J. Eur. Math. Soc. 17, 1927–2036

DOI 10.4171/JEMS/548

László Erdős · Horng-Tzer Yau

Gap universality of generalized Wigner and β-ensembles

Received February 5, 2014 and in revised form March 31, 2014 and May 5, 2015

Abstract. We consider generalized Wigner ensembles and general β-ensembles with analytic potentials for any $\beta \geq 1$. The recent universality results in particular assert that the local averages of consecutive eigenvalue gaps in the bulk of the spectrum are universal in the sense that they coincide with those of the corresponding Gaussian β-ensembles. In this article, we show that local averaging is not necessary for this result, i.e. we prove that the single gap distributions in the bulk are universal. In fact, with an additional step, our result can be extended to any $C^4(\mathbb{R})$ potential.

Keywords. β-ensembles, Wigner–Dyson–Gaudin–Mehta universality, gap distribution, log-gas

Contents

1. Introduction 1928
2. Main results 1931
3. Outline of the main ideas in the proof 1935
4. Local equilibrium measures 1938
   4.1. Basic properties of local equilibrium measures 1938
   4.2. Extensions and further results 1942
5. Gap universality for β-ensembles: proof of Theorem 2.3 1943
   5.1. Rigidity bounds and its consequences 1943
   5.2. Completing the proof of Theorem 2.3 1945
6. Gap universality for Wigner matrices: proof of Theorem 2.2 1948
7. Rigidity and level repulsion of local measures 1956
   7.1. Rigidity of $\mu_{\mathbf{y}}$: proof of Theorem 4.2 1956
   7.2. Level repulsion estimates of $\mu_{\mathbf{y}}$: proof of Theorem 4.3 1960
8. Proof of Theorem 4.1 1965
   8.1. Comparison of the local statistics of two local measures 1965
   8.2. Rigidity and level repulsion of the interpolating measure $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ 1968
9. Local statistics of interpolating measures: Proof of Theorem 8.1 1971
   9.1. Outline of the proof of Theorem 8.1 1971
   9.2. Random walk representation 1974
   9.3. Sets of good paths 1979

L. Erdős: Institute of Science and Technology Austria, Am Campus 1, A-3400 Klosterneuburg, Austria; e-mail: lerdos@ist.ac.at

H.-T. Yau: Department of Mathematics, Harvard University, Cambridge MA 02138, USA; e-mail: htyau@math.harvard.edu

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

9.4. Restrictions to the good paths . . . 1983
9.4.1. Restriction to the set $\mathcal{G}$ . . . 1983
9.4.2. Restriction to the set $\widetilde{\mathcal{Q}}$ and the decay estimates . . . 1983
9.5. Short time cutoff and finite speed of propagation . . . 1986
9.6. Proof of the finite speed of propagation estimate, Lemma 9.6 . . . 1987
9.7. Completing the proof of Theorem 8.1 . . . 1989
10. A discrete De Giorgi–Nash–Moser estimate . . . 1992
10.1. Hölder regularity . . . 1995
10.2. Proof of $(\mathbf{ST})_n \Rightarrow (\mathbf{OSC})_n$ . . . 1999
10.2.1. Statement of the generalized De Giorgi lemmas . . . 2000
10.2.2. Verifying the assumptions of Lemma 10.7 . . . 2003
10.2.3. Completing the proof of Proposition 10.5 . . . 2004
10.2.4. Summary of the choice of the parameters . . . 2009
10.3. Proof of $(\mathbf{ST})_n + (\mathbf{OSC})_n \Rightarrow (\mathbf{ST})_{n+1}$ . . . 2010
10.4. Proof of Lemma 10.6 (first De Giorgi lemma) . . . 2011
10.5. Proof of Lemma 10.7 (second De Giorgi lemma) . . . 2019
10.5.1. Dissipation with the good term . . . 2020
10.5.2. Time slices when the good term helps . . . 2022
10.5.3. Finding the intermediate set . . . 2024
References . . . 2034

1. Introduction

The fundamental vision that random matrices can be used as basic models for large quantum systems was due to E. Wigner [60]. He conjectured that the eigenvalue gap distributions of large random matrices were universal ("Wigner surmise") in the sense that large quantum systems and random matrices share the same gap distribution functions. The subsequent work of Dyson, Gaudin and Mehta clarified many related issues regarding this assertion and a thorough understanding of the Gaussian ensembles has thus emerged (see the classical book of Mehta [46] for a summary). There are two main categories of random matrices: the invariant and the noninvariant ensembles. The universality conjecture, which is also known as the Wigner–Dyson–Gaudin–Mehta (WDGM) conjecture, asserts that for both ensembles the eigenvalue gap distributions are universal up to symmetry classes. For invariant ensembles, the joint distribution function of the eigenvalues can be expressed explicitly in terms of one-dimensional particle systems with logarithmic interactions (i.e., *log-gases*) at an inverse temperature β. The values $\beta = 1, 2, 4$ correspond to the classical orthogonal, unitary and symplectic ensembles, respectively. Under various conditions on the external potential, the universality for the classical values $\beta = 1, 2, 4$ was proved, via analysis on the corresponding orthogonal polynomials, by Fokas–Its–Kitaev [37], Deift et al. [16, 19, 20], Bleher–Its [7], Pastur–Shcherbina [49, 50] and in many consecutive works (see e.g. [17, 18, 45, 51, 59]). For nonclassical values of β there is no matrix ensemble behind the model, except for the Gaussian cases [23] via tridiagonal matrices. One may still be interested in the local correlation functions of the log-gas as an interacting particle system. The orthogonal polynomial method is not applicable for nonclassical values of β even for the Gaussian case. For certain special potentials and even integer β, however, there are still explicit formulas for correlation functions [38].

Furthermore, for general β in the Gaussian case the local statistics were described very precisely with a different method by Valkó–Virág [57, 58]. The universality for general β-ensembles for any $\beta > 0$ was established only very recently [8, 9] by a new method based on dynamical methods using Dirichlet form estimates from [29, 30]. This method is important for this article and we will discuss it in more detail later on. All previous results achieved by this method, however, required in their statement to consider a local average of consecutive gaps. In the current paper we will prove universality of *each single* gap in the bulk.

Turning to the noninvariant ensembles, the most important class is the $N \times N$ Wigner matrices characterized by the independence of their entries. In general, there is no longer an explicit expression for the joint distribution function for the eigenvalues. However, there is a special class of ensembles, the Gaussian divisible ensembles, that interpolate between the general Wigner ensembles and the Gaussian ones. For these ensembles, at least in the special Hermitian case, there is still an explicit formula for the joint distribution of the eigenvalues based upon the Harish-Chandra–Itzykson–Zuber integral. This formula was first put into a mathematically useful form by Johansson [43] (see also the later work of Ben Arous–Péché [6]) to prove the universality of Gaussian divisible ensembles with a Gaussian component of size order one. In [26], the size of the Gaussian component needed for proving the universality was greatly reduced to $N^{-1/2+\varepsilon}$. More importantly, the idea of approximating Wigner ensembles by Gaussian divisible ones was first introduced and, after a perturbation argument, this resulted in the first proof of universality for Hermitian ensembles with general smooth distributions for matrix elements. The smoothness condition was later removed in [54, 27].

In his seminal paper [24], Dyson observed that the eigenvalue distribution of Gaussian divisible ensembles is the same as the solution of a special system of stochastic differential equations, commonly known now as the Dyson Brownian motion, at a fixed time t. For short times, t is comparable with the variance of the Gaussian component. He also conjectured that the time to "local equilibrium" of Dyson Brownian motion is of order $1/N$, which is then equivalent to the universality of Gaussian divisible ensembles with a Gaussian component of order slightly larger than $N^{-1/2}$. Thus the work [26] can be viewed as proving Dyson's conjecture for the Hermitian case. This method, however, completely tied with an explicit formula that is so far restricted to the Hermitian case.

A completely analytic approach to estimate the time to local equilibrium of Dyson's Brownian motion was initiated in [29] and further developed in [30, 35, 34] (see [32] for a detailed account). In these papers, Dyson's conjecture in full generality was proved [35] and universality was established for generalized Wigner ensembles for all symmetric classes. The idea of a dynamical approach in proving universality turns out to be a very powerful one. Dyson's Brownian motion can be viewed as the natural gradient flow for Gaussian β log-gases (we will often use the term β *log-gases* for β-ensembles to emphasize the logarithmic interaction). The gradient flow can be defined with respect to all β log-gases, not just the Gaussian ones. Furthermore, one can consider gradient flows of local log-gases with fixed "good boundary conditions". Here "local" refers to Gibbs measures on N^a, $0 < a < 1$, consecutive points of a log-gas with the locations of all other points fixed. By "good boundary conditions" we mean that these external points are *rigid*, i.e. their locations are close to their classical locations given by the limiting density of the original log-gas. Using this idea, we have proved the universality of general β-ensembles in [8, 9] for analytic potentials.

The main conclusion of these works is that the local gap distributions of either the generalized Wigner ensembles (in all symmetry classes) or the general β-ensembles are universal in the bulk of the spectrum (see [31] for a recent review). The dynamical approach based on Dyson's Brownian motion and related flows also provides a conceptual understanding for the origin of the universality. For technical reasons, however, these proofs apply to averages of consecutive gaps, i.e. cumulative statistics of N^{ε} consecutive gaps were proven to be universal. Averaging the statistics of the consecutive gaps is equivalent to averaging the energy parameter in the correlation functions. Thus, mathematically, the results were also formulated in terms of universality of the correlation functions with averaging in an energy window of size $N^{-1+\varepsilon}$.

The main goal of this paper is to remove the local averaging in the statistics of consecutive gaps in our general approach using Dyson's Brownian motion for both invariant and noninvariant ensembles. We will show that the distribution of *each single* gap in the bulk is universal, which we will refer to as the *single gap universality* or simply the *gap universality* whenever there is no confusion. The single gap universality was proved for a special class of Hermitian Wigner matrices with the property that the first four moments of the matrix elements match those of the standard Gaussian random variable [53] and no other results have been known before. In particular, *the single gap universality has not been proved even for the Gaussian orthogonal ensemble (GOE).*

The gap distributions are closely related to the correlation functions which were often used to state the universality of random matrices. These two concepts are equivalent in a certain average sense. However, there is no rigorous relation between correlation functions at a fixed energy and single gap distributions. Thus our results on single gap statistics do not automatically imply the universality of the correlation functions at a fixed energy, which up to very recently was rigorously proved only for Hermitian Wigner matrices [26, 54, 27, 32]; the real symmetry case was proven with a different method in [11].

The removal of a local average in the universality results proved in [29, 33, 34] is a technical improvement in itself and its physical meaning is not especially profound. Our motivation for taking seriously this endeavor is that the single gap distribution may be closely related to the distribution of a single eigenvalue in the bulk of the spectrum [40] or at the edge [55, 56]. Since our approach does not rely on any explicit formula involving Gaussian matrices, some extension of this method may provide a way to understand the distribution of an individual eigenvalue of Wigner matrices. In fact, partly based on the method in this paper, the edge universality for the β-ensembles and generalized Wigner ensembles was established in [10].

The main new idea in this paper is an analysis of the Dyson Brownian motion via parabolic regularity using the De Giorgi–Nash–Moser idea. Since the Hamiltonians of the local log-gases are convex, the correlation functions can be re-expressed in terms of a time average of certain random walks in random environments. The connection between correlation functions of general log-concave measures and random walks in random environments was already pointed out by Helffer and Sjöstrand [42] and Naddaf and

Spencer [47]. This connection was used as an effective way to estimate correlation functions for several models in statistical physics (see e.g. [3, 22, 39, 41, 21]), as well as to remove convexity assumptions in gradient interface models [14, 15].

In this paper we observe that the single gap universality is a consequence of the Hölder regularity of the solutions to these random walk problems. Due to the logarithmic interaction, the random walks are long ranged and their rates may be singular. Furthermore, the random environments themselves depend on the gap distributions, which were exactly the problems we want to analyze! If we view these random walks as (discrete) parabolic equations with random coefficients, we find that they are of divergence form and are in the form of the equations studied in the fundamental paper by Caffarelli, Chan and Vasseur [13]. The main difficulty in applying [13] to gain regularity is that the jump rates in our settings are random and they do not satisfy the uniform upper and lower bounds required in [13]. In fact, in some space-time regime the jump rates can be much more singular than allowed in [13]. To control the singularities of these coefficients, we prove an optimal level repulsion estimate for local log-gases. With these estimates, we are able to extend the method of [13] to prove Hölder regularity for the solution to these random walks problems. This shows that the single gap distributions are universal for local log-gases with good boundary conditions, which is the key result of this paper.

For β-ensembles, it is known that the rigidity of the eigenvalues ensures that boundary conditions are good with high probability. Thus we can apply the local universality of single gap distribution to get the single gap universality of the β-ensembles. We remark, however, that the current result holds only for $\beta \geq 1$ in contrast to $\beta > 0$ in [8, 9], since the current proof heavily relies on the dynamics of the gradient flow of local log-gases.[1] For noninvariant ensembles, a slightly longer argument using the local relaxation flow is needed to connect the local universality result with that for the original Wigner ensemble. This will be explained in Section 6.

In summary, we have recast the question of the single gap universality for random matrices, envisioned by Wigner in the sixties, into a problem concerning the regularity of a parabolic equation in divergence form studied by De Giorgi–Nash–Moser. Thanks to the insight of Dyson and the important progress by Caffarelli–Chan–Vasseur [13], we are able to establish the WDGM universality conjecture for each individual gap via De Giorgi–Nash–Moser's idea. We now introduce our models rigorously and state the main results.

