation estimate in Lemma B.1 (this time with $|a| \geq \frac{3}{4}LN^{\delta}$ and $|b| \leq \frac{1}{2}LN^{\delta}$; note that its proof only used that $|a-b| \gtrsim LN^{\delta}$), we have

$$\sup_{t\leq t_*} \mathcal{U}^{\mathcal{L}}_{aj_*}(0,t) \leq N^{-D} \quad \text{for all } |j_*| \geq \tfrac{3}{4}LN^{\delta'}$$

with very high probability. Hence we get from (B-35) with $j = j_*(s)$ that $\sup_{s\leq t} \mathcal{U}^{\mathcal{L}}_{ab}(s,t) \lesssim N^{-D+1}$ with very high probability. The estimate for $\mathcal{U}^{\mathcal{L}}_{ba}(s,t)$ follows in a similar way. This concludes the proof of Lemma B.3. □

Finally, we prove Lemma 6.5 and Proposition 6.8 which are versions of Lemma B.3 but for the short-range approximation on scale $L = N^{\frac{1}{2}+C_1\omega_1}$ needed in Section 6C2.

Proof of Lemma 6.5. Choosing $L = N^{\frac{1}{2}+C_1\omega_1}$, the proof of Lemma B.1 is exactly the same except for the estimate of Q in (B-28), since, for any $\alpha \in [0,1]$, $Q_i(t)$ from (6-41) is now defined as

$$Q_i(t) := \frac{\beta}{N} \sum_{j:|j-i|>L} \frac{1}{\bar{\gamma}^*_i - \bar{\gamma}^*_j} + \frac{1-\beta}{N} \sum_{j:|j-i|>L} \frac{1}{\tilde{z}_i - \tilde{z}_j}\, \mathrm{d}t + \Phi_\alpha(t), \tag{B-36}$$

with $\Phi_\alpha(t)$ given in (6-15) instead of (B-17). Then Lemmas B.2 and B.3 follow exactly in the same way.

By (B-36) it easily follows that

$$Q := \sup_{0\leq t\leq t_*} \sup_{|i|\leq LN^{\delta'}} |Q_i(t)| \lesssim \log N. \tag{B-37}$$

Hence, by an estimate similar to (B-29), we conclude that

$$\begin{aligned}\sup_{0\leq t\leq \tau} F(t) - F(0) &\lesssim \left(\frac{\nu^2 L N^{\frac{3}{4}\delta+\omega_1}}{N^{\frac{3}{2}}} + \frac{\nu L^{\frac{1}{4}} N^{\omega_1}}{N^{\frac{3}{4}}} + \frac{\nu Q N^{\omega_1}}{N^{\frac{1}{2}}} \right) \\ &\lesssim \frac{N^{\frac{3}{4}\delta+\omega_1}}{L^{\frac{1}{2}}N^{\delta'}} + \frac{N^{\omega_1}}{L^{\frac{1}{2}}N^{\delta'}} + \frac{N^{\frac{3}{4}+\omega_1}}{L^{\frac{3}{4}}N^{\frac{1}{2}}N^{\delta'}} \log N \leq 1 \end{aligned} \tag{B-38}$$

with very high probability. Note that in the last inequality we used that $L = N^{\frac{1}{2}+C_1\omega_1}$. □

Proof of Proposition 6.8. This proof is almost identical to the previous one, except that $Q_i(t)$ is now defined from (6-52) as

$$Q_i(t) := \beta\left[\frac{1}{N} \sum_{j:|j-i|>L} \frac{1}{\bar{\gamma}^*_i - \bar{\gamma}^*_j} + \Phi(t)\right] + (1-\beta)\left[\frac{\mathrm{d}}{\mathrm{d}t}\bar{\gamma}^*_i(t) - \frac{1}{N} \sum_{j:|j-i|\leq L} \frac{1}{\bar{\gamma}^*_i - \bar{\gamma}^*_j}\right],$$

which satisfies the same bound (B-37). The rest of the proof is unchanged. □

Appendix C: Short-long approximation

In this section we estimate the difference of the solution of the DBM $\tilde{z}(t,\alpha)$ and its short-range approximation $\hat{z}(t,\alpha)$, closely following the proof of Lemma 3.7 in [Landon and Yau 2017] and adapting it to the more complicated cusp situation. In particular, in Section C1 we estimate $|\tilde{z}(t,\alpha)-\hat{z}(t,\alpha)|$ for $0 \le t \le t_*$, i.e., until the formation of an exact cusp; in Section C2, instead, we estimate $|\tilde{z}(t,\alpha)-\hat{z}(t,\alpha)|$ for $t_* < t \le 2t_*$, i.e., after the formation of a small minimum. The precision of this approximation depends on the rigidity bounds we put as a condition. We consider a two-scale rigidity assumption, a weaker rigidity valid for all indices and a stronger rigidity valid for $1 \le |i| \lesssim i_* = N^{\frac{1}{2}+C_*\omega_1}$; both described by an exponent.

C1. ***Short-long approximation*: *small gap and exact cusp.*** In this subsection we estimate the difference of the solution of the DBM $\tilde{z}(t,\alpha)$ defined in (6-14) and its short-range approximation $\hat{z}(t,\alpha)$ defined by (6-97)–(6-100) for $0 \le t \le t_*$. We formulate Lemma C.1 (for $0 \le t \le t_*$) below a bit more generally than we need in order to indicate the dependence of the approximation precision on these two exponents. For our actual application in Lemmas 6.9 and 7.2 we use specific exponents.

Lemma C.1. *Let $\omega_1 \ll \omega_\ell \ll \omega_A \ll 1$. Let $0 < a_0 \le \frac{1}{4} + C\omega_1$, $C > 0$ a universal constant and $0 < a \le C\omega_1$. Let $i_* := N^{\frac{1}{2}+C_*\omega_1}$ with C_* defined in Proposition 6.3. We assume that*

$$|\tilde{z}_i(t,\alpha)-\bar{\gamma}_i(t)| \le \frac{N^{a_0}}{N^{\frac{3}{4}}}, \quad 1 \le |i| \le N,\ 0 \le t \le t_*, \tag{C-1}$$

and

$$|\tilde{z}_i(t,\alpha)-\bar{\gamma}_i(t)| \le \frac{N^{a}}{N^{\frac{3}{4}}}, \quad 1 \le |i| \le i_*,\ 0 \le t \le t_*. \tag{C-2}$$

Then, for any $\alpha \in [0,1]$, we have

$$\sup_{1\le|i|\le N}\ \sup_{0\le t\le t_*} |\hat{z}_i(t,\alpha)-\tilde{z}_i(t,\alpha)| \le \frac{N^a N^{C\omega_1}}{N^{\frac{3}{4}}}\left(\frac{1}{N^{2\omega_\ell}} + \frac{N^{\frac{1}{2}\omega_A}\log N}{N^{\frac{1}{6}}N^a} + \frac{N^{\frac{1}{2}\omega_A}\log N}{N^{\frac{1}{4}}N^a} + \frac{1}{N^{\frac{2}{5}a}i_*^{\frac{1}{5}}} + \frac{N^{a_0}}{N^a i_*^{\frac{1}{2}}} + \frac{1}{N^{\frac{1}{18}}N^a}\right) \tag{C-3}$$

with very high probability.

Proof of Lemma 6.9. Use Lemma C.1 with the choice $a_0 = \frac{1}{4} + C\omega_1$ and $a = C\omega_1$ for some universal constant $C > 0$. The conditions (C-1) and (C-2) are guaranteed by (6-21) and (6-22). □

Proof of Lemma C.1. Let $w_i := \hat{z}_i - \tilde{z}_i$; hence w is a solution of

$$\partial_t w = \mathcal{B}_1 w + \mathcal{V}_1 w + \zeta, \tag{C-4}$$

where the operator $\mathcal{B}_1$ is defined for any $f \in \mathbb{C}^{2N}$ by

$$(B_1 f)_i = \frac{1}{N}\sum_j^{\mathcal{A},(i)} \frac{f_j - f_i}{(\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha))(\hat{z}_i(t,\alpha)-\hat{z}_j(t,\alpha))}. \tag{C-5}$$

The diagonal operator $\mathcal{V}_1$ is defined by $(\mathcal{V}_1 f)_i = \mathcal{V}_1(i) f_i$, where

$$\mathcal{V}_1(i) := -\int_{\mathcal{I}_{y,i}(t)^c} \frac{\rho_{y,t}(E+\mathfrak{e}^+_{y,t})}{(\tilde{z}_i(t,\alpha)-E)(\hat{z}_i(t,\alpha)-E)}\,\mathrm{d}E \qquad \text{for } 0<|i|\le N^{\omega_A}, \tag{C-6}$$

$$\mathcal{V}_1(i) := -\int_{\mathcal{I}_{z,i}(t)^c\cap\mathcal{J}_z(t)} \frac{\bar{\rho}_t(E+\bar{\mathfrak{e}}^+_t)}{(\tilde{z}_i(t,\alpha)-E)(\hat{z}_i(t,\alpha)-E)}\,\mathrm{d}E \quad \text{for } N^{\omega_A}<|i|\le \tfrac{1}{2}i_*. \tag{C-7}$$

Finally, $\mathcal{V}_1(i)=0$ for $|i|\ge\frac{1}{2}i_*$. The vector ζ in (C-4) collects various error terms.