2. Main results

We will have two related results, one concerns the generalized Wigner ensembles, the other one the general β-ensembles. We first define the generalized Wigner ensembles. Let $H = (h_{ij})_{i,j=1}^N$ be an $N \times N$ Hermitian or symmetric matrix where the matrix elements $h_{ij} = \bar{h}_{ji}$, $i \leq j$, are independent random variables given by a probability measure ν_{ij}

[1] During the revision of this manuscript in the refereeing process, gap universality was proved for β-ensembles with $\beta > 0$ under slightly stronger restrictions on the potential in [4, 52].

with mean zero and variance $\sigma_{ij}^2 \geq 0$,

$$\mathbb{E}h_{ij} = 0, \qquad \sigma_{ij}^2 := \mathbb{E}|h_{ij}|^2. \tag{2.1}$$

The distribution ν_{ij} and its variance σ_{ij}^2 may depend on N, but we omit this fact in the notation. We also assume that the normalized matrix elements have a uniform subexponential decay,

$$\mathbb{P}(|h_{ij}| > x\sigma_{ij}) \leq \theta_1 \exp(-x^{\theta_2}), \qquad x > 0, \tag{2.2}$$

with some fixed constants $\theta_1, \theta_2 > 0$, uniformly in N, i, j. In fact, with minor modifications of the proof, an algebraic decay

$$\mathbb{P}(|h_{ij}| > x\sigma_{ij}) \leq C_M x^{-M}$$

with a large enough M is also sufficient.

Definition 2.1 ([33]). The matrix ensemble H defined above is called a *generalized Wigner matrix* if the following assumptions hold on the variances of the matrix elements (2.1):

(A) For any fixed j,

$$\sum_{i=1}^{N} \sigma_{ij}^2 = 1.$$

(B) There exist two positive constants, $C_{\inf}$ and $C_{\sup}$, independent of N such that

$$C_{\inf}/N \leq \sigma_{ij}^2 \leq C_{\sup}/N. \tag{2.3}$$

Let $\mathbb{P}$ and $\mathbb{E}$ denote the probability and the expectation with respect to this ensemble.

We will denote by $\lambda_1 \leq \cdots \leq \lambda_N$ the eigenvalues of H. In the special case when $\sigma_{ij}^2 = 1/N$ and h_{ij} is Gaussian, the joint probability distribution of the eigenvalues is given:

$$\mu = \mu_G^{(N)}(d\boldsymbol{\lambda}) = \frac{e^{-N\beta\mathcal{H}(\boldsymbol{\lambda})}}{Z_\beta} d\boldsymbol{\lambda}, \qquad \mathcal{H}(\boldsymbol{\lambda}) = \sum_{i=1}^{N} \frac{\lambda_i^2}{4} - \frac{1}{N} \sum_{i<j} \log|\lambda_j - \lambda_i|. \tag{2.4}$$

The value of β depends on the symmetry class of the matrix; $\beta = 1$ for GOE, $\beta = 2$ for GUE and $\beta = 4$ for GSE. Here Z_β is the normalization factor so that μ is a probability measure.

It is well known that the density (or one-point correlation function) of μ converges, as $N \to \infty$, to the Wigner semicircle law

$$\varrho(x) := \frac{1}{2\pi}\sqrt{(4 - x^2)_+}. \tag{2.5}$$

We use γ_j for the j-th quantile of this density, i.e. γ_j is defined by

$$\frac{j}{N} = \int_{-2}^{\gamma_j} \varrho_G(x)\, dx. \tag{2.6}$$

We now define a class of test functions. Fix an integer n. We say that $O = O_N : \mathbb{R}^n \to \mathbb{R}$, a possibly N-dependent sequence of differentiable functions, is an *n-particle observable* if

$$O_\infty := \sup_N \|O_N\|_\infty < \infty, \qquad \operatorname{supp} O_N \subset [-O_{\rm supp}, O_{\rm supp}]^n, \tag{2.7}$$

with some finite $O_{\rm supp}$, independent of N, but we allow $\|O_N'\|_\infty$ to grow with N. For any integers $A < B$ we also introduce the notation $[\![A, B]\!] := \{A, A+1, \dots, B\}$.

Our main result on generalized Wigner matrices asserts that the local gap statistics in the bulk of the spectrum are universal for any generalized Wigner matrix, in particular they coincide with those of the Gaussian case.

Theorem 2.2 (Gap universality for Wigner matrices). *Let H be a generalized Wigner ensemble with subexponentially decaying matrix elements (see (2.2)). Fix positive numbers α, O_∞, $O_{\rm supp}$ and an integer $n \in \mathbb{N}$. There exist $\varepsilon, C > 0$, depending only on α, $\mathcal{O}_\infty$ and $O_{\rm supp}$, such that for any n-particle observable $O = O_N$ satisfying (2.7) we have*

$$|[\mathbb{E} - \mathbb{E}^\mu] O(N(\lambda_j - \lambda_{j+1}), N(\lambda_j - \lambda_{j+2}), \dots, N(\lambda_j - \lambda_{j+n}))| \le C N^{-\varepsilon} \|O'\|_\infty \tag{2.8}$$

for any $j \in [\![\alpha N, (1-\alpha)N]\!]$ and for any sufficiently large $N \ge N_0$, where N_0 depends on all parameters of the model, as well as on n, α, $\mathcal{O}_\infty$ and $O_{\rm supp}$.

More generally, for any $k, m \in [\![\alpha N, (1-\alpha)N]\!]$ we have

$$\begin{aligned}\Big|&\mathbb{E} O\big((N\varrho_k)(\lambda_k - \lambda_{k+1}), (N\varrho_k)(\lambda_k - \lambda_{k+2}), \dots, (N\varrho_k)(\lambda_k - \lambda_{k+n})\big)\\ &- \mathbb{E}^\mu O\big((N\varrho_m)(\lambda_m - \lambda_{m+1}), (N\varrho_m)(\lambda_m - \lambda_{m+2}), \dots, (N\varrho_m)(\lambda_m - \lambda_{m+n})\big)\Big|\\ &\le C N^{-\varepsilon} \|O'\|_\infty,\end{aligned} \tag{2.9}$$

where the local density ϱ_k is defined by $\varrho_k := \varrho(\gamma_k)$.

It is well known that the gap distribution of Gaussian random matrices for all symmetry classes can be explicitly expressed via a Fredholm determinant provided that a certain local average is taken [16, 17, 18]. The result for a single gap, i.e. without local averaging, was only achieved recently in the special case of the Gaussian unitary ensemble (GUE) by Tao [53] (which then easily implies the same results for Hermitian Wigner matrices satisfying the four moment matching condition). It is not clear if a similar argument can be applied to the GOE case.

We now define β-ensembles with a general external potential. Let $\beta > 0$ be a fixed parameter. Let $V(x)$ be a real analytic[2] potential on $\mathbb{R}$ that grows faster than $(2+\varepsilon)\log|x|$ at infinity and satisfies

$$\inf_{\mathbb{R}} V'' > -\infty. \tag{2.10}$$

Consider the measure

$$\mu = \mu_{\beta,V}^{(N)}(d\boldsymbol{\lambda}) = \frac{e^{-N\beta\mathcal{H}(\boldsymbol{\lambda})}}{Z_\beta} d\boldsymbol{\lambda}, \qquad \mathcal{H}(\boldsymbol{\lambda}) = \frac{1}{2}\sum_{i=1}^N V(\lambda_i) - \frac{1}{N}\sum_{i<j} \log|\lambda_j - \lambda_i|. \tag{2.11}$$

[2] In fact, $V \in C^4(\mathbb{R})$ is sufficient: see Remark 5.1.

Since μ is symmetric in all its variables, we will mostly view it as a measure restricted to the cone

$$\Xi^{(N)} := \{\boldsymbol{\lambda} : \lambda_1 < \cdots < \lambda_N\} \subset \mathbb{R}^N. \tag{2.12}$$

Note that the Gaussian measure (2.4) is a special case of (2.11) with $V(\lambda) = \lambda^2/2$. In this case we use the notation μ_G for μ.

Let

$$\varrho_1^{(N)}(\lambda) := \mathbb{E}^{\mu} \frac{1}{N} \sum_{j=1}^{N} \delta(\lambda - \lambda_j)$$

denote the density, or the one-point correlation function, of μ. It is well known [1, 12] that $\varrho_1^{(N)}$ converges weakly to the equilibrium density $\varrho = \varrho_V$ as $N \to \infty$. The equilibrium density can be characterized as the unique minimizer (in the set of probability measures on $\mathbb{R}$ endowed with the weak topology) of the functional

$$I(\nu) := \int V(t)\, d\nu(t) - \iint \log|t-s|\, d\nu(t)\, d\nu(s). \tag{2.13}$$

In the case $V(x) = x^2/2$, the minimizer is the Wigner semicircle law $\varrho = \varrho_G$, defined in (2.5), where the subscript G refers to the Gaussian case. In the general case we assume that $\varrho = \varrho_V$ is supported on a single compact interval, $[A, B]$, and $\varrho \in C^2(A, B)$. Moreover, we assume that V is *regular* in the sense that ϱ is strictly positive on (A, B) and vanishes as a square root at the endpoints [9, (1.4)]. It is known that these conditions are satisfied if, for example, V is strictly convex.

For any $j \le N$ define the *classical location* of the j-th particle, $\gamma_{j,V}$, by

$$\frac{j}{N} = \int_A^{\gamma_{j,V}} \varrho_V(x)\, dx; \tag{2.14}$$

for the Gaussian case we have $[A, B] = [-2, 2]$ and we use the notation $\gamma_{j,G} = \gamma_j$ for the corresponding classical location, defined in (2.6). We set

$$\varrho_j^V := \varrho_V(\gamma_{j,V}) \quad \text{and} \quad \varrho_j^G := \varrho_G(\gamma_{j,G}) \tag{2.15}$$

to be the limiting density at the classical location of the j-th particle. Our main theorem on β-ensembles is the following:[3]

Theorem 2.3 (Gap universality for β-ensembles). *Let $\beta \ge 1$ and V be a real analytic*[4] *potential with* (2.10) *such that ϱ_V is supported on a single compact interval, $[A, B]$, $\varrho_V \in C^2(A, B)$, and V is regular. Fix positive numbers α, O_∞, O_{supp}, an integer $n \in \mathbb{N}$ and an n-particle observable $O = O_N$ satisfying* (2.7)*. Let $\mu = \mu_V = \mu_{\beta,V}^{(N)}$ be given by* (2.11) *and let μ_G denote the same measure for the Gaussian case. Then there exist an*

---

[3] During the revision of this manuscript in the refereeing process, gap universality was proved for β-ensembles with $\beta > 0$ under slightly stronger restrictions on the potential in [4, 52]. Furthermore, the fixed energy universality for the β-ensemble was also proved in [52].

[4] In fact, $V \in C^4(\mathbb{R})$ is sufficient: see Remark 5.1.

$\varepsilon > 0$, *depending only on* α, β *and the potential* V, *and a constant* C *depending on* $\mathcal{O}_\infty$ *and* $O_{\rm supp}$ *such that*

$$\begin{aligned}\Big|\mathbb{E}^{\mu_V} O\big((N\varrho_k^V)(\lambda_k-\lambda_{k+1}),(N\varrho_k^V)(\lambda_k-\lambda_{k+2}),\dots,(N\varrho_k^V)(\lambda_k-\lambda_{k+n})\big)\\ -\mathbb{E}^{\mu_G} O\big((N\varrho_m^G)(\lambda_m-\lambda_{m+1}),(N\varrho_m^G)(\lambda_m-\lambda_{m+2}),\dots,(N\varrho_m^G)(\lambda_m-\lambda_{m+n})\big)\Big|\\ \le CN^{-\varepsilon}\|O'\|_\infty \qquad (2.16)\end{aligned}$$

for any $k, m \in [\![\alpha N, (1-\alpha)N]\!]$ *and for any sufficiently large* $N \ge N_0$, *where* N_0 *depends on* V, β, *as well as on* n, α, O_∞ *and* $O_{\rm supp}$. *In particular, the distribution of the rescaled gaps with respect to* μ_V *does not depend on the index* k *in the bulk.*

Theorem 2.3, in particular, asserts that the single gap distribution in the bulk is independent of the index k. The special GUE case of this assertion is the content of [53] where the proof uses some special structures of GUE.

The proofs of both Theorems 2.2 and 2.3 rely on the uniqueness of the gap distribution for a localized version of the equilibrium measure (2.4) with a certain class of boundary conditions. This main technical result will be formulated in Theorem 4.1 in the next section after we introduce the necessary notation. A sketch of the content of the paper will be given at the end of Section 4.1.

We remark that Theorem 2.3 is stated only for $\beta \ge 1$; on the contrary, the universality with local averaging in [8, 9] was proved for $\beta > 0$. The main reason is that the current proof relies heavily on the dynamics of the gradient flow of local log-gases. Hence the well-posedness of the dynamics is crucial, and it is available only for $\beta \ge 1$. On the other hand, in [8, 9] we use only certain Dirichlet form inequalities (see e.g. [8, Lemma 5.9]), which we could prove with an effective regularization scheme for all $\beta > 0$. For $\beta < 1$ it is not clear if such a regularization can also be applied to the new inequalities we will prove here.

3. Outline of the main ideas in the proof

For the orientation of the reader we briefly outline the three main concepts in the proof without any technicalities.

1. Local Gibbs measures and their comparison

The first observation is that the macroscopic structure of the Gibbs measure $\mu_{\beta,V}^{(N)}$ (see (2.11)) heavily depends on V via the density ϱ_V. The microscopic structure, however, is essentially determined by the logarithmic interaction alone—the local density plays only the role of a scaling factor. Once the measure is localized, its dependence on V is reduced to a simple linear rescaling. This gives rise to the idea to consider the *local Gibbs measures*, defined on $\mathcal{K}$ consecutive particles (indexed by a set I) by conditioning on all other $N-\mathcal{K}$ particles. The frozen particles, denoted collectively by $\mathbf{y} = \{y_j\}_{j\notin I}$, play the role of the boundary conditions. The potential of the local Gibbs measure $\mu_{\mathbf{y}}$ is given by $\frac{1}{2}V_{\mathbf{y}}(x) = \frac{1}{2}V(x) - \frac{1}{N}\sum_{j\notin I}\log|x-y_j|$. From the rigidity property of the measure μ (see [8]), the frozen particles are typically very close to their classical locations determined by the appropriate quantiles of the equilibrium density ϱ_V. Moreover, from the Euler–Lagrange equation of (2.13) we have $V(x) = 2\int\log|x-y|\varrho_V(y)\,dy$. These properties together with the choice $\mathcal{K} \ll N$ ensure that $V_{\mathbf{y}}$ is small away from the boundary. Thus, apart from boundary effects, the local Gibbs measure is independent of the original potential V. In particular, its gap statistics can be compared with that of the Gaussian ensemble after an appropriate rescaling. For convenience, we scale all local measures so that the typical size of their gaps is one.

2. Random walk representation of covariance

The key technical difficulty is to estimate the boundary effects, which is given by the correlation between the external potential $\sum_i V_{\mathbf{y}}(x_i)$ and the gap observable $O(x_j - x_{j+1})$ (for simplicity we look at one gap only). We introduce the notation $\langle X; Y\rangle := \mathbb{E}XY - \mathbb{E}X\,\mathbb{E}Y$ to denote the covariance of two random variables X and Y. Following the more customary statistical physics terminology, we will refer to $\langle X; Y\rangle$ as *correlation*. Due to the long range of the logarithmic interaction, the two-point correlation function $\langle\lambda_i;\lambda_j\rangle$ of a log-gas decays only logarithmically in $|i-j|$, i.e. very slowly. What we really need is the correlation between a particle λ_i and a gap $\lambda_j - \lambda_{j+1}$ which decays faster, as $|i-j|^{-1}$, but we need quite precise estimates to exploit the gap structure.

For any Gibbs measure $\omega(d\mathbf{x}) = e^{-\beta\mathcal{H}(\mathbf{x})}d\mathbf{x}$ with strictly convex Hamiltonian, $\mathcal{H}'' \geq c > 0$, the correlation of any two observables F and G can be expressed as

$$\langle F(\mathbf{x}); G(\mathbf{x})\rangle_\omega = \frac{1}{2}\int_0^\infty ds\int d\omega(\mathbf{x})\,\mathbb{E}_{\mathbf{x}}[\nabla G(\mathbf{x}(s))\mathcal{U}(s,\mathbf{x}(\cdot))\nabla F(\mathbf{x})]. \tag{3.1}$$

Here $\mathbb{E}_{\mathbf{x}}$ is the expectation for the (random) paths $\mathbf{x}(\cdot)$ starting from $\mathbf{x}(0) = \mathbf{x}$ and solving the canonical SDE for the measure ω,

$$d\mathbf{x}(s) = d\mathbf{B}(s) - \beta\nabla\mathcal{H}(\mathbf{x}(s))ds,$$

and $\mathcal{U}(s) = \mathcal{U}(s,\mathbf{x}(\cdot))$ is the fundamental solution to the linear system of equations

$$\partial_s\mathcal{U}(s) = -\mathcal{U}(s)\mathcal{A}(s), \qquad \mathcal{A}(s) := \beta\mathcal{H}''(\mathbf{x}(s)), \tag{3.2}$$

with $\mathcal{U}(0) = I$. Notice that the coefficient matrix $\mathcal{A}(s)$, and thus the fundamental solution, depend on the random path $\mathbf{x}(s)$.

If G is a function of the gap, $G(\mathbf{x}) = O(x_j - x_{j+1})$, then (3.1) becomes

$$\begin{aligned}&\langle F(\mathbf{x}); O(x_j - x_{j+1})\rangle_\omega\\ &\quad = \frac{1}{2}\int_0^\infty ds\int d\omega(\mathbf{x})\sum_{i\in I}\mathbb{E}_{\mathbf{x}}[O'(x_j - x_{j+1})(\mathcal{U}_{i,j}(s) - \mathcal{U}_{i,j+1}(s))\partial_i F(\mathbf{x})].\end{aligned} \tag{3.3}$$

We will estimate the correlation (3.3) by showing that for a typical path $\mathbf{x}(\cdot)$ the solution $\mathcal{U}(s)$ is Hölder regular in the sense that $\mathcal{U}_{i,j}(s) - \mathcal{U}_{i,j+1}(s)$ is small if j is away from the boundary and s is not too small. The exceptional cases require various technical cutoff estimates.

3. Hölder regularity of the solution to (3.2)

We will apply (3.3) with the choice $\omega = \mu_{\mathbf{y}}$ and with a function F representing the effects of the boundary conditions. For any fixed realization of the path $\mathbf{x}(\cdot)$, we will view (3.2) as a finite-dimensional version of a parabolic equation. The coefficient matrix, the Hessian of the local Gibbs measure, is computed explicitly. It can be written as $\mathcal{A} = \mathcal{B} + \mathcal{W}$, where $\mathcal{W} \geq 0$ is diagonal and $\mathcal{B}$ is symmetric with quadratic form

$$\langle \mathbf{u}, \mathcal{B}(s)\mathbf{u}\rangle = \frac{1}{2}\sum_{i,j\in I} B_{ij}(s)(u_i - u_j)^2, \qquad B_{ij}(s) := \frac{\beta}{(x_i(s) - x_j(s))^2}.$$

After rescaling the problem and writing it in microscopic coordinates where the gap size is of order one, for a typical path and large $i - j$ we have

$$B_{ij}(s) \sim 1/(i-j)^2 \tag{3.4}$$

by rigidity. We also have a lower bound for any $i \neq j$,

$$B_{ij}(s) \gtrsim 1/(i-j)^2, \tag{3.5}$$

at least with a very high probability. If a matching upper bound were true for any $i \neq j$, then (3.2) would be the discrete analogue of the general equation

$$\partial_t u(t,x) = \int K(t,x,y)[u(t,y) - u(t,x)]\,dy, \qquad t > 0,\ x, y \in \mathbb{R}^d, \tag{3.6}$$

considered by Caffarelli–Chan–Vasseur [13]. It is assumed that the kernel K is symmetric and there is a constant $0 < s < 2$ such that the short distance singularity can be bounded by

$$C_1|x-y|^{-d-s} \leq K(t,x,y) \leq C_2|x-y|^{-d-s} \tag{3.7}$$

for some positive constants C_1, C_2. Roughly speaking, the integral operator corresponds to the behavior of the operator $|p|^s$, where $p = -i\nabla$. The main result of [13] asserts that for any $t_0 > 0$, the solution $u(t,x)$ is ε-Hölder continuous, $u \in C^{\varepsilon}((t_0,\infty), \mathbb{R}^d)$, for some positive exponent ε that depends only on t_0, C_1, C_2. Further generalizations and related local regularity results such as weak Harnack inequality can be found in [36].

Our equation (3.2) is of this type with $d = s = 1$, but it is discrete and in a finite interval I with a potential term. The key difference, however, is that the coefficient $B_{ij}(t)$ in the elliptic part of (3.2) can be singular in the sense that $B_{ij}(t)|i-j|^2$ is not uniformly bounded when i, j are close to each other. In fact, by extending the reasoning of Ben Arous and Bourgade [5], the minimal gap $\min_i(x_{i+1} - x_i)$ for GOE is typically of order $N^{-1/2}$ in the microscopic coordinates we are using now. Thus the analogue of the uniform upper bound (3.7) does not even hold for a fixed t. The only control we can guarantee for the singular behavior of B_{ij} with a large probability is the estimate

$$\sup_{0\leq s\leq\sigma}\ \sup_{0\leq M\leq CK\log K} \frac{1}{1+s}\int_0^s \frac{1}{M}\sum_{i\in I:\,|i-Z|\leq M} B_{i,i+1}(s')\,ds' \leq CK^{\rho} \tag{3.8}$$

with some small exponent ρ and for any $Z \in I$ far away from the boundary of I. This estimate essentially says that the space-time maximal function of $B_{i,i+1}(t)$ at a fixed space-time point $(Z, 0)$ is bounded by K^ρ. Our main generalization of the result in [13] is to show that the weak upper bound (3.8), together with (3.4) and (3.5) (holding up to a factor K^ξ), is sufficient for proving a discrete version of the Hölder continuity at $(Z, 0)$. More precisely, in Theorem 9.8 we essentially show that there exists a $\mathfrak{q} > 0$ such that for any fixed $\sigma \in [K^c, K^{1-c}]$, the solution to (3.2) satisfies

$$\sup_{|j-Z|+|j'-Z|\le\sigma^{1-\alpha}} |\mathcal{U}_{i,j}(\sigma) - \mathcal{U}_{i,j'}(\sigma)| \le CK^\xi\sigma^{-1-\frac{1}{2}\mathfrak{q}\alpha}, \tag{3.9}$$

with any $\alpha \in [0, 1/3]$ if we can guarantee that ρ and ξ are sufficiently small. The exponent $\mathfrak{q}$ is a universal positive number and it plays the role of the Hölder regularity exponent. In fact, to obtain Hölder regularity around one space-time point (Z, σ) as in (3.9), we need to assume the bound (3.8) around several (but not more than $(\log K)^C$) space-time points, which in our applications can be guaranteed with high probability.

Notice that $\mathcal{U}_{i,j}(\sigma)$ decays as σ^{-1}, hence (3.9) provides an additional decay for the discrete derivative. In particular, this guarantees that the ds integration in (3.3) is finite in the most critical intermediate regime $s \in [K^c, CK\log K]$.

The proof of Theorem 9.8 is given in Section 10. In that section we also formulate a Hölder regularity result for initial data in L^∞ (Theorem 10.1), which is the basis of all other results. Readers interested in the pure PDE aspect of our work are referred to Section 10 which can be read independently of the other sections of the paper.

4. Local equilibrium measures

4.1. Basic properties of local equilibrium measures

Fix two small positive numbers, $\alpha, \delta > 0$. Choose two positive integer parameters L, K such that

$$L \in [\![\alpha N, (1-\alpha)N]\!], \qquad N^\delta \le K \le N^{1/4}. \tag{4.1}$$

We consider the parameters L and K to be fixed and often we will not indicate them in the notation. All results will hold for any sufficiently small α, δ and for any sufficiently large $N \ge N_0$, where the threshold N_0 depends on α, δ and maybe on other parameters of the model. Throughout the paper we will use C and c to denote positive constants which, among others, may depend on α, δ and on the constants in (2.2) and (2.3), but we will not emphasize this dependence. Typically C denotes a large generic constant, while c denotes a small one; their values may change from line to line. These constants are independent of K and N, which are the limiting large parameters of the problem, but they may depend on each other. In most cases this interdependence is harmless since it only requires that a fresh constant C be sufficiently large or c be sufficiently small, depending on the size of the previously established generic constants. In some cases, however, the constants are related in a more subtle manner. In this case we will use $C_0, C_1, \dots$ and $c_0, c_1, \dots$ etc. to denote specific constants in order to be able to refer to them along the proof.

For convenience, we set

$$\mathcal{K} := 2K + 1.$$

Denote by $I = I_{L,K} := [\![L - K, L + K]\!]$ the set of $\mathcal{K}$ consecutive indices in the bulk. We will distinguish the inside and outside particles by renaming them as

$$(\lambda_1, \dots, \lambda_N) := (y_1, \dots, y_{L-K-1}, x_{L-K}, \dots, x_{L+K}, y_{L+K+1}, \dots, y_N) \in \Xi^{(N)}. \quad (4.2)$$

Note that the particles keep their original indices. The notation $\Xi^{(N)}$ refers to the simplex (2.12). For short we will write

$$\mathbf{x} = (x_{L-K}, \dots, x_{L+K}) \quad \text{and} \quad \mathbf{y} = (y_1, \dots, y_{L-K-1}, y_{L+K+1}, \dots, y_N).$$

These points are always listed in increasing order, i.e. $\mathbf{x} \in \Xi^{(\mathcal{K})}$ and $\mathbf{y} \in \Xi^{(N-\mathcal{K})}$. We will refer to the y's as the *external* points and to the x's as *internal* points.

We will fix the external points (often called *boundary conditions*) and study the conditional measures on the internal points. We first define the *local equilibrium measure* (or *local measure* for short) on $\mathbf{x}$ with boundary condition $\mathbf{y}$ by

$$\mu_{\mathbf{y}}(d\mathbf{x}) := \mu_{\mathbf{y}}(\mathbf{x}) d\mathbf{x}, \qquad \mu_{\mathbf{y}}(\mathbf{x}) := \mu(\mathbf{y}, \mathbf{x}) \left[\int \mu(\mathbf{y}, \mathbf{x})\, d\mathbf{x} \right]^{-1}, \quad (4.3)$$

where $\mu = \mu(\mathbf{y}, \mathbf{x})$ is the (global) equilibrium measure (2.11) (we do not distinguish between the measure μ and its density function $\mu(\mathbf{y}, \mathbf{x})$ in the notation). Note that for any fixed $\mathbf{y} \in \Xi^{(N-\mathcal{K})}$, all x_j lie in the *open configuration interval*, denoted by

$$J = J_{\mathbf{y}} := (y_{L-K-1}, y_{L+K+1}).$$

Denote by

$$\bar{y} := \tfrac{1}{2}(y_{L-K-1} + y_{L+K+1})$$

the midpoint of the configuration interval. We also introduce

$$\alpha_j := \bar{y} + \frac{j - L}{\mathcal{K} + 1}|J|, \qquad j \in I_{L,K}, \quad (4.4)$$

the $\mathcal{K}$ equidistant points within the interval J.