We define the stopping time

$$T := \max\Big\{t\in[0,t_*] : \sup_{0\le s\le t} |\tilde{z}_i(s,\alpha)-\hat{z}_i(s,\alpha)| \le \tfrac{1}{2}\min\{|\mathcal{I}_{z,i}(t)|, |\mathcal{I}_{y,i}(t)|\} \text{ for all } \alpha\in[0,1]\Big\}, \tag{C-8}$$

where we recall that

$$|\mathcal{I}_{z,i}(t)| \sim |\mathcal{I}_{y,i}(t)| \sim N^{-\frac{3}{4}+3\omega_\ell}.$$

For $0\le t\le T$ we have $\mathcal{V}_1\le 0$. Therefore, since

$$\sum_i (\mathcal{B}f)_i = 0,$$

by the symmetry of $\mathcal{A}$, the semigroup of $\mathcal{B}_1+\mathcal{V}_1$, denoted by $\mathcal{U}^{\mathcal{B}_1+\mathcal{V}_1}$, is a contraction on every ℓ^p space. Hence, since $w(0)=0$ by (6-100), we have

$$w(t) = \int_0^t \mathcal{U}^{\mathcal{B}_1+\mathcal{V}_1}(s,t)\zeta(s)\,\mathrm{d}s,$$

and so

$$\|w(t)\|_\infty \le t \sup_{0\le s\le t}\|\zeta(s)\|_\infty \le N^{-\frac{1}{2}+\omega_1}\sup_{0\le s\le t}\|\zeta(s)\|_\infty. \tag{C-9}$$

Thus, to prove (C-3) it is enough to estimate $\|\zeta(s)\|_\infty$, for all $0\le s\le t_*$.

The error term ζ is given by $\zeta_i=0$ for $|i|>\frac{1}{2}i_*$; then for $1\le|i|\le N^{\omega_A}$, ζ_i is defined as

$$\zeta_i = \int_{\mathcal{I}_{y,i}(t)^c} \frac{\rho_{y,t}(E+\mathfrak{e}^+_{y,t})}{\tilde{z}_i(t,\alpha)-E}\,\mathrm{d}E - \frac{1}{N}\sum_j^{\mathcal{A}^c,(i)} \frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)} + \Phi_\alpha(t) - \Re[m_{y,t}(\mathfrak{e}^+_{y,t})], \tag{C-10}$$

with $\Phi_\alpha(t)$ defined in (6-15), and for $N^{\omega_A}<|i|\le\frac{1}{2}i_*$ as

$$\zeta_i = \int_{\mathcal{I}_{z,i}(t)^c\cap\mathcal{J}_z(t)} \frac{\bar{\rho}_t(E+\bar{\mathfrak{e}}^+_t)}{\tilde{z}_i(t,\alpha)-E}\,\mathrm{d}E - \frac{1}{N}\sum_{1\le|j|<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)} \frac{1}{\tilde{z}_i(t,\alpha)-\tilde{z}_j(t,\alpha)}. \tag{C-11}$$

Note that in the sum in (C-11) we do not have the summation over $|j|\ge\frac{3}{4}i_*$ since if $1\le|i|\le\frac{1}{2}i_*$ and $|j|\ge\frac{3}{4}i_*$ then $(i,j)\in\mathcal{A}^c$.

In the following we will often omit the t- and the α-arguments from $\tilde{z}_i$ and $\bar{\gamma}_i$ for notational simplicity.

First, we consider the error term (C-11) for $N^{\omega_A} < |i| \le \frac{1}{2} i_*$. We start with the estimate

$$
\begin{aligned}
|\zeta_i| &= \left| \int_{\mathcal{I}^c_{z,i}(t)\cap\mathcal{J}_z(t)} \frac{\bar\rho_t(E+\bar{\mathfrak{e}}_t^+)}{\tilde z_i - E}\,\mathrm{d}E - \frac{1}{N}\sum_{1\le|j|<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)} \frac{1}{\tilde z_i - \tilde z_j}\right| \\
&\lesssim \left| \sum_{1\le|j|<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)} \int_{\bar\gamma_j}^{\bar\gamma_{j+1}} \frac{\bar\rho_t(E+\bar{\mathfrak{e}}_t^+)(E-\bar\gamma_j)}{(\tilde z_i - E)(\tilde z_i - \bar\gamma_j)}\,\mathrm{d}E\right| + \left|\frac{1}{N}\sum_{1\le|j|<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)} \frac{\tilde z_j - \bar\gamma_j}{(\tilde z_i - \tilde z_j)(\tilde z_i - \bar\gamma_j)}\right| \\
&\quad + \left|\int_{\bar\gamma_{j_+}}^{\bar\gamma_{j_+ +1}} \frac{\bar\rho_t(E+\bar{\mathfrak{e}}_t^+)}{\tilde z_i - E}\,\mathrm{d}E\right| + \left|\int_{\bar\gamma_{-(3/4)i_*}}^{\bar\gamma_{-(3/4)i_*+1}} \frac{\bar\rho_t(E+\bar{\mathfrak{e}}_t^+)}{\tilde z_i - E}\,\mathrm{d}E\right| + \left|\int_0^{\bar\gamma_1} \frac{\bar\rho_t(E+\bar{\mathfrak{e}}_t^+)}{\tilde z_i - E}\,\mathrm{d}E\right|. \qquad \text{(C-12)}
\end{aligned}
$$

Since $|j_+ - i| \ge N^{4\omega_\ell} + N^{\omega_\ell}|i|^{\frac{3}{4}}$ and N^{ω_A}, i.e.,

$$
|\bar\gamma_{j_+} - \bar\gamma_i| \ge \frac{N^{\omega_\ell}|i|^{\frac{1}{2}}}{N^{\frac{3}{4}}}
$$

is bigger than the rigidity scale (C-2), all terms in the last line of (C-12) are bounded by $N^{-\frac{1}{4}-3\omega_\ell}$.

Then, using the rigidity estimate in (C-2) for the first and the second term of the right-hand side of (C-12), we conclude that

$$
|\zeta_i| \lesssim \frac{N^a}{N^{\frac{7}{4}}} \sum_{1\le|j|<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)} \frac{1}{(\bar\gamma_i - \bar\gamma_j)^2} + N^{-\frac{1}{4}-3\omega_\ell}. \qquad \text{(C-13)}
$$

The sum on the right-hand side of (C-13) is over all the j, negative and positive, but the main contribution comes from i and j with the same sign, because if i and j have opposite signs then

$$
\frac{1}{(\bar\gamma_i - \bar\gamma_j)^2} \le \frac{1}{(\bar\gamma_{-i} - \bar\gamma_j)^2}.
$$

Hence, assuming that i is positive (for negative i's we proceed exactly in the same way), we conclude that

$$
|\zeta_i| \lesssim \frac{N^a}{N^{\frac{7}{4}}} \sum_{1\le j<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)} \frac{1}{(\bar\gamma_i - \bar\gamma_j)^2} + N^{-\frac{1}{4}-3\omega_\ell}. \qquad \text{(C-14)}
$$

From now we assume that both i and j are positive. In order to estimate (C-14) we use the explicit expression of the quantiles from (4-13a), i.e.,

$$
\bar\gamma_j \sim \max\left\{ \left(\frac{j}{N}\right)^{\frac{2}{3}} \bar\Delta_t^{\frac{1}{9}}, \left(\frac{j}{N}\right)^{\frac{3}{4}} \right\},
$$

where $\bar\Delta_t \lesssim t_*^{\frac{3}{2}}$ denotes the length of the small gap of $\bar\rho_t$, for all $|j| \le i_* \sim N^{\frac{1}{2}}$. A simple calculation from (4-13a) shows that in the regime $i \ge N^{\omega_A}$ and $j \in \mathcal{A}^c$ we may replace $|\bar\gamma_i - \bar\gamma_j| \sim |\gamma_{y,i}(t) - \gamma_{y,j}(t)| \sim$

$|i^{\frac{3}{4}} - j^{\frac{3}{4}}|/N^{\frac{3}{4}}$; hence

$$|\zeta_i| \lesssim \frac{N^a}{N^{\frac{1}{4}}} \sum_{1 \le j < \frac{3}{4} i_*}^{\mathcal{A}^c, (i)} \frac{i^{\frac{1}{2}} + j^{\frac{1}{2}}}{(i-j)^2} + N^{-\frac{1}{4} - 3\omega_\ell}. \tag{C-15}$$

In fact, the same replacement works if either $i \ge N^{4\omega_\ell}$ or $j \ge N^{4\omega_\ell}$ and at least one of these two inequalities always holds as $(i, j) \in \mathcal{A}^c$. Using $i \le \frac{1}{2} i_*$ and that by the restriction $(i, j) \in \mathcal{A}^c$ we have $|j - i| \ge \ell(\ell^3 + i^{\frac{3}{4}})$, elementary calculation gives

$$|\zeta_i| \lesssim \frac{N^a}{N^{\frac{1}{4}} N^{2\omega_\ell}}. \tag{C-16}$$

Since analogous computations hold for i and j both negative, we have

$$|\zeta_i| \lesssim \frac{N^a}{N^{\frac{1}{4}} N^{2\omega_\ell}} \quad \text{for any } N^{\omega_A} < |i| \le \tfrac{1}{2} i_* \tag{C-17}$$

with very high probability.

Next, we proceed with the bound for ζ_i for $|i| \le N^{\omega_A}$. From (C-10) we have

$$\begin{aligned} \zeta_i = &\left(\int_{\mathcal{I}_{z,i}(t)^c \cap \mathcal{J}_z(t)} \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}_t^+)}{\tilde{z}_i - E} \, \mathrm{d}E - \frac{1}{N} \sum_{|j| < \frac{3}{4} i_*}^{\mathcal{A}^c, (i)} \frac{1}{\tilde{z}_i - \tilde{z}_j} \right) \\ &+ \left(\int_{\mathcal{J}_z(t)^c} \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}_t^+)}{\tilde{z}_i - E} \, \mathrm{d}E - \frac{1}{N} \sum_{|j| \ge \frac{3}{4} i_*}^{\mathcal{A}^c, (i)} \frac{1}{\tilde{z}_i - \tilde{z}_j} \right) \\ &+ \Phi_\alpha(t) - \Re[\bar{m}_t(\tilde{z}_i + \bar{\mathfrak{e}}_t^+)] + \Re[m_{y,t}(\tilde{z}_i + \mathfrak{e}_{y,t}^+)] - \Re[m_{y,t}(\mathfrak{e}_{y,t}^+)] \\ &+ \left(\int_{\mathcal{I}_{z,i}(t)} \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}_t^+)}{\tilde{z}_i - E} \, \mathrm{d}E - \int_{\mathcal{I}_{y,i}(t)} \frac{\rho_{y,t}(E + \mathfrak{e}_{y,t}^+)}{\tilde{z}_i - E} \, \mathrm{d}E \right) =: A_1 + A_2 + A_3 + A_4. \end{aligned} \tag{C-18}$$

By the remark after (C-15), the estimate of A_1 proceeds as in (C-15) and so we conclude that

$$|A_1| \lesssim \frac{N^a}{N^{\frac{1}{4}} N^{2\omega_\ell}}. \tag{C-19}$$

To estimate A_2, we first notice that the restriction $(i, j) \in \mathcal{A}^c$ in the summation is superfluous for $|i| \le N^{\omega_A}$ and $|j| \ge \frac{3}{4} i_*$. Let $\eta_1 \in [N^{-\frac{3}{4} + \frac{3}{4}\omega_A}, N^{-\delta}]$, for some small fixed $\delta > 0$, be an auxiliary scale we will determine later in the proof; then we write A_2 as

$$\begin{aligned} A_2 = &\left(\int_{\mathcal{J}_z(t)^c} \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}_t^+)}{\tilde{z}_i - E} \, \mathrm{d}E - \int_{\mathcal{J}_z(t)^c} \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}_t^+)}{\tilde{z}_i - E + \mathrm{i}\eta_1} \, \mathrm{d}E \right) \\ &+ \left(\frac{1}{N} \sum_{|j| \ge \frac{3}{4} i_*} \frac{1}{\tilde{z}_i - \tilde{z}_j + \mathrm{i}\eta_1} - \frac{1}{N} \sum_{|j| \ge \frac{3}{4} i_*} \frac{1}{\tilde{z}_i - \tilde{z}_j} \right) \\ &+ \left(\frac{1}{N} \sum_{|j| < \frac{3}{4} i_*} \frac{1}{\tilde{z}_i - \tilde{z}_j + \mathrm{i}\eta_1} - \int_{\mathcal{J}_z(t)} \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}_t^+)}{\tilde{z}_i - E + \mathrm{i}\eta_1} \, \mathrm{d}E \right) \\ &+ (\bar{m}_t(\tilde{z}_i + \mathrm{i}\eta_1) - m_{2N}(\tilde{z}_i + \mathrm{i}\eta_1, t, \alpha)) =: A_{2,1} + A_{2,2} + A_{2,3} + A_{2,4}, \end{aligned} \tag{C-20}$$

where we introduced

$$m_{2N}(z,t,\alpha) := \frac{1}{N}\sum_{|j|\le N}\frac{1}{z_j(t,\alpha)-z},\quad z\in\mathbb{H}.$$

For $1\le |i|\le N^{\omega_A}$ and $|j|>\frac{3}{4}i_*$, the term $A_{2,2}$ is bounded by the crude rigidity (C-1) as

$$|A_{2,2}|\le \frac{1}{N}\sum_{|j|>\frac{3}{4}i_*}\frac{\eta_1}{(\tilde{z}_i-\tilde{z}_j)^2}\lesssim \frac{N^{\frac{1}{2}}\eta_1}{i_*^{\frac{1}{2}}}. \tag{C-21}$$

Exactly the same estimate holds for $A_{2,1}$.

Next, using the rigidity estimates in (C-1) and (C-2) we conclude that

$$\begin{aligned}|A_{2,4}| &\lesssim \frac{1}{N}\sum_{1\le|j|\le i_*}\frac{|\tilde{z}_j-\bar{\gamma}_j|}{|\tilde{z}_i-\tilde{z}_j+\mathrm{i}\eta_1|^2}+\frac{1}{N}\sum_{i_*\le|j|\le N}\frac{|\tilde{z}_j-\bar{\gamma}_j|}{|\tilde{z}_i-\tilde{z}_j+\mathrm{i}\eta_1|^2}\\ &\lesssim \frac{N^a}{N^{\frac{3}{4}}\eta_1}\Im m_N(\bar{\gamma}_i+\mathrm{i}\eta_1)+\frac{N^{a_0}}{N^{\frac{7}{4}}}\sum_{i_*\le|j|\le N}\frac{1}{(\bar{\gamma}_i-\bar{\gamma}_j)^2}\\ &\lesssim \frac{N^a}{N^{\frac{3}{4}}\eta_1}\left(\frac{N^{\frac{3}{4}\omega_A}}{N^{\frac{3}{4}}}+\eta_1\right)^{\frac{1}{3}}+\frac{N^{a_0}}{N^{\frac{1}{4}}i_*^{\frac{1}{2}}}\lesssim \frac{N^a}{N^{\frac{3}{4}}\eta_1^{\frac{2}{3}}}+\frac{N^{a_0}}{i_*^{\frac{1}{2}}N^{\frac{1}{4}}}.\end{aligned} \tag{C-22}$$

Here we used that the rigidity scale near i for $1\le|i|\le N^{\omega_A}$ is much smaller than $\eta_1\ge N^{-\frac{3}{4}+\frac{3}{4}\omega_A}$. In particular, we know that $\Im m_N(\bar{\gamma}_i+\mathrm{i}\eta_1)$ can be bounded by the density $\bar{\rho}_t(\bar{\gamma}_i+\eta_1)$, which in turn is bounded by $(\bar{\gamma}_i+\eta_1)^{\frac{1}{3}}$. Similarly we conclude that

$$|A_{2,3}|\le\frac{N^a}{N^{\frac{3}{4}}\eta_1^{\frac{2}{3}}}.$$

Optimising (C-21) and (C-22) for η_1, we choose $\eta_1=(i_*^{\frac{1}{2}}N^{a-\frac{5}{4}})^{\frac{3}{5}}$, which falls into the required interval for η_1. Collecting all estimates for the parts of A_2 in (C-20), we therefore conclude that

$$|A_2|\le\frac{N^{\frac{3}{5}a}}{i_*^{\frac{1}{5}}N^{\frac{1}{4}}}+\frac{N^{a_0}}{i_*^{\frac{1}{2}}N^{\frac{1}{4}}}. \tag{C-23}$$

Next, we treat A_3 from (C-18). By (6-16), we have $\Phi_\alpha(t)=\Re[\bar{m}_t(\bar{\mathfrak{e}}_t^+)]+\mathcal{O}(N^{-1})$, and so by (4-22a) we conclude that

$$\begin{aligned}|A_3| &= \left|\Re[\bar{m}_t(\bar{\mathfrak{e}}_t^+)]-\Re[\bar{m}_t(\tilde{z}_i+\bar{\mathfrak{e}}_t^+)]+\Re[m_{y,t}(\tilde{z}_i+\mathfrak{e}_{y,t}^+)]-\Re[m_{y,t}(\mathfrak{e}_{y,t}^+)]\right|\\ &\lesssim\left(\frac{|i|^{\frac{1}{4}}N^{\frac{7}{18}\omega_1}}{N^{\frac{1}{4}}N^{\frac{1}{6}}}+\frac{|i|^{\frac{1}{2}}}{N^{\frac{1}{2}}}\right)|\log|\bar{\gamma}_i||\lesssim\frac{N^{\frac{1}{4}\omega_A}N^{\frac{7}{18}\omega_1}\log N}{N^{\frac{1}{4}}N^{\frac{1}{6}}}+\frac{N^{\frac{1}{2}\omega_A}\log N}{N^{\frac{1}{2}}}.\end{aligned} \tag{C-24}$$

We proceed writing A_4 as

$$\begin{aligned}A_4 &= \left(\int_{\mathcal{I}_{z,i}(t)}\frac{\bar{\rho}_t(E+\bar{\mathfrak{e}}_t^+)}{\tilde{z}_i-E}\,\mathrm{d}E-\int_{\mathcal{I}_{z,i}(t)}\frac{\rho_{y,t}(E+\mathfrak{e}_{y,t}^+)}{\tilde{z}_i-E}\,\mathrm{d}E\right)\\ &\quad+\left(\int_{\mathcal{I}_{z,i}(t)}\frac{\rho_{y,t}(E+\mathfrak{e}_{y,t}^+)}{\tilde{z}_i-E}\,\mathrm{d}E-\int_{\mathcal{I}_{y,i}(t)}\frac{\rho_{y,t}(E+\mathfrak{e}_{y,t}^+)}{\tilde{z}_i-E}\,\mathrm{d}E\right)=:A_{4,1}+A_{4,2}.\end{aligned} \tag{C-25}$$

We start with the estimate for $A_{4,2}$. By (6-96) and the comparison estimates between $\bar{\gamma}_{z,i}$ and $\hat{\gamma}_{y,i}$ by (4-18) we have

$$|\mathcal{I}_{z,i}(t)\Delta\mathcal{I}_{y,i}(t)| \lesssim |\bar{\gamma}_{z,i-j_-(i)} - \hat{\gamma}_{y,i-j_-(i)}| + |\bar{\gamma}_{z,i+j_+(i)} - \hat{\gamma}_{y,i+j_+(i)}| \lesssim \frac{N^{\frac{1}{2}\omega_1}(\ell^3 + |i|^{\frac{3}{4}})}{N^{\frac{11}{12}}}, \quad \text{(C-26)}$$

where Δ is the symmetric difference. In the second inequality of (C-26) we used that $|i \pm j_\pm(i)| \lesssim N^{\omega_A}$ and $\omega_A \ll 1$. For $E \in \mathcal{I}_{z,i}\Delta\mathcal{I}_{y,i}$ we have

$$\left|\frac{\rho_{y,t}(E + \mathfrak{e}^+_{y,t})}{\tilde{z}_i - E}\right| \lesssim \frac{N^{\frac{1}{2}}(\ell^2 + |i|^{\frac{1}{2}})}{\ell^3 + |i|^{\frac{3}{4}}}, \quad \text{(C-27)}$$

and so, using $|i| \le N^{\omega_A}$,

$$|A_{4,2}| \lesssim \frac{N^{\frac{1}{2}\omega_1}N^{\frac{1}{2}\omega_A}}{N^{\frac{5}{12}}} = \frac{N^{\frac{1}{2}\omega_1}N^{\frac{1}{2}\omega_A}}{N^{\frac{1}{4}}N^{\frac{1}{6}}} \quad \text{(C-28)}$$

with very high probability.