For any fixed $L, K, \mathbf{y}$, the equilibrium measure can also be written as a Gibbs measure,

$$\mu_{\mathbf{y}} = \mu^{(N)}_{\mathbf{y},\beta,V} = Z_{\mathbf{y}}^{-1} e^{-N\beta \mathcal{H}_{\mathbf{y}}}, \quad (4.5)$$

with Hamiltonian

$$\begin{aligned} \mathcal{H}_{\mathbf{y}}(\mathbf{x}) &:= \sum_{i \in I} \frac{1}{2} V_{\mathbf{y}}(x_i) - \frac{1}{N} \sum_{\substack{i,j \in I \\ i<j}} \log|x_j - x_i|, \\ V_{\mathbf{y}}(x) &:= V(x) - \frac{2}{N} \sum_{j \notin I} \log|x - y_j|. \end{aligned} \quad (4.6)$$

Here $V_{\mathbf{y}}(x)$ can be viewed as the external potential of a β-log-gas of the points $\{x_i : i \in I\}$ in the configuration interval J.

Our main technical result, Theorem 4.1 below, asserts that, for K, L chosen according to (4.1), the local gap statistics are essentially independent of V and $\mathbf{y}$ as long as the boundary conditions $\mathbf{y}$ are regular. This property is expressed by defining the following set of "good" boundary conditions with some given positive parameters ν, α:

$$\begin{aligned}\mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\nu,\alpha) := &\{\mathbf{y} : |y_k - \gamma_k| \le N^{-1+\nu},\ \forall k \in [\![\alpha N, (1-\alpha)N]\!] \setminus I_{L,K}\}\\ &\cap \{\mathbf{y} : |y_k - \gamma_k| \le N^{-4/15+\nu},\ \forall k \in [\![N^{3/5+\nu}, N - N^{3/5+\nu}]\!]\}\\ &\cap \{\mathbf{y} : |y_k - \gamma_k| \le 1,\ k \in [\![1, N]\!] \setminus I_{L,K}\}. \end{aligned} \tag{4.7}$$

In Section 5 we will see that this definition is tailored to the previously proven rigidity bounds for the β-ensemble (see (5.4)). The rigidity bounds for the generalized Wigner matrices are stronger (see (6.1)), so this definition will suit the needs of both proofs.

Theorem 4.1 (Gap universality for local measures). *Fix L, $\widetilde{L}$ and $\mathcal{K} = 2K+1$ satisfying (4.1) with an exponent $\delta > 0$. Consider two boundary conditions $\mathbf{y}, \widetilde{\mathbf{y}}$ such that the configuration intervals coincide,*

$$J = (y_{L-K-1}, y_{L+K+1}) = (\widetilde{y}_{\widetilde{L}-K-1}, \widetilde{y}_{\widetilde{L}+K+1}). \tag{4.8}$$

Consider the measures $\mu = \mu_{\mathbf{y},\beta,V}$ and $\widetilde{\mu} = \mu_{\widetilde{\mathbf{y}},\beta,\widetilde{V}}$ defined as in (4.5), with possibly two different external potentials V and $\widetilde{V}$. Let $\xi > 0$ be a small constant. Assume that

$$|J| = \frac{\mathcal{K}}{N\varrho(\bar{y})} + O\left(\frac{K^{\xi}}{N}\right). \tag{4.9}$$

Suppose that $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2)$ and

$$\max_{j\in I_{L,K}} |\mathbb{E}^{\mu_{\mathbf{y}}} x_j - \alpha_j| + \max_{j\in I_{\widetilde{L},K}} |\mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}} x_j - \alpha_j| \le C N^{-1} K^{\xi}. \tag{4.10}$$

Let the integer p satisfy $|p| \le K - K^{1-\xi^*}$ for some small $\xi^* > 0$. Then there exists $\xi_0 > 0$, depending on δ, such that if $\xi, \xi^* \le \xi_0$ then for any n fixed and any n-particle observable $O = O_N$ satisfying (2.7) with fixed control parameters O_∞ and $O_{\rm supp}$, we have

$$\begin{aligned}&\big|\mathbb{E}^{\mu_{\mathbf{y}}} O\big(N(x_{L+p} - x_{L+p+1}), \dots, N(x_{L+p} - x_{L+p+n})\big)\\ &\quad - \mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}} O\big(N(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+1}), \dots, N(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+n})\big)\big| \le C K^{-\varepsilon} \|O'\|_\infty \end{aligned} \tag{4.11}$$

for some $\varepsilon > 0$ depending on δ, α and for some C depending on O_∞ and $O_{\rm supp}$. This holds for any $N \ge N_0$ sufficiently large, where N_0 depends on the parameters ξ, ξ^*, α, and C in (4.10).

In the following two theorems we establish rigidity and level repulsion estimates for the local log-gas $\mu_{\mathbf{y}}$ with good boundary conditions $\mathbf{y}$. While both rigidity and level repulsion are basic questions for log-gases, our main motivation to prove these theorems is to use them in the proof of Theorem 4.1. The current form of the level repulsion estimate is new; a weaker form was proved in [8, (4.11)]. The rigidity estimate was proved for the global equilibrium measure μ in [8]. From this estimate, one can conclude that $\mu_{\mathbf{y}}$ has a good rigidity bound for a set of boundary conditions with high probability with respect to the global measure μ. However, we will need a rigidity estimate for $\mu_{\mathbf{y}}$ for a set of $\mathbf{y}$'s with high probability with respect to some different measure, which may be asymptotically singular to μ for large N. For example, in the proof for the gap universality of Wigner matrices such a measure is given by the time evolved measure $f_t\mu$ (see Section 6). The following result asserts that a rigidity estimate holds for $\mu_{\mathbf{y}}$ provided that $\mathbf{y}$ itself satisfies a rigidity bound and an extra condition, (4.12), holds. This provides explicit criteria to describe the set of "good" $\mathbf{y}$'s, whose measure with respect to $f_t\mu$ can then be estimated with different methods.

Theorem 4.2 (Rigidity estimate for local measures). *Let L and K satisfy (4.1) with δ the exponent appearing in (4.1). Let ξ,α be any fixed positive constants. For $\mathbf{y}$ in $\mathcal{R}_{L,K}(\xi\delta/2,\alpha)$ consider the local equilibrium measure $\mu_{\mathbf{y}}$ defined in (4.5) and assume that*

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_j - \alpha_j| \le CN^{-1}K^{\xi}, \qquad j \in I = I_{L,K}. \tag{4.12}$$

Then there are positive constants C, c, depending on ξ, such that for any $k \in I$ and $u > 0$,

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \ge uK^{\xi}N^{-1}) \le Ce^{-cu^2}. \tag{4.13}$$

Now we state the level repulsion estimates which will be proven in Section 7.2.

Theorem 4.3 (Level repulsion estimate for local measures). *Let L and K satisfy (4.1) and let ξ,α be any fixed positive constants. Then for $\mathbf{y} \in \mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\xi^2\delta/2,\alpha)$ we have the following estimates:*

(i) [Weak form of level repulsion] *For any $s > 0$,*

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+1} - x_i \le s/N] \le C(Ns)^{\beta+1}, \qquad i \in [\![L-K-1, L+K]\!], \tag{4.14}$$

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+2} - x_i \le s/N] \le C(Ns)^{2\beta+1} \qquad i \in [\![L-K-1, L+K-1]\!]. \tag{4.15}$$

(*Here we use the convention that $x_{L-K-1} := y_{L-K-1}, x_{L+K+1} := y_{L+K+1}$.*)

(ii) [Strong form of level repulsion] *Suppose that there exist positive constants C, c such that the following rigidity estimate holds for any $k \in I$:*

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \ge CK^{\xi^2}N^{-1}) \le C\exp(-K^c). \tag{4.16}$$

Then there exists a small constant θ, depending on C, c in (4.16), such that for any $s \ge \exp(-K^{\theta})$,

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+1} - x_i \le s/N] \le C(K^{\xi}s\log N)^{\beta+1}, \qquad i \in [\![L-K-1, L+K]\!], \tag{4.17}$$

$$\mathbb{P}^{\mu_{\mathbf{y}}}[x_{i+2} - x_i \le s/N] \le C(K^{\xi}s\log N)^{2\beta+1}, \qquad i \in [\![L-K-1, L+K-1]\!]. \tag{4.18}$$

We remark that the estimates (4.18) and (4.15) on the second gap are not needed for the main proof; we listed them only for possible reference. The exponents are not optimal; one would expect them to be $3\beta + 3$. With some extra work, it should not be difficult to get the optimal exponents. Moreover, our results can be extended to $x_{i+k} - x_i$ for any k finite. We also mention that the assumption (4.16) required in part (ii) is weaker than what we prove in (4.13). In fact, the weaker form (4.16) of the rigidity would be enough throughout the paper except at one place, at the end of the proof of Lemma 8.4.

Theorem 4.1 is our key result. In Sections 5 and 6 we will show how to use Theorem 4.1 to prove the main Theorems 2.2 and 2.3. Although the basic structure of the proof of Theorem 2.3 is similar to the one given in [8] where a locally averaged version of this theorem was proved under a locally averaged version of Theorem 4.1, here we have to verify the assumption (4.10), which will be done in Lemma 5.2. The proof of Theorem 2.2, on the other hand, is very different from the recent proof of universality in [33, 34]. This will be explained in Section 6.

The proofs of the auxiliary Theorems 4.2 and 4.3 will be given in Section 7. The proof of Theorem 4.1 will start from Section 8.1 and will continue until the end of the paper. At the beginning of Section 8.1 we will explain the main ideas of the proof. The readers interested in the proof of Theorem 4.1 can skip Sections 5 and 6.

4.2. *Extensions and further results*

We formulated Theorems 4.1–4.3 with assumptions requiring that the boundary conditions $\mathbf{y}$ are "good". In fact, all these results hold in a more general setting.

Definition 4.4. An external potential U of a β-log-gas of K points in a configuration interval $J = (a, b)$ is called K^ξ*-regular* if the following bounds hold:

$$|J| = \frac{\mathcal{K}}{N\varrho(\bar{y})} + O\left(\frac{K^\xi}{N}\right), \tag{4.19}$$

$$U'(x) = \varrho(\bar{y}) \log \frac{d_+(x)}{d_-(x)} + O\left(\frac{K^\xi}{Nd(x)}\right), \qquad x \in J, \tag{4.20}$$

$$U''(x) \geq \inf V'' + \frac{c}{d(x)}, \qquad x \in J, \tag{4.21}$$

with some positive $c > 0$ and for some small $\xi > 0$, where

$$d(x) := \min\{|x-a|, |x-b|\}$$

is the distance to the boundary of J and

$$d_-(x) := d(x) + \varrho(\bar{y})N^{-1}K^\xi, \qquad d_+(x) := \max\{|x-a|, |x-b|\} + \varrho(\bar{y})N^{-1}K^\xi.$$

The following lemma, proven in Appendix A, asserts that "good" boundary conditions $\mathbf{y}$ give rise to a regular external potential $V_{\mathbf{y}}$.

Lemma 4.5. *Let L and K satisfy (4.1) and let δ be the exponent appearing in (4.1). Then for any $\mathbf{y} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha/2)$ the external potential $V_{\mathbf{y}}$ (4.6) on the configuration interval $J_{\mathbf{y}}$ is K^ξ-regular:*

$$|J_{\mathbf{y}}| = \frac{\mathcal{K}}{N\varrho(\bar{y})} + O\left(\frac{K^\xi}{N}\right), \tag{4.22}$$

$$V_{\mathbf{y}}'(x) = \varrho(\bar{y}) \log \frac{d_+(x)}{d_-(x)} + O\left(\frac{K^\xi}{Nd(x)}\right), \qquad x \in J_{\mathbf{y}}, \tag{4.23}$$

$$V_{\mathbf{y}}''(x) \geq \inf V'' + \frac{c}{d(x)}, \qquad x \in J_{\mathbf{y}}. \tag{4.24}$$

Remark. The proofs of Theorems 4.1–4.3 and 7.3 do not use the explicit form of $V_{\mathbf{y}}$ and $J_{\mathbf{y}}$; they only depend on the property that $V_{\mathbf{y}}$ on $J_{\mathbf{y}}$ is regular.

5. Gap universality for β-ensembles: proof of Theorem 2.3

5.1. Rigidity bounds and its consequences

The aim of this section is to use Theorem 4.1 to prove Theorem 2.3. In order to verify the assumptions of Theorem 4.1, we first recall the rigidity estimate with respect to μ defined in (2.11). Recall that $\gamma_k = \gamma_{k,V}$ denotes the classical location of the k-th point (see (2.14)). For the case of convex potential, in [8, Theorem 3.1] it was proved that for any fixed $\alpha, \nu > 0$, there are constants $C_0, c_1, c_2 > 0$ such that for any $N \geq 1$ and $k \in [\![\alpha N, (1-\alpha)N]\!]$,

$$\mathbb{P}^\mu\left(|\lambda_k - \gamma_k| > N^{-1+\nu}\right) \leq C_0 \exp\left(-c_1 N^{c_2}\right). \tag{5.1}$$

The same estimate holds for the nonconvex case (see [9, Theorem 1.1]) by using a convexification argument.

Near the spectral edges, a somewhat weaker control was proven for the convex case: Lemma 3.6 of [8] states that for any $\nu > 0$ there are $C_0, c_1, c_2 > 0$ such that

$$\mathbb{P}^\mu\left(|\lambda_k - \gamma_k| > N^{-4/15+\nu}\right) \leq C_0 \exp\left(-c_1 N^{c_2}\right) \tag{5.2}$$

for any $N^{3/5+\nu} \leq k \leq N - N^{3/5+\nu}$, if $N \geq N_0(\nu)$ is sufficiently large. We can choose C_0, c_1, c_2 to be the same in (5.1) and (5.2). Combining this result with the convexification argument in [9], one can show that the estimate (5.2) also holds for the nonconvex case.

Finally, we have a very weak control that holds for all points (see [9, (1.7)]): for any $C > 0$ there are positive constants C_0, c_1 and c_2 such that

$$\mathbb{P}^\mu\left(|\lambda_k - \gamma_k| > C\right) \leq C_0 \exp\left(-c_1 N^{c_2}\right). \tag{5.3}$$

Given C, we can choose C_0, c_1, c_2 to be the same in (5.1)–(5.3).

The set $\mathcal{R}_{L,K}$ in (4.7) was exactly defined as the set of events that these three rigidity estimates hold. From (5.1)–(5.3) we have

$$\mathbb{P}^\mu\left(\mathcal{R}_{L,K}(\nu, \alpha)\right) \geq 1 - C_0 \exp\left(-c_1 N^{c_2}\right) \tag{5.4}$$

for any $\nu, \alpha > 0$ with some positive constants C_0, c_1, c_2 that depend on ν and α.

Remark 5.1. The real analyticity of V in this paper is used only to obtain the rigidity results (5.1)–(5.3) by applying earlier results from [8, 9]. After the first version of the current work appeared in the arXiv, jointly with P. Bourgade we proved the following stronger rigidity result [10, Theorem 2.4]). For any $\beta, \xi > 0$ and $V \in C^4(\mathbb{R})$, regular with equilibrium density supported on a single interval $[A, B]$, there are $c > 0$ and N_0 such that

$$\mathbb{P}^{\mu}(|\lambda_k - \gamma_k| > N^{-2/3+\xi}(\hat{k})^{-1/3}) \le e^{-N^c}, \qquad \forall k \in [\![1, N]\!],$$

for any $N \ge N_0$, where $\hat{k} = \min\{k, N + 1 - k\}$. This result allows us to relax the original real analyticity condition on $V \in C^4(\mathbb{R})$. It would also allow us to redefine the set $\mathcal{R}_{L,K}$ in (4.7) to the more transparent set appearing in (4.13), but this generalization does not affect the rest of the proof.

Lemma 5.2. *Let L and K satisfy (4.1) and let δ be the exponent in (4.1). Then for any small ξ and α there exists a set $\mathcal{R}^* = \mathcal{R}^*_{L,K,\mu}(\xi^2\delta/2, \alpha/2) \subset \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2)$ such that*

$$\mathbb{P}^{\mu}(\mathcal{R}^*) \ge 1 - C_0 \exp\big(-\tfrac{1}{2}c_1 N^{c_2}\big) \tag{5.5}$$

with the constants C_0, c_1, c_2 from (5.1). Moreover, for any $\mathbf{y} \in \mathcal{R}^*$ we have

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_k - \alpha_k| \le C N^{-1} K^{\xi}, \qquad k \in I_{L,K}, \tag{5.6}$$

where α_k was defined in (4.4).

Proof. For any $\nu > 0$ define

$$\mathcal{R}^*_{L,K,\mu}(\nu, \alpha) := \big\{\mathbf{y} \in \mathcal{R}(\nu, \alpha) : \mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \gamma_k| > N^{-1+\nu}) \le \exp\big(-\tfrac{1}{2}c_1 N^{c_2}\big),\ \forall k \in I_{L,K}\big\} \tag{5.7}$$

with the ν-dependent constants $c_1, c_2 > 0$ from (5.4). Note that $\mathcal{R}^*$, unlike $\mathcal{R}$, depends on the underlying measure μ through the family of its conditional measures $\mu_{\mathbf{y}}$. Applying (5.4) for $\nu = \xi^2\delta/2$ and setting $\mathcal{R} = \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2)$, $\mathcal{R}^* = \mathcal{R}^*_{L,K,\mu}(\xi^2\delta/2, \alpha/2)$, we have

$$\mathbb{P}^{\mu}(\mathcal{R}^*) \ge 1 - C_0 \exp\big(-\tfrac{1}{2}c_1 N^{c_2}\big)$$

with some C_0, c_1, c_2. Now if $\mathbf{y} \in \mathcal{R}^*$, then

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_k - \gamma_k| \le C_0 e^{-c_1 N^{c_2}/3} + C N^{-1} K^{\xi^2}, \qquad k \in I_{L,K}. \tag{5.8}$$

In order to prove (5.6), it remains to show that $|\alpha_k - \gamma_k|$ is bounded by $CN^{-1}K^{\xi}$ for any $k \in I_{L,K}$. To see this, we can use the fact that $\varrho \in C^1$ away from the edge, thus $\varrho(x) = \varrho(\bar{y}) + O(x - \bar{y})$ (recall that $\bar{y}$ is the midpoint of J). By Taylor expansion,

$$\begin{aligned} k - (L - K - 1) &= N \int_{\gamma_{L-K-1}}^{\gamma_k} \varrho = N \int_{y_{L-K-1}}^{\gamma_k} \varrho + O(N^{\xi\delta/2}) \\ &= N|\gamma_k - y_{L-K-1}|\varrho(\bar{y}) + O(N|J|^2 + N^{\xi\delta/2}), \end{aligned}$$

i.e.

$$\gamma_k = y_{L-K-1} + \frac{k-L+K+1}{N\varrho(\bar{y})} + O(N^{-1}K^{\xi}). \tag{5.9}$$

Here we have used the fact that $J = J_{\mathbf{y}}$ satisfies (4.22), since $\mathbf{y} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2) \subset \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$. Comparing (5.9) with the definition (4.4) of α_k, and using (4.22) and the fact that $\bar{y} - y_{L-K-1} = \frac{1}{2}|J|$, we have

$$|\alpha_k - \gamma_k| \le CN^{-1}K^{\xi}. \tag{5.10}$$

Together with (5.8) this implies (5.6). □

5.2. *Completing the proof of Theorem 2.3*

We first notice that it is sufficient to prove Theorem 2.3 for the special case $m = N/2$, i.e. when the local statistics for the Gaussian measure is considered at the central point of the spectrum. Indeed, once Theorem 2.3 is proved for any V, k and $m = N/2$, then with the choice $V(x) = x^2/2$ we can use it to establish that the local statistics for the Gaussian measure around any fixed index k in the bulk coincide with the local statistics in the middle. So from now on we assume $m = N/2$, but we keep the notation m for simplicity.

Given k and $m = N/2$ as in (2.16), we first choose $L, \widetilde{L}, K$, satisfying (4.1) (maybe with a smaller α than given in Theorem 2.3), so that $k = L + p$, $m = \widetilde{L} + p$ for some $|p| \le K/2$. In particular

$$|\widetilde{L} - N/2| \le K. \tag{5.11}$$

For brevity, we use $\mu = \mu_V$ and $\widetilde{\mu} = \mu^G$ in accordance with Theorem 4.1.

We consider $\mathbf{y} \in R^*_{L,K,\mu}(\xi^2\delta/2, \alpha)$ and $\widetilde{\mathbf{y}} \in R^*_{\widetilde{L},K,\widetilde{\mu}}(\xi^2\delta/2, \alpha)$, where δ is the exponent in (4.1). We omit the arguments and recall that

$$\mu(R^*_{L,K,\mu}) \ge 1 - C_0\exp\big(-\tfrac{1}{2}c_1N^{c_2}\big), \qquad \widetilde{\mu}(R^*_{\widetilde{L},K,\widetilde{\mu}}) \ge 1 - C_0\exp\big(-\tfrac{1}{2}c_1N^{c_2}\big) \tag{5.12}$$

with some positive constants.

Proposition 5.3. *With the above choice of the parameters, for any* $\mathbf{y} \in R^*_{L,K,\mu}(\xi^2\delta/2, \alpha)$ *and* $\widetilde{\mathbf{y}} \in R^*_{\widetilde{L},K,\widetilde{\mu}}(\xi^2\delta/2, \alpha)$, *we have*

$$\begin{aligned}\Big|\mathbb{E}^{\mu_{\mathbf{y}}} O\big((N\varrho^V_{L+p})(x_{L+p} - x_{L+p+1}), \ldots, (N\varrho^V_{L+p})(x_{L+p} - x_{L+p+n})\big) \\ - \mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}} O\big((N\varrho^G_{\widetilde{L}+p})(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+1}), \ldots, (N\varrho^G_{\widetilde{L}+p})(x_{\widetilde{L}+p} - x_{\widetilde{L}+p+n})\big)\Big| \\ \le CK^{-\varepsilon}\|O'\|_{\infty},\end{aligned} \tag{5.13}$$

where ε is from Theorem 4.1.

Theorem 2.3 follows immediately from (5.12) and this proposition. □

Proof of Proposition 5.3. We will apply Theorem 4.1, but first we have to bring the two measures to the same configuration interval J in order to satisfy (4.8). This will be done in three steps. First, using the scale invariance of the Gaussian log-gas we rescale it so that the local density approximately matches that of μ_V. This will guarantee that the two configuration intervals have almost the same length. In the second step we adjust the local Gaussian log-gas $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$ so that $J_{\widetilde{\mathbf{y}}}$ has exactly the correct length. Finally, we shift the two intervals so that they coincide. This allows us to apply Theorem 4.1 to conclude that the local statistics are identical.

The local densities ϱ_V around $\gamma_{L+p,V}$ and ϱ_G around $\gamma_{\widetilde{L}+p,G}$ may differ considerably. So in the first step we rescale the Gaussian log-gas so that

$$\varrho_V(\gamma_{L+p,V}) = \varrho_G(\gamma_{\widetilde{L}+p,G}). \tag{5.14}$$

To do so, recall that we defined the Gaussian log-gas with the standard $V(x) = x^2/2$ external potential, but we could choose $V_s(x) = s^2x^2/2$ with any fixed $s > 0$ and consider the Gaussian log-gas

$$\mu_G^s(\boldsymbol{\lambda}) \sim \exp(-N\beta\mathcal{H}_s(\boldsymbol{\lambda})), \qquad \mathcal{H}_s(\boldsymbol{\lambda}) := \frac{1}{2}\sum_{i=1}^N V_s(\lambda_i) - \frac{1}{N}\sum_{i<j}\log|\lambda_j - \lambda_i|.$$

This results in rescaling the semicircle density ϱ_G to $\varrho_G^s(x) := s\varrho_G(sx)$ and $\gamma_{i,G}$ to $\gamma_{i,G}^s := s^{-1}\gamma_{i,G}$ for any i, so $\varrho_G(\gamma_{i,G})$ gets rescaled to $\varrho_G^s(\gamma_{i,G}^s) = s\varrho_G(\gamma_{i,G})$. In particular, $\varrho_G(\gamma_{\widetilde{L}+p,G})$ is rescaled to $s\varrho_G(\gamma_{\widetilde{L}+p,G})$, and thus choosing s appropriately, we can achieve that (5.14) holds (keeping the left hand side fixed). Set

$$\mathcal{O}_s(\mathbf{x}) := O\big((N\varrho_G^s(\gamma_{m,G}^s))(x_m - x_{m+1}), \dots, (N\varrho_G^s(\gamma_{m,G}^s))(x_m - x_{m+n})\big), \qquad m = \widetilde{L} + p,$$

and notice that $\mathcal{O}_s(\mathbf{x}) = \mathcal{O}(s\mathbf{x})$. This means that the local gap statistics $\mathbb{E}^{\mu_G^s}\mathcal{O}_s$ is independent of the scaling parameter s, since the product $(N\varrho_m^G)(x_m - x_{m+a})$ (notation defined in (2.15)) is unchanged under the scaling. So we can work with the rescaled Gaussian measure. For notational simplicity we will not carry the s parameter further and we just assume that (5.14) holds with the original Gaussian $V(x) = x^2/2$.

We have now achieved that the two densities coincide at some points of the configuration intervals, but the lengths of these two intervals still slightly differ. In the second step we match them exactly. Since $\mathbf{y} \in R_{L,K}(\xi\delta/2, \alpha)$ and $\widetilde{\mathbf{y}} \in R_{\widetilde{L},K}(\xi\delta/2, \alpha)$, from (4.22) we see that

$$|J_{\mathbf{y}}| = |y_{L+K+1} - y_{L-K-1}| = \frac{\mathcal{K}}{N\varrho_V(\bar{y})} + O(N^{-1}K^{\xi}), \tag{5.15}$$

$$|J_{\widetilde{\mathbf{y}}}| = |\widetilde{y}_{L+K+1} - \widetilde{y}_{L-K-1}| = \frac{\mathcal{K}}{N\varrho_G(\bar{\widetilde{y}})} + O(N^{-1}K^{\xi}). \tag{5.16}$$

Since ϱ_V is C^1, for any $|j| \le K$ we have

$$\begin{aligned}|\varrho_V(\bar{y}) - \varrho_V(\gamma_{L+j,V})| &\le C|\bar{y} - \gamma_{L+j,V}| \\ &\le C|\bar{y} - y_{L,V}| + C|\gamma_{L+j,V} - \gamma_{L,V}| + O(N^{-1}K^{\xi}) \le CKN^{-1},\end{aligned}$$

and similarly for $\varrho_G(\bar{\widetilde{y}})$.