To estimate the integral in $A_{4,1}$ we have to deal with the logarithmic singularity due to the values of E close to $\tilde{z}_i(t)$. For $\max\{\bar{\mathfrak{e}}^-_t, \mathfrak{e}^-_{y,t}\} < E \le 0$ we have

$$\rho_{y,t}(E + \mathfrak{e}^+_{y,t}) = \bar{\rho}_t(E + \bar{\mathfrak{e}}^+_t) = 0. \quad \text{(C-29)}$$

For $\min\{\bar{\mathfrak{e}}^-_t, \mathfrak{e}^-_{y,t}\} \le E \le \max\{\bar{\mathfrak{e}}^-_t, \mathfrak{e}^-_{y,t}\}$, using the $\frac{1}{3}$-Hölder continuity of $\bar{\rho}_t$ and $\rho_{y,t}$ and (4-6a) we have

$$|\rho_{y,t}(E + \mathfrak{e}^+_{y,t}) - \bar{\rho}_t(E + \bar{\mathfrak{e}}^+_t)| \lesssim \Delta^{\frac{1}{3}}_{y,t}(t_* - t)^{\frac{1}{9}} \lesssim \frac{N^{\frac{11}{18}\omega_1}}{N^{\frac{11}{36}}} \quad \text{(C-30)}$$

for all $0 \le t \le t_*$. In the last inequality we used that $\Delta_{y,t} \le \Delta_{y,0} \lesssim N^{-\frac{3}{4}+\frac{1}{2}3\omega_1}$ for all $t \le t_*$. Similarly, for $E \le \min\{\bar{\mathfrak{e}}^-_t, \mathfrak{e}^-_{y,t}\}$ we have

$$|\rho_{y,t}(E + \mathfrak{e}^+_{y,t}) - \bar{\rho}_t(E + \bar{\mathfrak{e}}^+_t)| \lesssim |\rho_{y,t}(E' + \mathfrak{e}^-_{y,t}) - \bar{\rho}_t(E' + \bar{\mathfrak{e}}^-_t)| + \Delta^{\frac{1}{3}}_{y,t}(t_* - t)^{\frac{1}{9}}, \quad \text{(C-31)}$$

with $E' \le 0$.

Using (4-6b) for $E \ge 0$ and combining (4-6b) with (C-29)–(C-31) for $E < 0$, we have

$$\begin{aligned}|A_{4,1}| \lesssim \left(\frac{(\ell + |i|^{\frac{1}{4}})N^{\frac{1}{3}\omega_1}}{N^{\frac{1}{4}}N^{\frac{1}{6}}} + \frac{(\ell^2 + |i|^{\frac{1}{2}})}{N^{\frac{1}{2}}} + \frac{N^{\frac{11}{18}\omega_1}}{N^{\frac{11}{36}}}\right) \int_{\mathcal{I}_{,i}(t)\cap\{|E-\tilde{z}_i|>N^{-60}\}} \frac{1}{|\tilde{z}_i - E|}\,\mathrm{d}E \\ + \left|\int_{|E-\tilde{z}_i|\le N^{-60}} \frac{\bar{\rho}_t(E + \bar{\mathfrak{e}}^+_t) - \rho_{y,t}(E + \mathfrak{e}^+_{y,t})}{\tilde{z}_i - E}\,\mathrm{d}E\right|. \quad \text{(C-32)}\end{aligned}$$

The two singular integrals in the second line are estimated separately. By the $\frac{1}{3}$-Hölder continuity $\rho_{y,t}$ we conclude that

$$\begin{aligned}\left|\int_{|E-\tilde{z}_i|\le N^{-60}} \frac{\rho_{y,t}(E + \mathfrak{e}^+_{y,t})}{\tilde{z}_i - E}\,\mathrm{d}E\right| &= \left|\int_{|E-\tilde{z}_i|\le N^{-60}} \frac{\rho_{y,t}(E + \mathfrak{e}^+_{y,t}) - \rho_{y,t}(\tilde{z}_i + \mathfrak{e}^+_{y,t})}{\tilde{z}_i - E}\,\mathrm{d}E\right| \\ &\lesssim \int_{|E-\tilde{z}_i|\le N^{-60}} \frac{1}{|\tilde{z}_i - E|^{\frac{2}{3}}}\,\mathrm{d}E \lesssim N^{-20}.\end{aligned}$$

The same bound holds for the other singular integral in (C-32) by using the $\frac{1}{3}$-Hölder continuity of $\bar{\rho}_t$. Hence, for $1 \le |i| \le N^{\omega_A}$, by (C-32) we have

$$|A_{4,1}| \le \frac{N^{\frac{1}{4}\omega_A} N^{\frac{1}{3}\omega_1} \log N}{N^{\frac{1}{4}} N^{\frac{1}{6}}} + \frac{N^{\frac{1}{2}\omega_A} \log N}{N^{\frac{1}{2}}} + \frac{N^{\frac{11}{18}\omega_1} \log N}{N^{\frac{11}{36}}} \tag{C-33}$$

with very high probability.

Collecting all the estimates (C-17), (C-19), (C-23), (C-24), (C-28) and (C-33), and recalling $\omega_1 \ll \omega_\ell \ll \omega_A \ll 1$, we see that (C-19) is the largest term and thus $|\zeta| \lesssim N^{-\frac{1}{4}-2\omega_\ell} N^{C\omega_1}$ as $a \le C\omega_1$. Thus, using (C-9), we conclude that the estimate in (C-3) is satisfied for all $0 \le t \le T$. In particular, this means that

$$|\hat{z}_i(t,\alpha) - \tilde{z}_i(t,\alpha)| \le N^{-\frac{3}{4}+C\omega_1}, \quad 0 \le t \le T,$$

for some small constant $C > 0$. We conclude the proof of this lemma by showing that $T \ge t_*$.

Suppose by contradiction that $T < t_*$; then, since the solution of the DBM have continuous paths (see Theorem 12.2 of [Erdős and Yau 2017]), we have

$$|\hat{z}_i(T+\tilde{t},\alpha) - \tilde{z}_i(T+\tilde{t},\alpha)| \le \frac{N^a N^{c\omega_1}}{N^{\frac{3}{4}} N^{2\omega_\ell}}$$

for some tiny $\tilde{t} > 0$ and for any $\alpha \in [0,1]$. This bound is much smaller than the threshold $|\mathcal{I}_{y,i}(t)|, |\mathcal{I}_{z,i}(t)| \sim N^{-\frac{3}{4}} + 3\omega\ell$ in the definition of T. But this is a contradiction by the maximality in the definition of T; hence $T = t_*$, proving (C-3) for all $0 \le t \le t_*$. This completes the proof of Lemma C.1. □

Proof of Lemma 7.2. The proof of this lemma is very similar to that of Lemma C.1; hence we will only sketch the proof by indicating the differences. The main difference is that in this lemma we have optimal i-dependent rigidity for all $1 \le |i| \le i_*$. Hence, we can give a better estimate on the first two terms in (C-12) as follows (recall that $N^{\omega_A} \le i \le \frac{1}{2} i_*$):

$$|\zeta_i| \lesssim \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}} \sum_{|j|<\frac{3}{4}i_*} \frac{1}{(\bar{\gamma}_i - \bar{\gamma}_j)^2 |j|^{\frac{1}{4}}} \lesssim \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{3}{4}}} \sum_{|j|<\frac{3}{4}i_*} \frac{|i|^{\frac{1}{2}} + |j|^{\frac{1}{2}}}{(|i|-|j|)^2 |j|^{\frac{1}{4}}} \lesssim \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{1}{4}} N^{3\omega_\ell}}.$$

Compared with (C-16), the additional N^{ω_ℓ} factor in the denominator comes from the $|j|^{\frac{1}{4}}$ factor beforehand that is due to the optimal dependence of the rigidity on the index. Consequently, using the optimal rigidity in (6-3), we improve the denominator in the first term on the right-hand side of (C-3) from $N^{2\omega_\ell}$ to $N^{3\omega_\ell}$ with respect Lemma C.1.

Furthermore, by (6-3),

$$|A_{2,3}|, |A_{2,4}| \le \frac{N^\xi}{N\eta_1} \quad \text{and} \quad |A_{2,1}|, |A_{2,2}| \lesssim \frac{N^{\frac{1}{2}}\eta_1}{i_*^{\frac{1}{2}}} \lesssim N^{\frac{1}{4}-\frac{1}{2}C_*\omega_1} \eta_1,$$

since $i_* = N^{\frac{1}{2}+C_*\omega_1}$; hence, choosing $\eta_1 = N^{-\frac{5}{8}}$, we conclude that

$$|A_1| + |A_2| \lesssim \frac{N^\xi N^{\frac{1}{6}\omega_1}}{N^{\frac{1}{4}} N^{3\omega_\ell}} + \frac{N^\xi}{N^{\frac{3}{8}}}.$$

All other estimates follow exactly in the same way of the proof of Lemma C.1. This concludes the proof of Lemma 7.2. □

C2. ***Short-long approximation: small minimum.*** In this subsection we estimate the difference of the solution of the DBM $\tilde{z}(t,\alpha)$ defined by (8-10) and its short-range approximation $\hat{z}(t,\alpha)$ defined by (8-12)–(8-15) for $t_* \le t \le 2t_*$.