Using (5.15), (5.14) and the fact that the densities are separated away from zero, we easily see that

$$s := |J_{\mathbf{y}}|/|J_{\widetilde{\mathbf{y}}}| \quad \text{satisfies} \quad s = s_{\mathbf{y},\widetilde{\mathbf{y}}} = 1 + O(K^{-1+\xi}). \tag{5.17}$$

Note that this s is different from the scaling parameter in the first step, but it will play a similar role so we use the same notation. For each fixed $\mathbf{y}, \widetilde{\mathbf{y}}$ we can now scale the conditional Gaussian log-gas $\mu_{\widetilde{\mathbf{y}}}$ by a factor of s, i.e. change $\widetilde{\mathbf{y}}$ to $s\widetilde{\mathbf{y}}$, so that after rescaling $|J_{\mathbf{y}}| = |J_{s\widetilde{\mathbf{y}}}|$.

We will now show that this rescaling does not alter the gap statistics:

Lemma 5.4. *Suppose that s satisfies (5.17) and let $\mu = \mu_G$ be the Gaussian log-gas. Then*

$$|[\mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}} - \mathbb{E}^{\mu_{s\widetilde{\mathbf{y}}}}]\mathcal{O}(\mathbf{x})| \leq CK^{-1+\xi} \tag{5.18}$$

with

$$\mathcal{O}(\mathbf{x}) := O\big((N\varrho_m^G)(x_m - x_{m+1}), \ldots, (N\varrho_m^G)(x_m - x_{m+n})\big)$$

for any $\widetilde{L} - K \leq m \leq \widetilde{L} + K - n$ (note that the observable is not rescaled).

Proof. Define the Gaussian log-gas $\mu_{\widetilde{\mathbf{y}}}^s \sim e^{-N\beta\mathcal{H}_{\widetilde{\mathbf{y}}}^s}$ with $\mathcal{H}_{\widetilde{\mathbf{y}}}^s$ defined exactly as in (4.6) but $V_{\mathbf{y}}(x)$ replaced with

$$V_{\widetilde{\mathbf{y}}}^s(x) = V_s(x) - \frac{2}{N}\sum_{j \notin \widetilde{I}} \log|x - \widetilde{y}_j|, \qquad V_s(x) = \tfrac{1}{2}s^2x^2, \qquad \widetilde{I} := [\![\widetilde{L} - K, \widetilde{L} + K]\!].$$

Then by scaling

$$\mathbb{E}^{\mu_{s\widetilde{\mathbf{y}}}}\mathcal{O}(\mathbf{x}) = \mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}^s}\mathcal{O}(\mathbf{x}/s) = \mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}^s}\mathcal{O}(\mathbf{x}) + O(\|O'\|_\infty|s-1|), \tag{5.19}$$

where in the last step we have used the fact that the observable O is a differentiable function with compact support. The error term is negligible by (5.17) and (4.1).

In order to control $[\mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}^s} - \mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}}]\mathcal{O}(\mathbf{x})$, it is sufficient to bound the relative entropy $S(\mu_{\widetilde{\mathbf{y}}}^s|\mu_{\widetilde{\mathbf{y}}})$. However, for any $\mathbf{y} \in R_{L,K}$ we have

$$\mathcal{H}_{\mathbf{y}}'' \geq \min_{x \in J_{\mathbf{y}}} \frac{1}{N}\sum_{j \notin I} \frac{1}{|x - y_j|^2} \geq \frac{cN}{K} \tag{5.20}$$

with a positive constant. Applying this for $\widetilde{\mathbf{y}}$, we see that $\mu_{\widetilde{\mathbf{y}}}$ satisfies the *logarithmic Sobolev inequality* (LSI)

$$S(\mu_{\widetilde{\mathbf{y}}}^s|\mu_{\widetilde{\mathbf{y}}}) \leq \frac{CK}{N} D(\mu_{\widetilde{\mathbf{y}}}^s|\mu_{\widetilde{\mathbf{y}}}),$$

where

$$S(\mu|\omega) := \int \left(\frac{d\mu}{d\omega}\log\frac{d\mu}{d\omega}\right)d\omega, \qquad D(\mu|\omega) := \frac{1}{2N}\int \left|\nabla\sqrt{\frac{d\mu}{d\omega}}\right|^2 d\omega$$

is the relative entropy and the relative Dirichlet form of two probability measures. Therefore

$$S(\mu_{\widetilde{\mathbf{y}}}^{s}|\mu_{\widetilde{\mathbf{y}}}) \leq \frac{CK}{N^2}\mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}}\sum_{i\in\widetilde{I}}|NV_s'(x_i)-NV'(x_i)|^2 = CK(s^2-1)^2\mathbb{E}^{\mu_{\widetilde{\mathbf{y}}}}\sum_{i\in\widetilde{I}}x_i^2$$
$$\leq CK^4N^{-2}(s-1)^2.$$

In the last step we have used (5.11), which, by rigidity for the Gaussian log-gas, guarantees that $|x_i| \leq CK/N$ with very high probability for any $i \in \widetilde{I}$. Together with (5.19) and (5.17) we obtain (5.18). □

Summarizing, we can from now on assume that (5.14) holds and $|J_{\mathbf{y}}| = |J_{\widetilde{\mathbf{y}}}|$. By a straightforward shift we can also assume that $J_{\mathbf{y}} = J_{\widetilde{\mathbf{y}}}$ so that condition (4.8) is satisfied. Condition (4.9) has already been proved in Lemma 4.5. Condition (4.10) follows from the definition of $\mathcal{R}^*_{L,K,\mu}$ and $\mathcal{R}^*_{\widetilde{L},K,\widetilde{\mu}}$ (see Lemma 5.2). Thus all conditions of Theorem 4.1 are verified. Finally, the multiplicative factors ϱ^V_{L+p} and $\varrho^G_{\widetilde{L}+p}$ in (5.13) coincide by (5.14) and (2.15). Then Theorem 4.1 (with an observable O rescaled by the common factor $\varrho^V_{L+p} = \varrho^G_{\widetilde{L}+p}$) implies Proposition 5.3. □

6. Gap universality for Wigner matrices: proof of Theorem 2.2

In our recent results on the universality of Wigner matrices [29, 33, 34], we established the universality for Gaussian divisible matrices by establishing the local ergodicity of the Dyson Brownian motion (DBM). By *local ergodicity* we meant an effective estimate on the time to equilibrium for a local average of observables depending on the gap. In fact, we gave an almost optimal estimate on this time. Then we used the Green function comparison theorem to connect Gaussian divisible matrices to general Wigner matrices. The local ergodicity of DBM was shown by studying the flow of the global Dirichlet form. The estimate on the global Dirichlet form in all these works was sufficiently strong to imply "ergodicity for locally averaged observables" without having to go through local equilibrium measures. In an earlier work [28], however, we used an approach common in the hydrodynamical limits by studying the properties of local equilibrium measures. Since by Theorem 4.1 we know the local equilibrium measures very well, we will now combine the virtues of both methods to prove Theorem 2.2. To explain the new method we will be using, we first recall the standard approach to universality from [29, 33, 34], which consists of the following three steps:

(i) Rigidity estimates on the precise location of eigenvalues.
(ii) Dirichlet form estimates and local ergodicity of DBM.
(iii) Green function comparison theorem to remove the small Gaussian convolution.

In order to prove single gap universality, we will need to apply a similar strategy for the local equilibrium measure $\mu_{\mathbf{y}}$. However, apart from establishing rigidity for $\mu_{\mathbf{y}}$, we will need to strengthen Step (ii). The idea is to use Dirichlet form estimates as in the previous approach, but then apply these estimates to show that the "local structure" after the evolution of the DBM for a short time is characterized by the local equilibrium $\mu_{\mathbf{y}}$ in a strong sense, i.e. without averaging. Since Theorem 4.1 provides single gap universality for the local equilibrium $\mu_{\mathbf{y}}$, this proves single gap universality after a short time DBM evolution, and thus yields the strong form of Step (ii) without averaging the observables. Notice that the key input here is Theorem 4.1 which contains an effective estimate on the time to equilibrium for each single gap. We will call this property the *strong local ergodicity of DBM*. In particular, our result shows that the local averaging taken in our previous works is not essential.

We now recall the rigidity estimate which asserts that the eigenvalues $\lambda_1, \dots, \lambda_N$ of a generalized Wigner matrix follow the Wigner semicircle law $\varrho_G(x)$ (2.5) in a very strong local sense. More precisely, Theorem 2.2 of [35] states that the eigenvalues are near their classical locations, $\{\gamma_j\}_{j=1}^N$, (2.6), in the sense that

$$\mathbb{P}\{\exists j : |\lambda_j - \gamma_j| \ge (\log N)^{\zeta} [\min(j, N-j+1)]^{-1/3} N^{-2/3}\} \le C \exp[-c(\log N)^{\phi\zeta}] \tag{6.1}$$

for any exponent ζ satisfying

$$A_0 \log\log N \le \zeta \le \frac{\log(10N)}{10\log\log N},$$

where the positive constants C, ϕ, A_0 depend only on $C_{\inf}, C_{\sup}, \theta_1, \theta_2$ (see (2.2), (2.3)). In particular, for any fixed $\alpha, \nu > 0$, there are constants $C_0, c_1, c_2 > 0$ such that for any $N \ge 1$ and $k \in [\![\alpha N, (1-\alpha)N]\!]$ we have

$$\mathbb{P}(|\lambda_k - \gamma_k| > N^{-1+\nu}) \le C_0 \exp(-c_1 N^{c_2}), \tag{6.2}$$

and (6.1) also implies

$$\mathbb{E}\sum_{k=1}^N (\lambda_k - \gamma_k)^2 \le N^{-1+2\nu} \tag{6.3}$$

for any $\nu > 0$. The constants C_0, c_1, c_2 may be different from the ones in (5.1) but they play a similar role so we keep their notation. With a slight abuse of notation, we introduce the set $\mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\xi, \alpha)$ from (4.7) in the generalized Wigner setup as well, just γ_k denote the classical locations with respect to the semicircle law (see (2.6)). In particular (6.1) implies that for any $\xi, \alpha > 0$,

$$\mathbb{P}(\mathcal{R}_{L,K}(\xi, \alpha)) \ge 1 - C_0 \exp(-c_1 N^{c_2}) \tag{6.4}$$

with some positive constants C_0, c_1, c_2, analogously to (5.4). We remark that the rigidity bound (6.1) for the generalized Wigner matrices is optimal (up to logarithmic factors) throughout the spectrum and it gives a stronger control than the estimate used in the intermediate regime in the second line of the definition (4.7). For the forthcoming argument the weaker estimates are sufficient, so for notational simplicity we will not modify the definition of $\mathcal{R}$.

The Dyson Brownian motion (DBM) describes the evolution of the eigenvalues of a flow of Wigner matrices, $H = H_t$, if each matrix element h_{ij} evolves according to independent (up to symmetry restriction) Brownian motions. The dynamics of the matrix elements are given by an Ornstein–Uhlenbeck (OU) process which leaves the standard Gaussian distribution invariant. In the Hermitian case, the OU process for the rescaled matrix elements $v_{ij} := N^{1/2}h_{ij}$ is given by the stochastic differential equation

$$dv_{ij} = d\beta_{ij} - \tfrac{1}{2}v_{ij}dt, \quad i, j = 1, \dots, N,$$

where β_{ij}, $i < j$, are independent complex Brownian motions with variance one, and β_{ii} are real Brownian motions of the same variance. The real symmetric case is analogous, just β_{ij} are real Brownian motions.

Denote the distribution of the eigenvalues $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_N)$ of H_t at time t by $f_t(\boldsymbol{\lambda})\mu(d\boldsymbol{\lambda})$ where the Gaussian measure μ is given by (2.4). The density $f_t = f_{t,N}$ satisfies the forward equation

$$\partial_t f_t = \mathcal{L} f_t, \tag{6.5}$$

where

$$\mathcal{L} = \mathcal{L}_N := \sum_{i=1}^N \frac{1}{2N}\partial_i^2 + \sum_{i=1}^N \Big(-\frac{\beta}{4}\lambda_i + \frac{\beta}{2N}\sum_{j\neq i}\frac{1}{\lambda_i - \lambda_j}\Big)\partial_i, \quad \partial_i = \frac{\partial}{\partial \lambda_i},$$

with $\beta = 1$ for the real symmetric case and $\beta = 2$ in the complex Hermitian case. The initial data f_0 is given by the original generalized Wigner matrix.

Now we define a useful technical tool that was first introduced in [29]. For any $\tau > 0$ denote by $W = W^\tau$ an auxiliary potential defined by

$$W^\tau(\boldsymbol{\lambda}) := \sum_{j=1}^N W_j^\tau(\lambda_j), \qquad W_j^\tau(\lambda) := \frac{1}{2\tau}(\lambda_j - \gamma_j)^2,$$

i.e. it is a quadratic confinement on scale $\sqrt{\tau}$ for each eigenvalue near its classical location, where the parameter $\tau > 0$ will be chosen later.

Definition 6.1. We define the probability measure $d\mu^\tau := Z_\tau^{-1} e^{-N\beta\mathcal{H}^\tau}$, where the total Hamiltonian is given by

$$\mathcal{H}^\tau := \mathcal{H} + W^\tau.$$

Here $\mathcal{H}$ is the Gaussian Hamiltonian given by (2.4) and $Z_\tau = Z_{\mu^\tau}$ is the partition function. The measure μ^τ will be referred to as the *relaxation measure* with *relaxation time* τ.

Denote

$$Q := \sup_{0\le t\le 1} \frac{1}{N}\int \sum_{j=1}^N (\lambda_j - \gamma_j)^2 f_t(\boldsymbol{\lambda})\,\mu(d\boldsymbol{\lambda}).$$

Since H_t is a generalized Wigner matrix for all t, (6.3) implies that

$$Q \le N^{-2+2\nu} \tag{6.6}$$

for any $\nu > 0$ if $N \ge N_0(\nu)$ is large enough.

Recall the definition of the *Dirichlet form* with respect to a probability measure ω:

$$D^\omega(\sqrt{g}) := \sum_{i=1}^N D_i^\omega(\sqrt{g}), \qquad D_i^\omega(\sqrt{g}) := \frac{1}{2N}\int |\partial_i \sqrt{g}|^2\, d\omega = \frac{1}{8N}\int |\partial_i \log g|^2 g\, d\omega,$$

and the definition of the *relative entropy* of two probability measures $g\omega$ and ω:

$$S(g\omega|\omega) := \int g \log g\, d\omega.$$

Now we recall Theorem 2.5 from [31]:

Theorem 6.2. *For any $\tau > 0$ and for the local relaxation measure μ^τ, set $\psi := d\mu^\tau/d\mu$ and let $g_t := f_t/\psi$. Suppose there is a constant m such that*

$$S(f_\tau\mu^\tau|\mu^\tau) \le CN^m.$$

Then for any $t \ge \tau N^{\varepsilon'}$ the entropy and the Dirichlet form satisfy the estimates

$$S(g_t\mu^\tau|\mu^\tau) \le CN^2 Q\tau^{-1}, \qquad D^{\mu^\tau}(\sqrt{g_t}) \le CN^2 Q\tau^{-2},$$

where the constants depend on ε' and m.

Corollary 6.3. *Fix $\mathfrak{a} > 0$ and let $\tau \ge N^{-\mathfrak{a}}$. Under the assumptions of Theorem 6.2, for any $t \ge \tau N^{\varepsilon'}$ the entropy and the Dirichlet form satisfy*

$$D^\mu(\sqrt{f_t}) \le CN^2 Q\tau^{-2}. \tag{6.7}$$

Furthermore, if the initial data of the DBM, f_0, is given by a generalized Wigner ensemble, then

$$D^\mu(\sqrt{f_t}) \le CN^{2\mathfrak{a}+2\nu} \tag{6.8}$$

for any $\nu > 0$.

Proof. Since $g_t = f_t/\psi$, we have

$$D^\mu(\sqrt{f_t}) = \sum_{i=1}^N \frac{1}{8N}\int |\partial_i \log g_t + \partial_i \log \psi|^2 f_t\, d\mu$$

$$\le \frac{1}{4N}\sum_{i=1}^N \int |\partial_i \log g_t|^2 f_t\, d\mu + \frac{1}{4N}\sum_{i=1}^N \int |\partial_i \log \psi|^2 f_t\, d\mu \le 2D^{\mu^\tau}(\sqrt{g_t}) + 2N^2 Q\tau^{-2}.$$

Thus (6.7) follows from Theorem 6.2. Finally, (6.8) follows from (6.7) and (6.6). □

Define $f_{\mathbf{y}}$ to be the conditional density of $f\mu$ given $\mathbf{y}$ with respect to $\mu_{\mathbf{y}}$, i.e. $f_{\mathbf{y}}\mu_{\mathbf{y}} = (f\mu)_{\mathbf{y}}$. For any $\mathbf{y} \in \mathcal{R}_{L,K}$ we have the convexity bound (5.20). Thus we have the logarithmic Sobolev inequality

$$S(f_{\mathbf{y}}\mu_{\mathbf{y}}|\mu_{\mathbf{y}}) \le C\frac{K}{N}\sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}}) \tag{6.9}$$

and the bound

$$\int d\mu_{\mathbf{y}}\,|f_{\mathbf{y}} - 1| \le \sqrt{S(f_{\mathbf{y}}\mu_{\mathbf{y}}|\mu_{\mathbf{y}})} \le C\sqrt{\frac{K}{N}\sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}})}. \tag{6.10}$$

To control the Dirichlet forms D_i for most external configurations $\mathbf{y}$, we need the following lemma.

Lemma 6.4. *Fix $\mathfrak{a}, \nu > 0$, and $\tau \ge N^{-\mathfrak{a}}$. Suppose the initial data f_0 of the DBM is given by a generalized Wigner ensemble. Then, with some small $\varepsilon' > 0$, for any $t \ge \tau N^{\varepsilon'}$ there exists a set $\mathcal{G}_{L,K} \subset \mathcal{R}_{L,K}$ of good boundary conditions $\mathbf{y}$ with*

$$\mathbb{P}^{f_t\mu}(\mathcal{G}_{L,K}) \ge 1 - CN^{-\varepsilon'} \tag{6.11}$$

such that for any $\mathbf{y} \in \mathcal{G}_{L,K}$ we have

$$\sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{t,\mathbf{y}}}) \le CN^{3\varepsilon'+2\mathfrak{a}+2\nu}, \qquad f_{t,\mathbf{y}} = (f_t)_{\mathbf{y}},\ I = I_{L,K}, \tag{6.12}$$

and for any bounded observable O,

$$|[\mathbb{E}^{f_{t,\mathbf{y}}\mu_{\mathbf{y}}} - \mathbb{E}^{\mu_{\mathbf{y}}}]O(\mathbf{x})| \le CK^{1/2}N^{2\varepsilon'+\mathfrak{a}+\nu-1/2}. \tag{6.13}$$

Furthermore, for any $k \in I$ we also have

$$|\mathbb{E}^{f_{t,\mathbf{y}}\mu_{\mathbf{y}}} x_k - \gamma_k| \le CN^{-1+\nu}. \tag{6.14}$$

Proof. In this proof, we omit the subscript t, i.e. we write $f = f_t$. By definition of the conditional measure and by (6.8), we have

$$\mathbb{E}^{f\mu}\sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}}) = \sum_{i\in I} D_i^{\mu}(\sqrt{f}) \le D^{\mu}(\sqrt{f}) \le CN^{2\mathfrak{a}+2\nu}.$$

By the Markov inequality, (6.12) holds for all $\mathbf{y}$ in a set $\mathcal{G}^1_{L,K}$ with $\mathbb{P}^{f\mu}(\mathcal{G}^1_{L,K}) \ge 1 - CN^{-3\varepsilon'}$. Without loss of generality, by (6.4) we can assume that $\mathcal{G}^1_{L,K} \subset \mathcal{R}_{L,K}$. The estimate (6.13) now follows from (6.12) and (6.10).

Similarly, the rigidity bound (6.2) with respect to $f\mu$ can be translated to the measure $f_{\mathbf{y}}\mu_{\mathbf{y}}$, i.e. there exists a set $\mathcal{G}^2_{L,K}$ with

$$\mathbb{P}^{f\mu}(\mathcal{G}^2_{L,K}) \ge 1 - C_0\exp\big(-\tfrac{1}{2}c_1N^{c_2}\big)$$

such that for any $\mathbf{y} \in \mathcal{G}^2_{L,K}$ and for any $k \in I$,

$$\mathbb{P}^{f_{\mathbf{y}}\mu_{\mathbf{y}}}(|x_k - \gamma_k| \ge N^{-1+\nu}) \le \exp\big(-\tfrac{1}{2}c_1N^{c_2}\big).$$

In particular, we deduce (6.14) for any $\mathbf{y} \in \mathcal{G}^2_{L,K}$. Setting $\mathcal{G}_{L,K} := \mathcal{G}^1_{L,K} \cap \mathcal{G}^2_{L,K}$ concludes the proof. ◻

Lemma 6.5. *Fix $\mathfrak{a}, \nu > 0$ and $\tau \geq N^{-\mathfrak{a}}$. Suppose the initial data f_0 of the DBM is given by a generalized Wigner ensemble. Then, with some small $\varepsilon' > 0$, for any $t \geq \tau N^{\varepsilon'}$, $k \in I$ and $\mathbf{y} \in \mathcal{G}_{L,K}$, we have*

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_k - \mathbb{E}^{f_{t,\mathbf{y}}\mu_{\mathbf{y}}} x_k| \leq K N^{-3/2+\nu+\mathfrak{a}+2\varepsilon'}. \tag{6.15}$$

In particular, if the parameters are such that

$$K N^{-3/2+\nu+\mathfrak{a}+2\varepsilon'} \leq N^{-1} K^{\xi} \quad \textit{and} \quad N^{-1+\nu} \leq N^{-1} K^{\xi}$$

with some small $\xi > 0$, then

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_k - \alpha_k| \leq C N^{-1} K^{\xi}, \qquad k \in I, \tag{6.16}$$

where α_k is defined in (4.4). *In other words, the analogue of* (4.10) *is satisfied.*

Notice that if we apply (6.13) with the special choice $O(\mathbf{x}) = x_k$ then the error estimate will be much worse than (6.15). We wish to emphasize that (6.16) is not an obvious fact although we know that it holds for $\mathbf{y}$ with high probability with respect to the equilibrium measure μ. The key point of (6.16) is that it holds for any $\mathbf{y} \in \mathcal{G}_{L,K}$ and thus with "high probability" with respect to $f_t\mu$.

Proof of Lemma 6.5. Once again, we omit the subscript t. The estimate (6.16) is a simple consequence of (6.15), (6.14) and (5.10). To prove (6.15), we run the reversible dynamics

$$\partial_s q_s = \mathcal{L}_{\mathbf{y}} q_s$$

starting from initial data $q_0 = f_{\mathbf{y}}$, where the generator $\mathcal{L}_{\mathbf{y}}$ is the unique reversible generator with the Dirichlet form $D^{\mu_{\mathbf{y}}}$, i.e.,

$$-\int f \mathcal{L}_{\mathbf{y}} g \, d\mu_{\mathbf{y}} = \sum_{i \in I} \frac{1}{2N} \int \nabla_i f \cdot \nabla_i g \, d\mu_{\mathbf{y}}.$$

Recall that from the convexity bound (5.20), $\tau_K = K/N$ is the time to equilibrium of this dynamics. After differentiation and integration we get

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_k - \mathbb{E}^{f_{\mathbf{y}}\mu_{\mathbf{y}}} x_k| = \left| \int_0^{K^{\varepsilon'}\tau_K} du \, \frac{1}{2N} \int (\partial_k q_u) \, d\mu_{\mathbf{y}} \right| + O(\exp(-cK^{\varepsilon'})).$$

From the Schwarz inequality with a free parameter R, we can bound the last line by

$$\frac{1}{N} \int_0^{K^{\varepsilon'}\tau_K} du \int (R(\partial_k \sqrt{q_u})^2 + R^{-1}) \, d\mu_{\mathbf{y}} + O(\exp(-cK^{\varepsilon'})).$$

Dropping the trivial subexponential error term and using the fact that the time integral of the Dirichlet form is bounded by the initial entropy, we can bound the last line by

$$R S(f_{\mathbf{y}}\mu_{\mathbf{y}} | \mu_{\mathbf{y}}) + \frac{K^{\varepsilon'}\tau_K}{NR}.$$

Using the logarithmic Sobolev inequality for $\mu_{\mathbf{y}}$ and optimizing the parameter R, we can bound

$$
\begin{aligned}
|\mathbb{E}^{\mu_{\mathbf{y}}} x_k - \mathbb{E}^{f_{\mathbf{y}}\mu_{\mathbf{y}}} x_k| &\le \tau_K R \sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}}) + \frac{K^{\varepsilon'}\tau_K}{NR} + O(\exp(-cK^{\varepsilon'})) \\
&\le \frac{K^{\varepsilon'}\tau_K}{N^{1/2}} \Big(\sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{f_{\mathbf{y}}})\Big)^{1/2} + O(\exp(-cK^{\varepsilon'})).
\end{aligned}
$$

Combining this with (6.12), we obtain (6.15). □

We now prove the following comparison for the local statistics of μ and $f_t\mu$, where μ is the Gaussian β-ensemble, (2.11), with quadratic V, and f_t is the solution of (6.5) with initial data f_0 given by the original generalized Wigner matrix.

Lemma 6.6. *Fix n, $\mathfrak{a} > 0$ and $\tau \ge N^{-\mathfrak{a}}$. Then for $\mathfrak{a}$ sufficiently small there exist $\varepsilon, \varepsilon' > 0$ such that for any $t \ge \tau N^{\varepsilon'}$, any n and any n-particle observable O we have*

$$
\begin{aligned}
\big|[\mathbb{E}^{f_t\mu} - \mathbb{E}^{\mu}] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \dots, N(x_j - x_{j+n})\big)\big| \\
\le CN^{-\varepsilon}\|O'\|_\infty \qquad (6.17)
\end{aligned}
$$

for any $j \in [\![\alpha N, (1-\alpha)N]\!]$ and any sufficiently large N.