Lemma C.2. *Under the same assumption as Section C1 and assuming that the rigidity bounds* (C-1) *and* (C-2) *hold for the* $\tilde{z}(t,\alpha)$ *dynamics* (8-10) *for all* $t_* \le t \le 2t_*$, *we conclude that*

$$\sup_{1\le|i|\le N}\sup_{t_*\le t\le 2t_*}|\tilde{z}_i(t,\alpha)-\hat{z}_i(t,\alpha)| \lesssim \frac{N^a N^{C\omega_1}}{N^{\frac{3}{4}}}\left(\frac{1}{N^{2\omega_\ell}}+\frac{1}{N^{\frac{2}{5}a}i_*^{\frac{1}{5}}}+\frac{N^{a_0}}{N^a i_*^{\frac{1}{2}}}+\frac{1}{N^a N^{\frac{1}{24}}}\right) \tag{C-34}$$

with very high probability, *for any* $\alpha\in[0,1]$.

Proof. The proof of this lemma is similar to the proof of Section C1, but some estimates for the semicircular flow are slightly different mainly because in this lemma the $\tilde{z}_i(t,\alpha)$ are shifted by $\bar{\mathfrak{m}}_t$ instead of $\bar{\mathfrak{e}}_t^+$. Hence, we will skip some details in this proof, describing carefully only the estimates that are different with respect to Section C1.

Let $w_i := \hat{z}_i - \tilde{z}_i$; hence w is a solution of

$$\partial_t = \mathcal{B}_1 w + \mathcal{V}_1 w + \zeta,$$

where $\mathcal{B}_1$ and $\mathcal{V}_1$ are defined as in (C-5)–(C-7) substituting $\bar{\mathfrak{e}}_t^+$ by $\bar{\mathfrak{m}}_t$.

Without loss of generality we assume that $\mathcal{V}_1 \le 0$ for all $t_* \le t \le T$ (see (C-8) in the proof of Section C1 but now we have $t_* \le t \le 2t_*$ in the definition of the stopping time). This implies that $\mathcal{U}^{\mathcal{B}_1+\mathcal{V}_1}$ is a contraction semigroup and so in order to prove (C-34) it is enough to estimate

$$\sup_{t_*\le t\le T}\|\zeta(s)\|_\infty.$$

At the end, exactly as at the end of the proof of Lemma C.1, by continuity of the paths, we can easily establish $T = 2t_*$ for the stopping time.

The error term ζ is given by $\zeta_i = 0$ for $|i| > \frac{1}{2}i_*$; then ζ_i for $1 \le |i| \le N^{\omega_A}$ is defined as

$$\zeta_i = \int_{\mathcal{I}_{y,i}(t)^c}\frac{\rho_{y,t}(E+\tilde{\mathfrak{m}}_{y,t})}{\tilde{z}_i - E}\,\mathrm{d}E - \frac{1}{N}\sum_j^{\mathcal{A}^c,(i)}\frac{1}{\tilde{z}_i-\tilde{z}_j} + \Psi_\alpha(t) + \frac{\mathrm{d}}{\mathrm{d}t}\tilde{\mathfrak{m}}_{y,t}, \tag{C-35}$$

with $\Psi_\alpha(t)$ defined in (8-8), and for $N^{\omega_A} < |i| \le \frac{1}{2}i_*$ as

$$\zeta_i = \int_{\mathcal{I}_{z,i}(t)^c\cap\mathcal{J}_z(t)}\frac{\bar{\rho}_t(E+\bar{\mathfrak{m}}_t)}{\tilde{z}_i-E}\,\mathrm{d}E - \frac{1}{N}\sum_{|j|<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)}\frac{1}{\tilde{z}_i-\tilde{z}_j}. \tag{C-36}$$

We start to estimate the error term for $N^{\omega_A} < |i| \le \frac{1}{2}i_*$. By a similar computation to the one leading to (C-17) in Section C1, using (C-2), we conclude that

$$|\zeta_i| = \left|\int_{\mathcal{I}_{i,z}^c(t)\cap\mathcal{J}_z(t)}\frac{\bar{\rho}_t(E+\bar{\mathfrak{m}}_t)}{\tilde{z}_i-E}\,\mathrm{d}E - \frac{1}{N}\sum_{|j|<\frac{3}{4}i_*}^{\mathcal{A}^c,(i)}\frac{1}{\tilde{z}_i-\tilde{z}_j}\right| \lesssim \frac{N^a}{N^{\frac{1}{4}}N^{2\omega_\ell}}, \quad N^{\omega_A} < |i| \le \tfrac{1}{2}i_*. \tag{C-37}$$

Next, we proceed with the bound for ζ_i for $1 \leq |i| \leq N^{\omega_A}$. We rewrite ζ_i as

$$\zeta_i = \left(\int_{\mathcal{I}_{i,z}^c(t)} \frac{\bar{\rho}_t(E + \bar{\mathfrak{m}}_t)}{\tilde{z}_i - E} \, \mathrm{d}E - \frac{1}{N} \sum_j^{\mathcal{A}^c, (i)} \frac{1}{\tilde{z}_i - \tilde{z}_j} \right) \\ + \Re[m_{y,t}(\tilde{z}_i + \tilde{\mathfrak{m}}_{y,t})] + \frac{\mathrm{d}}{\mathrm{d}t} \tilde{\mathfrak{m}}_{y,t} + \Psi_\alpha(t) - \Re[\bar{m}_t(\tilde{z}_i + \bar{\mathfrak{m}}_t)] \\ + \left(\int_{\mathcal{I}_{z,i}(t)} \frac{\bar{\rho}_t(E + \bar{\mathfrak{m}}_t)}{\tilde{z}_i - E} \, \mathrm{d}E - \int_{\mathcal{I}_{y,i}(t)} \frac{\rho_{y,t}(E + \tilde{\mathfrak{m}}_{y,t})}{\tilde{z}_i - E} \, \mathrm{d}E \right) =: (A_1 + A_2) + A_3 + A_4, \quad \text{(C-38)}$$

where $(A_1 + A_2)$ indicates that for the actual estimates we split the first line in (C-38) into two terms as in (C-18). By similar computations to those in Section C1, see (C-19) and (C-23), we conclude that

$$|A_1| + |A_2| \lesssim \frac{N^a}{N^{\frac{1}{4}} N^{2\omega_\ell}} + \frac{N^{\frac{3}{5}a}}{N^{\frac{1}{4}} i_*^{\frac{1}{5}}} + \frac{N^{a_0}}{i_*^{\frac{1}{2}} N^{\frac{1}{4}}}. \quad \text{(C-39)}$$

By (4-14b), (4-14d), (4-22b) and the definition of $\Psi_\alpha(t)$ in (8-8) it follows that

$$|A_3| \lesssim \left| \Re[m_{y,t}(\tilde{z}_i + \tilde{\mathfrak{m}}_{y,t}) - m_{y,t}(\tilde{\mathfrak{m}}_{y,t})] - \Re[\bar{m}_t(\bar{\mathfrak{m}}_t) - \bar{m}_t(\tilde{z}_i + \bar{\mathfrak{m}}_t)] \right| + \frac{N^{\omega_1}}{N} \\ \lesssim \left(\frac{N^{\frac{1}{4}\omega_A} N^{\frac{1}{4}\omega_1}}{N^{\frac{1}{4}} N^{\frac{1}{8}}} + \frac{N^{\frac{3}{4}\omega_1}}{N^{\frac{3}{8}}} + \frac{N^{\frac{1}{2}\omega_A}}{N^{\frac{1}{2}}} \right) |\log|\hat{\gamma}_i(t)|| + \frac{N^{\frac{7}{12}\omega_1}}{N^{\frac{7}{24}}} \lesssim \frac{N^{\frac{7}{12}\omega_1}}{N^{\frac{7}{24}}}. \quad \text{(C-40)}$$

We proceed writing A_4 as

$$A_4 = \left(\int_{\mathcal{I}_{z,i}(t)} \frac{\bar{\rho}_t(E + \bar{\mathfrak{m}}_t)}{\tilde{z}_i - E} \, \mathrm{d}E - \int_{\mathcal{I}_{z,i}(t)} \frac{\rho_{y,t}(E + \tilde{\mathfrak{m}}_{y,t})}{\tilde{z}_i - E} \, \mathrm{d}E \right) \\ + \left(\int_{\mathcal{I}_{z,i}(t)} \frac{\rho_{y,t}(E + \tilde{\mathfrak{m}}_{y,t})}{\tilde{z}_i - E} \, \mathrm{d}E - \int_{\mathcal{I}_{y,i}(t)} \frac{\rho_{y,t}(E + \tilde{\mathfrak{m}}_{y,t})}{\tilde{z}_i - E} \, \mathrm{d}E \right) =: A_{4,1} + A_{4,2}. \quad \text{(C-41)}$$

We start with the estimate for $A_{4,2}$.

By (4-19) we have

$$|\mathcal{I}_{z,i}(t) \Delta \mathcal{I}_{y,i}(t)| \lesssim \frac{N^\xi (\ell + |i|)}{N}, \quad \text{(C-42)}$$

where Δ is the symmetric difference. Note that this bound is somewhat better than the analogous (C-26) due to the better bound in (4-19) compared with (4-18). For $E \in \mathcal{I}_{z,i}(t) \Delta \mathcal{I}_{y,i}(t)$ we have

$$\left| \frac{\rho_{y,t}(E + \bar{\mathfrak{m}}_t)}{\tilde{z}_i - E} \right| \lesssim \frac{N^{\frac{1}{2}} (\ell^2 + |i|^{\frac{1}{2}})}{\ell^3 + |i|^{\frac{3}{4}}}, \quad \text{(C-43)}$$

and so

$$|A_{4,2}| \lesssim \frac{N^{\frac{3}{4}\omega_A}}{N^{\frac{1}{2}}} \quad \text{(C-44)}$$

with very high probability.

To estimate the integral in $A_{4,1}$, we combine (4-6d) and (4-14b) to obtain

$$|\bar{\rho}_t(\bar{\mathfrak{m}}_t+E)-\rho_{y,t}(\tilde{\mathfrak{m}}_{y,t}+E)|\leq|\rho_{x,t}(\alpha\mathfrak{m}_{x,t}+(1-\alpha)\mathfrak{m}_{y,t}+E)-\rho_{y,t}(\mathfrak{m}_{y,t}+E)|+(t-t_*)^{\frac{7}{12}}. \quad \text{(C-45)}$$

Proceeding similarly to the estimate of $|A_{4,1}|$ at the end of the proof of Section C1, we conclude that

$$|A_{4,1}|\lesssim\left(\frac{N^{\xi}(\ell^2+|i|^{\frac{1}{2}})}{N^{\frac{1}{2}}}+\frac{N^{\frac{7}{12}\omega_1}}{N^{\frac{7}{24}}}\right)\int_{\mathcal{I}_{z,i}(t)\cap\{|E-\tilde{z}_i|>N^{-60}\}}\frac{1}{|\tilde{z}_i-E|}\,\mathrm{d}E + \left|\int_{|E-\tilde{z}_i|\leq N^{-60}}\frac{\bar{\rho}_t(E+\bar{\mathfrak{m}}_t)-\rho_{y,t}(E+\tilde{\mathfrak{m}}_{y,t})}{\tilde{z}_i-E}\,\mathrm{d}E\right|. \quad \text{(C-46)}$$

Furthermore, similarly to the estimate in the singular integral in (C-32), but substituting $\bar{\mathfrak{e}}_t^+$ and $\mathfrak{e}_{y,t}^+$ by $\bar{\mathfrak{m}}_t$ and $\tilde{\mathfrak{m}}_{y,t}$ respectively, we conclude that the last term in (C-46) is bounded by N^{-20}. Therefore,

$$|A_{41}|\lesssim\frac{N^{\xi}(\ell^2+|i|^{\frac{1}{2}})}{N^{\frac{1}{2}}}+\frac{N^{\frac{7}{12}\omega_1}}{N^{\frac{7}{24}}}\lesssim\frac{N^{\frac{7}{12}\omega_1}}{N^{\frac{7}{24}}} \quad \text{(C-47)}$$

for any $|i|\leq N^{\omega_A}$. Collecting (C-39), (C-40), (C-44) and (C-47) completes the proof of Lemma C.2. □

Appendix D: Sobolev-type inequality

The proof of the Sobolev-type inequality in the cusp case is essentially identical to that in the edge case presented in Appendix B of [Bourgade et al. 2014]; only the exponents need adjustment to the cusp scaling. We give some details for completeness.

Proof of Lemma 7.5. We will prove only the first inequality in (7-20). The proof for the second one is exactly the same. We start by proving a continuous version of (7-20) and then we will conclude the proof by linear interpolation. We claim that for any small η there exists a constant $c_\eta>0$ such that for any real function $f\in L^p(\mathbb{R}_+)$ we have

$$\int_0^{+\infty}\int_0^{+\infty}\frac{(f(x)-f(y))^2}{|x^{\frac{3}{4}}-y^{\frac{3}{4}}|^{2-\eta}}\,\mathrm{d}x\,\mathrm{d}y\geq c_\eta\left(\int_0^{+\infty}|f(x)|^p\,\mathrm{d}x\right)^{\frac{2}{p}}. \quad \text{(D-1)}$$

We recall the representation formula for fractional powers of the Laplacian: for any $0<\alpha<2$ and for any function $f\in L^p(\mathbb{R})$ for some $p\in[1,\infty)$ we have

$$\langle f,|p|^\alpha f\rangle=C(\alpha)\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{(f(x)-f(y))^2}{|x-y|^{1+\alpha}}\,\mathrm{d}x\,\mathrm{d}y, \quad \text{(D-2)}$$

with some explicit constant $C(\alpha)$, where $|p|:=\sqrt{-\Delta}$.

Since for $0<x<y$ we have

$$y^{\frac{3}{4}}-x^{\frac{3}{4}}=\frac{4}{3}\int_x^y s^{-\frac{1}{4}}\,\mathrm{d}s\leq C(y-x)(xy)^{-\frac{1}{8}},$$

in order to prove (D-1) it is enough to show that

$$\int_0^{+\infty}\int_0^{+\infty}\frac{(f(x)-f(y))^2}{|x-y|^{2-\eta}}(xy)^q\,\mathrm{d}x\,\mathrm{d}y \ge c_\eta\left(\int_0^{+\infty}|f(x)|^p\,\mathrm{d}x\right)^{\frac{2}{p}}, \tag{D-3}$$

where $q := \frac{1}{4}-\frac{1}{8}\eta$ and $p := 8/(2+3\eta)$. Let $\tilde{f}(x)$ be the symmetric extension of f to the whole real line; i.e., $\tilde{f}(x) := f(x)$ for $x > 0$ and $\tilde{f}(x) := f(-x)$ for $x < 0$. Then, by a simple calculation we have

$$4\int_0^{+\infty}\int_0^{+\infty}\frac{(f(x)-f(y))^2}{|x-y|^{2-\eta}}(xy)^q\,\mathrm{d}x\,\mathrm{d}y \ge \int_{\mathbb{R}}\int_{\mathbb{R}}\frac{(\tilde{f}(x)-\tilde{f}(y))^2}{|x-y|^{2-\eta}}|xy|^q\,\mathrm{d}x\,\mathrm{d}y.$$

Introducing $\phi(x) := |x|^q$ and dropping the tilde for f, the estimate in (D-3) would follow from

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{(f(x)-f(y))^2}{|x-y|^{2-\eta}}\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y \ge c'_\eta\left(\int_{\mathbb{R}}|f(x)|^p\,\mathrm{d}x\right)^{\frac{2}{p}}. \tag{D-4}$$

By the same computation as in the proof of Proposition 10.5 in [Bourgade et al. 2014] we conclude that

$$\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{(f(x)-f(y))^2}{|x-y|^{2-\eta}}\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y = \langle\phi f, |p|^{1-\eta}\phi f\rangle + C_0(\eta)\int_{\mathbb{R}}\frac{|\phi(x)f(x)|^2}{|x|^{1-\eta}}\,\mathrm{d}x,$$

with some $C_0(\eta) > 0$; hence for the proof of (D-4) it is enough to show that

$$\langle\phi f, |p|^{1-\eta}\phi f\rangle \ge c_\eta\left(\int_{\mathbb{R}}|f|^p\right)^{\frac{2}{p}}.$$

Let $g := |p|^{\frac{1}{2}(1-\eta)}|x|^q f$; we need to prove that

$$\|g\|_2 \ge c_\eta\||x|^{-q}|p|^{-\frac{1}{2}(1-\eta)}g\|_p.$$

By the n-dimensional Hardy–Littlewood–Sobolev inequality in [Stein and Weiss 1958] we have

$$\left\||x|^{-q}\int|x-y|^{-a}g(y)\,\mathrm{d}y\right\|_p \le C\|g\|_r,$$

where

$$\frac{1}{r}+\frac{a+q}{n} = 1+\frac{1}{p}, \quad 0 \le q < \frac{n}{p} \quad \text{and} \quad 0 < a < n.$$

In our case $a = \frac{1}{2}(1+\eta)$, $r = 2$, $n = 1$ and all the conditions are satisfied if we take $0 < \eta < 1$. This completes the proof of (D-1).

Next, in order to prove (7-20), we proceed by linear interpolation as in Proposition B.2 in [Erdős and Yau 2015]. Given $u : \mathbb{Z} \to \mathbb{R}$, let $\psi : \mathbb{R} \to \mathbb{R}$ be its linear interpolation; i.e., $\psi(i) := u_i$ for $i \in \mathbb{Z}$ and

$$\psi(x) := u_i + (u_{i+1}-u_i)(x-i) = u_{i+1} - (u_{i+1}-u_i)(i+1-x) \tag{D-5}$$

for $x \in [i, i+1]$. It is easy to see that for each $p \in [2, +\infty]$ $\left(\text{i.e., } \eta \le \frac{2}{3}\right)$ there exists a constant C_p such that

$$C_p^{-1}\|\psi\|_{L^p(\mathbb{R})} \le \|u\|_{L^p(\mathbb{Z})} \le C_p\|\psi\|_{L^p(\mathbb{R})}. \tag{D-6}$$

In order to prove (7-20) we claim that

$$\int_0^{+\infty}\int_0^{+\infty} \frac{|\psi(x)-\psi(y)|^2}{|x^{\frac34}-y^{\frac34}|^{2-\eta}}\,\mathrm{d}x\,\mathrm{d}y \le c_\eta \sum_{i\neq j\in\mathbb{Z}_+} \frac{(u_i-u_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}} \tag{D-7}$$

for some constant $c_\eta > 0$. Indeed, combining (D-6) and (D-7) with (D-1) we conclude (7-20). Finally, the proof of (D-7) is a simple exercise along the lines of the proof of Proposition B.2 in [Erdős and Yau 2015]. □

Appendix E: Heat-kernel estimates

The proof of the heat kernel estimates relies on the Nash method. In the edge scaling regime a similar bound was proven in [Bourgade et al. 2014] for a compact interval, extended to a noncompact interval but with compactly supported initial data w_0 in [Landon and Yau 2017]. Here we closely follow the latter proof, adjusted to the cusp regime, where interactions on both sides of the cusp play a role unlike in the edge regime.

Proof of Lemma 7.6. We start by proving (7-21); then (7-22) follows by (7-21) by duality. Without loss of generality we assume $\|w_0\|_1 = 1$ and that

$$\|w(\tilde{s})\|_p \ge N^{-100} \tag{E-1}$$

for each $s \le \tilde{s} \le t$, where $w(\tilde{s}) = \mathcal{U}^{\mathcal{L}}(s,\tilde{s})w_0$. Otherwise, by ℓ^p-contraction we have $\|w(\tilde{s})\|_p \le N^{-100}$ implying (7-21) directly.

In the following we use the convention $w := w(\tilde{s})$ if there is no confusion. By (7-20), we have

$$\|w\|_p^2 \lesssim \sum_{\substack{i,j\ge1\\ i\neq j}} \frac{(w_i-w_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}} + \sum_{\substack{i,j\le -1\\ i\neq j}} \frac{(w_i-w_j)^2}{||i|^{\frac34}-|j|^{\frac34}|^{2-\eta}}.$$

First we assume that both i and j are positive. Let $\delta_4<\delta_2<\delta_3<\frac12\delta_1$. We start with the estimate

$$\sum_{\substack{i,j\ge1\\ i\neq j}} \frac{(w_i-w_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}} \lesssim \sum_{\substack{(i,j)\in\mathcal{A}\\ i,j\ge1}} \frac{(w_i-w_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}} + \sum_{i\ge1}\sum_{j\ge1}^{\mathcal{A}^c,(i)} \frac{w_i^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}}. \tag{E-2}$$

We proceed by writing

$$\sum_{\substack{(i,j)\in\mathcal{A}\\ i,j\ge1}} \frac{(w_i-w_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}} \lesssim \sum_{\substack{(i,j)\in\mathcal{A}:\, i,j\ge1\\ i \text{ or } j\le \ell^4N^{\delta_2}}} \frac{(w_i-w_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}} + \sum_{\substack{(i,j)\in\mathcal{A}\\ i,j\ge \ell^4N^{\delta_2}}} \frac{(w_i-w_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}}. \tag{E-3}$$

By Lemma B.3 we have

$$\sum_{\substack{(i,j)\in\mathcal{A}\\ i,j\ge \ell^4N^{\delta_2}}} \frac{(w_i-w_j)^2}{|i^{\frac34}-j^{\frac34}|^{2-\eta}} \lesssim N^{-200}, \tag{E-4}$$

since $i \ge \ell^4 N^{\delta_2}$ and $|(w_0)_j| \le N^{-100}$ for $j \ge \ell^4 N^{\delta_4}$ by our hypotheses. Indeed, for $i \ge \ell^4 N^{\delta_2}$, we have

$$w_i = (\mathcal{U}^{\mathcal{L}}(s,\tilde{s})w_0)_i = \sum_{j=-N}^{N} \mathcal{U}^{\mathcal{L}}_{ij}(w_0)_j = \sum_{j=-\ell^4 N^{\delta_4}}^{\ell^4 N^{\delta_4}} \mathcal{U}^{\mathcal{L}}_{ij}(w_0)_j + N^{-100} \lesssim N^{-100} \tag{E-5}$$

with very high probability. If $(i,j) \in \mathcal{A}$, $i,j \ge 1$ and i or j is smaller than $\ell^4 N^{\delta_2}$, then both i and j are smaller than $\ell^4 N^{\delta_3}$. Hence, for such i and j, by (7-19), we have

$$|\hat{z}_i(t,\alpha) - \hat{z}_j(t,\alpha)| \lesssim \frac{N^{\frac{1}{6}\omega_1}|i^{\frac{3}{4}} - j^{\frac{3}{4}}|}{N^{\frac{3}{4}}} \tag{E-6}$$

for any fixed $\alpha \in [0,1]$ and for all $0 \le t \le t_*$, where $\hat{z}_i(t,\alpha)$ is defined by (6-106)–(6-107).

If i and j are both negative the estimates in (E-2)–(E-6) follow in the same way.

In the remainder of the proof $\mathcal{B}$, $\mathcal{B}_{ij}$ and $\mathcal{V}_i$ are defined as in (6-106)–(6-107). By (E-6) it follows that

$$\begin{aligned}\sum_{\substack{(i,j)\in\mathcal{A}:\, i,j\ge 1\\ i \text{ or } j \le \ell^4 N^{\delta_2}}} \frac{(w_i - w_j)^2}{|i^{\frac{3}{4}} - j^{\frac{3}{4}}|^{2-\eta}} + \sum_{\substack{(i,j)\in\mathcal{A}:\, i,j\le -1\\ i \text{ or } j \ge -\ell^4 N^{\delta_2}}} \frac{(w_i - w_j)^2}{|i^{\frac{3}{4}} - j^{\frac{3}{4}}|^{2-\eta}} &\lesssim -N^{-\frac{1}{2}} N^{\frac{1}{3}\omega_1 + C\eta} \sum_{(i,j)\in\mathcal{A}} \mathcal{B}_{ij}(w_i - w_j)^2 \\ &= -2N^{-\frac{1}{2}} N^{\frac{1}{3}\omega_1 + C\eta} \langle w, \mathcal{B}w\rangle. \end{aligned} \tag{E-7}$$

Furthermore, since $1 \le |i| \le \ell^4 N^{\delta_3}$, we have

$$\sum_j^{\mathcal{A}^c,(i)} \frac{1}{||i|^{\frac{3}{4}} - |j|^{\frac{3}{4}}|^{2-\eta}} \lesssim \frac{N^{\frac{1}{3}\omega_1 + C\eta}}{N^{\frac{3}{2}}} \sum_j^{\mathcal{A}^c,(i)} \frac{1}{(\hat{z}_i - \hat{z}_j)^2}. \tag{E-8}$$

By the rigidity (7-16), (7-17) and (7-19), we can replace $\hat{z}_j$ by $\bar{\gamma}_j$ in the sum on the right-hand side of (E-8) and so approximate it by an integral; then using that $\bar{\rho}_t(E) \lesssim \rho_{y,t}(E)$ in the cusp regime, i.e., $|E| \le \delta_*$, with δ_* defined in Definition 4.1, we conclude that

$$\frac{1}{N} \sum_j^{\mathcal{A}^c,(i)} \frac{1}{(\hat{z}_i(t) - \hat{z}_j(t))^2} \lesssim \int_{I_{i,y}(t)^c} \frac{\rho_{y,t}(E + \mathfrak{e}^+_{y,t})}{(\hat{z}_i(t) - E)^2}\, \mathrm{d}E = -\mathcal{V}_i. \tag{E-9}$$

Hence, by (E-9), we conclude that

$$\begin{aligned}\sum_i \sum_j^{\mathcal{A}^c,(i)} \frac{w_i^2}{||i|^{\frac{3}{4}} - |j|^{\frac{3}{4}}|^{2-\eta}} &\lesssim \sum_{1\le|i|\le\ell^4 N^{\delta_3}} \sum_j^{\mathcal{A}^c,(i)} \frac{w_i^2}{||i|^{\frac{3}{4}} - |j|^{\frac{3}{4}}|^{2-\eta}} + N^{-200} \\ &\lesssim -N^{-\frac{1}{2}} N^{\frac{1}{3}\omega_1 + C\eta} \sum_{|i|\le\ell^4 N^{\delta_3}} w_i^2 \mathcal{V}_i + N^{-200} \\ &\lesssim -N^{-\frac{1}{2}} N^{\frac{1}{3}\omega_1 + C\eta} \langle w, \mathcal{V}w\rangle + N^{-200}. \end{aligned} \tag{E-10}$$

Note that in the first inequality of (E-10) we used (E-5).

Summarising (E-4), (E-7) and (E-10) and rewriting N^{-200} into an ℓ^p-norm using (E-1), we obtain

$$\|w\|_p^2 \le -N^{-\frac{1}{2}} N^{\frac{1}{3}\omega_1 + C\eta} \langle w, \mathcal{L}w\rangle + \tfrac{1}{10}\|w\|_p^2.$$

Hence, using Hölder inequality, we have

$$\begin{aligned}\partial_t \|w\|_2^2 = \langle w, \mathcal{L}w\rangle &\le -c_\eta N^{\frac{1}{2}} N^{-\frac{1}{3}\omega_1 - C\eta} \|w\|_p^2 \\ &\le -c_\eta N^{\frac{1}{2}} N^{-\frac{1}{3}\omega_1 - C\eta} \|w\|_2^{\frac{1}{2}(6-3\eta)} \|w\|_1^{-\frac{1}{2}(2-3\eta)} \\ &\le -c_\eta N^{\frac{1}{2}} N^{-\frac{1}{3}\omega_1 - C\eta} \|w\|_2^{\frac{1}{2}(6-3\eta)} \|w_0\|_1^{-\frac{1}{2}(2-3\eta)}. \end{aligned} \tag{E-11}$$

In the last inequality of (E-11) we used the ℓ^1-contraction of $\mathcal{U}^{\mathcal{L}}$. Integrating (E-11) back in time, it easily follows that

$$\|\mathcal{U}^{\mathcal{L}}(s,t)w_0\|_2 \le \left(\frac{N^{C\eta+\frac{1}{3}\omega_1}}{c_\eta N^{\frac{1}{2}}(t-s)}\right)^{1-3\eta} \|w_0\|_1, \tag{E-12}$$

proving (7-21). The same bound also holds for the transpose operator $(\mathcal{U}^{\mathcal{L}})^T$.

In order to prove (7-22) we follow Lemma 3.11 of [Landon and Yau 2017]. Let $\chi(i) := \mathbf{1}_{\{|i|\le \ell^4 N^{\delta_5}\}}$, with $\delta_4 < \delta_5 < \frac{1}{2}\delta_1$, and $v \in \mathbb{R}^{2N}$. Then, we have

$$\langle \mathcal{U}^{\mathcal{L}}(0,t)w_0, v\rangle = \langle w_0, (\mathcal{U}^{\mathcal{L}})^T \chi v\rangle + \langle w_0, (\mathcal{U}^{\mathcal{L}})^T (1-\chi)v\rangle.$$

By Lemma B.3 we have

$$|\langle w_0, (\mathcal{U}^{\mathcal{L}})^T (1-\chi)v\rangle| \le N^{-100} \|w_0\|_2 \|v\|_1. \tag{E-13}$$

By (7-21) and the Cauchy–Schwarz inequality we have

$$|\langle w_0, (\mathcal{U}^{\mathcal{L}})^T \chi v\rangle| \le \|w_0\|_2 \|(\mathcal{U}^{\mathcal{L}})^T \chi v\|_2 \le \|w_0\|_2 \left(\frac{N^{C\eta+\frac{1}{3}\omega_1}}{c_\eta N^{\frac{1}{2}} t}\right)^{1-3\eta} \|v\|_1. \tag{E-14}$$

Hence, combining (E-13) and (E-14), we conclude that

$$\|\mathcal{U}^{\mathcal{L}}(0,t)w_0\|_\infty \le \left(\frac{N^{C\eta+\frac{1}{3}\omega_1}}{c_\eta N^{\frac{1}{2}} t}\right)^{1-3\eta} \|w_0\|_2, \tag{E-15}$$

and so, by (E-12), that

$$\begin{aligned}\|\mathcal{U}^{\mathcal{L}}(0,t)w_0\|_\infty = \left\|\mathcal{U}^{\mathcal{L}}\left(\tfrac{1}{2}t, t\right)\mathcal{U}^{\mathcal{L}}\left(0, \tfrac{1}{2}t\right)w_0\right\|_\infty &\lesssim \left(\frac{N^{C\eta+\frac{1}{3}\omega_1}}{c_\eta N^{\frac{1}{2}} t}\right)^{1-3\eta} \left\|\mathcal{U}^{\mathcal{L}}\left(0, \tfrac{1}{2}t\right)w_0\right\|_2 \\ &\lesssim \left(\frac{N^{C\eta+\frac{1}{3}\omega_1}}{c_\eta N^{\frac{1}{2}} t}\right)^{2(1-3\eta)} \|w_0\|_1, \end{aligned} \tag{E-16}$$

where in the first inequality we used that $\mathcal{U}^{\mathcal{L}}\left(0, \frac{1}{2}t\right)w_0$ satisfies the hypothesis of Lemma 7.6, since $\left|\left(\mathcal{U}^{\mathcal{L}}\left(0, \frac{1}{2}t\right)w_0\right)_i\right| \le N^{-100}$ for $|i| \ge \ell^4 N^{2\delta_4}$ by the finite-speed estimate of Lemma B.3. Combining (E-15) and (E-16), inequality (7-22) follows by interpolation. □

Acknowledgements

The authors are very grateful to Johannes Alt for his invaluable contribution in helping improve several results of [Alt et al. 2018a] tailored to the needs of this paper. Cipolloni, Erdős, and Schröder were partially supported by ERC Advanced Grant No. 338804. Cipolloni received funding for this project from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 665385. Krüger was partially supported by the Hausdorff Center for Mathematics.

References

[Adhikari and Huang 2018] A. Adhikari and J. Huang, "Dyson Brownian motion for general β and potential at the edge", preprint, 2018. arXiv 1810.08308

[Adler et al. 2010] M. Adler, P. L. Ferrari, and P. van Moerbeke, "Airy processes with wanderers and new universality classes", *Ann. Probab.* **38**:2 (2010), 714–769. MR Zbl

[Ajanki et al. 2015] O. Ajanki, L. Erdős, and T. Krüger, "Quadratic vector equations on complex upper half-plane", preprint, 2015. arXiv 1506.05095

[Ajanki et al. 2017a] O. Ajanki, L. Erdős, and T. Krüger, "Singularities of solutions to quadratic vector equations on the complex upper half-plane", *Comm. Pure Appl. Math.* **70**:9 (2017), 1672–1705. MR Zbl

[Ajanki et al. 2017b] O. H. Ajanki, L. Erdős, and T. Krüger, "Universality for general Wigner-type matrices", *Probab. Theory Related Fields* **169**:3-4 (2017), 667–727. MR Zbl

[Alt et al. 2018a] J. Alt, L. Erdős, and T. Krüger, "The Dyson equation with linear self-energy: spectral bands, edges and cusps", preprint, 2018. arXiv 1804.07752

[Alt et al. 2018b] J. Alt, L. Erdős, T. Krüger, and D. Schröder, "Correlated random matrices: band rigidity and edge universality", preprint, 2018. arXiv 1804.07744

[Bourgade and Yau 2017] P. Bourgade and H.-T. Yau, "The eigenvector moment flow and local quantum unique ergodicity", *Comm. Math. Phys.* **350**:1 (2017), 231–278. MR Zbl

[Bourgade et al. 2014] P. Bourgade, L. Erdős, and H.-T. Yau, "Edge universality of beta ensembles", *Comm. Math. Phys.* **332**:1 (2014), 261–353. MR Zbl

[Bourgade et al. 2016] P. Bourgade, L. Erdős, H.-T. Yau, and J. Yin, "Fixed energy universality for generalized Wigner matrices", *Comm. Pure Appl. Math.* **69**:10 (2016), 1815–1881. MR Zbl

[Brézin and Hikami 1998a] E. Brézin and S. Hikami, "Level spacing of random matrices in an external source", *Phys. Rev. E* (3) **58**:6 (1998), 7176–7185. MR

[Brézin and Hikami 1998b] E. Brézin and S. Hikami, "Universal singularity at the closure of a gap in a random matrix theory", *Phys. Rev. E* (3) **57**:4 (1998), 4140–4149. MR

[Claeys et al. 2018] T. Claeys, A. B. J. Kuijlaars, K. Liechty, and D. Wang, "Propagation of singular behavior for Gaussian perturbations of random matrices", *Comm. Math. Phys.* **362**:1 (2018), 1–54. MR Zbl

[Dyson 1962] F. J. Dyson, "A Brownian-motion model for the eigenvalues of a random matrix", *J. Math. Phys.* **3** (1962), 1191–1198. MR Zbl

[Erdős and Schnelli 2017] L. Erdős and K. Schnelli, "Universality for random matrix flows with time-dependent density", *Ann. Inst. Henri Poincaré Probab. Stat.* **53**:4 (2017), 1606–1656. MR Zbl

[Erdős and Yau 2015] L. Erdős and H.-T. Yau, "Gap universality of generalized Wigner and β-ensembles", *J. Eur. Math. Soc.* **17**:8 (2015), 1927–2036. MR Zbl

[Erdős and Yau 2017] L. Erdős and H.-T. Yau, *A dynamical approach to random matrix theory*, Courant Lect. Notes Math. **28**, Courant Inst. Math. Sci., New York, 2017. MR Zbl

[Erdős et al. 2018] L. Erdős, T. Krüger, and D. Schröder, "Cusp universality for random matrices I: Local law and the complex Hermitian case", preprint, 2018. arXiv 1809.03971

[Huang and Landon 2016] J. Huang and B. Landon, "Local law and mesoscopic fluctuations of Dyson Brownian motion for general β and potential", preprint, 2016. arXiv 1612.06306

[Landon and Yau 2017] B. Landon and H.-T. Yau, "Convergence of local statistics of Dyson Brownian motion", *Comm. Math. Phys.* **355**:3 (2017), 949–1000. MR Zbl

[Landon and Yau 2017] B. Landon and H.-T. Yau, "Edge statistics of Dyson Brownian motion", preprint, 2017. arXiv 1712.03881

[Landon et al. 2019] B. Landon, P. Sosoe, and H.-T. Yau, "Fixed energy universality of Dyson Brownian motion", *Adv. Math.* **346** (2019), 1137–1332. MR Zbl

[Stein and Weiss 1958] E. M. Stein and G. Weiss, "Fractional integrals on n-dimensional Euclidean space", *J. Math. Mech.* **7** (1958), 503–514. MR Zbl

[Tracy and Widom 2006] C. A. Tracy and H. Widom, "The Pearcey process", *Comm. Math. Phys.* **263**:2 (2006), 381–400. MR Zbl

Received 28 Jan 2019. Revised 17 Jun 2019. Accepted 22 Jul 2019.

GIORGIO CIPOLLONI: giorgio.cipolloni@ist.ac.at
IST Austria, Klosterneuburg, Austria

LÁSZLÓ ERDŐS: lerdos@ist.ac.at
IST Austria, Klosterneuburg, Austria

TORBEN KRÜGER: torben-krueger@uni-bonn.de
University of Bonn, Bonn, Germany

DOMINIK SCHRÖDER: dschroed@ist.ac.at
IST Austria, Klosterneuburg, Austria

PURE and APPLIED ANALYSIS

msp.org/paa

Cover image: The figure shows the outgoing scattered field produced by scattering a plane wave, coming from the northwest, off of the (stylized) letters P A A. The total field satisfies the homogeneous Dirichlet condition on the boundary of the letters. It is based on a numerical computation by Mike O’Neil of the Courant Institute.

See inside back cover or msp.org/paa for submission instructions.

The subscription price for 2019 is US $495/year for the electronic version, and $555/year (+$25, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Pure and Applied Analysis (ISSN 2578-5885 electronic, 2578-5893 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

PAA peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PURE and APPLIED ANALYSIS

vol. 1 no. 4 2019

2578-5893 (2019)1:4;1-3