Proof. We will apply Lemma 6.4, and we choose $L = j$. Since $K \le N^{1/4}$, the right hand side of (6.13) is smaller than $N^{-\varepsilon}$. Thus

$$
\big|[\mathbb{E}^{f_{t,\mathbf{y}}\mu_{\mathbf{y}}} - \mathbb{E}^{\mu_{\mathbf{y}}}] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \dots, N(x_j - x_{j+n})\big)\big| \le CN^{-\varepsilon} \qquad (6.18)
$$

for all $\mathbf{y} \in \mathcal{G}_{L,K}$ with the probability of $\mathcal{G}_{L,K}$ satisfying (6.11). Choose any $\tilde{\mathbf{y}} \in \mathcal{R}^*$ defined in Lemma 5.2. We now apply Theorem 4.1 with both $\mu_{\mathbf{y}}$ and $\mu_{\tilde{\mathbf{y}}}$ given by the local Gaussian β-ensemble. Thus (4.10) is guaranteed by (5.6) and (6.16). Since $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R} = \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha)$ and Lemma 4.5 guarantees (4.9), we can apply Theorem 4.1 so that

$$
\big|[\mathbb{E}^{\mu_{\mathbf{y}}} - \mathbb{E}^{\mu_{\tilde{\mathbf{y}}}}] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \dots, N(x_j - x_{j+n})\big)\big| \le CN^{-\varepsilon}\|O'\|_\infty
$$

for all $\mathbf{y} \in \mathcal{G}_{L,K}$ and $\tilde{\mathbf{y}} \in \mathcal{R}^*$. Since $\mathbb{P}^{\mu}(\mathcal{R}^*) \ge 1 - N^{-\varepsilon}$ (see (5.5)), we thus have

$$
\big|[\mathbb{E}^{\mu_{\mathbf{y}}} - \mathbb{E}^{\mu}] O\big(N(x_j - x_{j+1}), N(x_j - x_{j+2}), \dots, N(x_j - x_{j+n})\big)\big| \le CN^{-\varepsilon}\|O'\|_\infty, \qquad (6.19)
$$

for all $\mathbf{y} \in \mathcal{G}_{L,K}$. From (6.18), (6.19) and the probability estimate (6.11) for $\mathcal{G}_{L,K}$, possibly reducing ε so that $\varepsilon \le \varepsilon'$, we obtain (6.17). □

Recall that H_t is the generalized Wigner matrix whose matrix elements evolve by independent OU processes. Thus in Lemma 6.6 we have proved that the local statistics of H_t, for $t \geq N^{-2\mathfrak{a}+\varepsilon'}$, is the same as the corresponding Gaussian one for any initial generalized matrix H_0. Finally, we need to approximate the generalized Wigner ensembles by Gaussian divisible ones. The idea of approximation first appeared in [26] via a "reverse heat flow" argument and was also used in [54] via a four moment theorem. We will follow the Green function comparison theorem of [33, 34], and in particular the result in [44], since these results were formulated and proved for generalized Wigner matrices.

Theorem 1.10 from [44] implies that if the first four moments of two generalized Wigner ensembles $H^{\mathbf{v}}$ and $H^{\mathbf{w}}$ are the same then

$$\lim_{N\to\infty}[\mathbb{E}^{\mathbf{v}}-\mathbb{E}^{\mathbf{w}}]O\big(N(x_j-x_{j+1}),N(x_j-x_{j+2}),\dots,N(x_j-x_{j+n})\big)=0, \tag{6.20}$$

provided that one of the ensembles, say $H^{\mathbf{w}}$, satisfies the following *level repulsion estimate*: For any $\kappa>0$, there is an $\alpha_0>0$ such that for any α satisfying $0<\alpha\leq\alpha_0$ there exists a $\nu>0$ such that

$$\mathbb{P}^{\mathbf{w}}\big(\mathcal{N}(E-N^{-1-\alpha},E+N^{-1-\alpha})\geq 2\big)\leq N^{-\alpha-\nu} \tag{6.21}$$

for all $E\in[-2+\kappa,2-\kappa]$, where $\mathcal{N}(a,b)$ denotes the number of eigenvalues in the interval (a,b). Although this theorem was stated with the assumption that all four moments of the matrix elements of the two ensembles match exactly, it in fact only requires that the first three moments match exactly and the differences of the fourth moments are less than $N^{-c''}$ for some small $c''>0$. The relaxation of the fourth moment assumption was carried out in detail in [33, 25] and we will not repeat it here.

We now apply (6.20) with $H^{\mathbf{v}}$ being the generalized Wigner ensemble for which we wish to prove the universality and $H^{\mathbf{w}}=H_t$ with $t=N^{-c'}$ for some small c'. The necessary estimate (6.21) on the level repulsion follows from the gap universality and the rigidity estimate for H_t. More precisely, for any energy E in the bulk, choose k such that $|\gamma_k-E|\leq C/N$. Then from the rigidity estimate (6.1), for any $c>0$ we have

$$\begin{aligned}
&\mathbb{P}^{\mathbf{w}}\big(\mathcal{N}(E-N^{-1-\alpha},E+N^{-1-\alpha})\geq 2\big)\\
&\qquad\leq\sum_{j:\,|j-k|\leq N^{c_0}}\mathbb{P}^{\mathbf{w}}(\lambda_{j+1}-\lambda_j\leq N^{-1-\alpha})+e^{-N^c}\\
&\qquad\leq\sum_{j:\,|j-k|\leq N^{c}}[\mathbb{P}^{\mu}(\lambda_{j+1}-\lambda_j\leq N^{-1-\alpha})+CN^{-\varepsilon}]+e^{-N^c}\\
&\qquad\leq CN^{\xi}N^{-(\beta+1)\alpha+c}+CN^{\alpha-\varepsilon}.
\end{aligned}$$

Here in the first inequality we have used the rigidity (6.1) and in the second inequality we have used (6.17) with an observable O that is a smoothed version of the characteristic function on scale $N^{-\alpha}$, i.e. $\|O'\|_\infty\leq CN^{\alpha}$. In the last step we have used the level repulsion bound for GOE/GUE for $\beta=1$ or 2, respectively. The level repulsion bound for GOE/GUE is well known; it also follows from of Theorem 4.3(ii) and the fact that (4.18) holds for all $\mathbf{y}\in\mathcal{R}_{L,K}$, i.e. with a very high probability (see (5.4)). Finally, we choose $\alpha_0\leq\varepsilon/4$. Then for any $\alpha<\alpha_0$, there exist small exponents ν,c,ξ such that $\nu+c+\xi<\alpha$. This proves (6.21) for the ensemble H_t.

Following [34], we construct an auxiliary Wigner matrix H_0 such that the first three moments of H_t and of the *original* matrix $H^{\mathbf{v}}$ are identical while the differences of the fourth moments are less than $N^{-c''}$ for some small $c'' > 0$ depending on c' (see Lemma 3.4 of [34]). The gap statistics of $H^{\mathbf{v}}$ and $H^{\mathbf{w}} = H_t$ coincide by (6.20) and the gap statistics of H_t coincides with those of GUE/GOE by Lemma 6.6. This completes the proof of (2.8) showing that the local gap statistics with the same gap-label j is identical for the generalized Wigner matrix and the Gaussian case. Now (2.9) follows directly from Theorem 2.3, which, in particular, compares the local gap statistics for different gap labels (k and m) in the Gaussian case. This completes the proof of Theorem 2.2. □

7. Rigidity and level repulsion of local measures

7.1. Rigidity of $\mu_{\mathbf{y}}$: proof of Theorem 4.2

We will prove Theorem 4.2 using a method similar to the proof of [8, Theorem 3.1]. Theorem 3.1 of [8] was proved by a quite complicated argument involving induction on scales and the loop equation. The loop equation, however, requires analyticity of the potential and it cannot be applied to prove Theorem 4.2 for a local measure whose potential $V_{\mathbf{y}}$ is not analytic. We note, however, that in [8] the loop equation was used only to estimate the *expected locations* of the particles. Now this estimate is given as a condition by (4.12) and thus we can adapt the proof in [8] to the current setting. For later application, however, we will need a stronger form of the rigidity bound, namely we will establish that the tail of the gap distribution has a Gaussian decay. This stronger statement requires some modifications to the argument from [8] which therefore we partially repeat here. We now introduce the notation needed to prove Theorem 4.2.

Let θ be a continuously differentiable nonnegative function with $\theta = 0$ on $[-1, 1]$ and $\theta'' \geq 1$ for $|x| > 1$. We can take for example $\theta(x) = (x-1)^2\mathbf{1}(x > 1) + (x+1)^2\mathbf{1}(x < -1)$.

For any $m \in [\![\alpha N, (1-\alpha)N]\!]$ and any integer $1 \leq M \leq \alpha N$, we denote $I^{(m,M)} := [\![m-M, m+M]\!]$ and $\mathcal{M} := |I^{(m,M)}| = 2M+1$. Let $\eta := \xi/3$. For any k, M with $|k-L| \leq K-M$, define

$$\phi^{(k,M)}(\mathbf{x}) := \sum_{i<j,\, i,j \in I^{(k,M)}} \theta\left(\frac{N(x_i - x_j)}{\mathcal{M}K^{2\eta}}\right). \tag{7.1}$$

Let

$$\omega_{\mathbf{y}}^{(k,M)} := Z_{\mathbf{y},\phi}\mu_{\mathbf{y}} e^{-\phi^{(k,M)}},$$

where $Z_{\mathbf{y},\phi}$ is a normalization constant. Choose an increasing sequence of integers, $M_1 < \cdots < M_A$, such that $M_1 = K^{\xi}$, $M_A = CK^{1-2\eta}$ with a large constant C, and $M_\gamma/M_{\gamma-1} \sim K^\eta$ (meaning that $cK^\eta \leq M_\gamma/M_{\gamma-1} \leq CK^\eta$). We can choose the sequence so that $A \leq C\eta^{-1}$. We set $\omega_\gamma := \omega_{\mathbf{y}}^{(k,M_\gamma)}$ and we study the rigidity properties of the measures $\omega_A, \omega_{A-1}, \ldots, \omega_1$ in this order. Note that $\mu_{\mathbf{y}} = \omega_A$ since $\mathbf{y} \in \mathcal{R}_{L,K} = \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$ guarantees that $|x_i - x_j| \leq |J_{\mathbf{y}}| \leq CK/N$ (see (4.22)), thus for $M = M_A = CK^{1-2\eta}$

the argument of θ in (7.1) is smaller than 1, so $\phi \equiv 0$ in this case. We also introduce the notation

$$x_k^{[M]} := \frac{1}{2M+1} \sum_{j=k-M}^{k+M} x_j .$$

Definition 7.1. We say that $\mu_{\mathbf{y}}$ has *exponential rigidity on scale* ℓ if there are constants C, c such that for any $k \in I$,

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \geq \ell + uK^{\xi}N^{-1}) \leq Ce^{-cu^2}, \qquad u > 0.$$

First we prove that $\mu_{\mathbf{y}}$ has exponential rigidity on scale $M_A N^{-1}$. Starting from $\gamma = A$, by the Herbst bound and the logarithmic Sobolev inequality (6.9) for $\mu_{\mathbf{y}}$ with LSI constant of order K/N, for any $k \in [\![L - K + M_A, L + K - M_A]\!]$ we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k^{[M_A]} - \mathbb{E}^{\mu_{\mathbf{y}}} x_k^{[M_A]}| \geq b/\sqrt{M_A}) \leq e^{-c(N/K)Nb^2}, \qquad b \geq 0,$$

i.e.

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k^{[M_A]} - \mathbb{E}^{\mu_{\mathbf{y}}} x_k^{[M_A]}| \geq uK^{\eta}/N) \leq Ce^{-cu^2}.$$

Using the estimate (6.16) we see that

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_k^{[M_A]} - \alpha_k^{[M_A]}| \leq CN^{-1}K^{\xi}.$$

Thus we obtain

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k^{[M_A]} - \alpha_k^{[M_A]}| \geq CN^{-1}K^{\xi} + uK^{\eta}/N) \leq Ce^{-cu^2}. \tag{7.2}$$

Since $x_{k-M}^{[M]} \leq x_k \leq x_{k+M}^{[M]}$ and the α_k's are regular with spacing of order $1/N$, we get

$$x_k - \alpha_k \leq x_{k+M}^{[M]} - \alpha_{k-M}^{[M]} \leq x_{k+M}^{[M]} - \alpha_{k+M}^{[M]} + CMN^{-1},$$

and we also have a similar lower bound. Thus

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \geq CM_A N^{-1} + uK^{\eta}/N) \leq Ce^{-cu^2} \tag{7.3}$$

for any $k \in [\![L - K + 2M_A, L + K - 2M_A]\!]$, where we have used $M_A \geq K^{\xi}$. If $k \in [\![L - K, L - K + 2M_A]\!]$, then

$$x_k - \alpha_k \leq x_{L-K+2M_A} - \alpha_{L-K+2M_A} + CM_A N^{-1},$$
$$x_k - \alpha_k \geq y_{L-K-1} - \alpha_k \geq -CM_A N^{-1}.$$

Thus

$$|x_k - \alpha_k| \leq |x_{L-K+2M_A} - \alpha_{L-K+2M_A}| + CM_A N^{-1}.$$

Since (7.3) holds for the difference $x_{L-K+2M_A} - \alpha_{L-K+2M_A}$, it holds for $x_k - \alpha_k$ as well (with at most an adjustment of C) for any $k \in [\![L - K, L - K + 2M_A]\!]$. A similar argument holds for $k \in [\![L + K - 2M_A, L + K]\!]$. Thus we have proved (7.3) for all $k \in [\![L - K, L + K]\!]$, i.e. we have shown exponential rigidity on scale $M_A N^{-1}$.

Now we use an induction on scales to show that if

(i) for any $k \in [\![L - K + M_\gamma, L + K - M_\gamma]\!]$ we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k^{[M_\gamma]} - \alpha_k^{[M_\gamma]}| \geq uK^\xi N^{-1}) \leq Ce^{-cu^2}, \qquad u \geq 0; \tag{7.4}$$

(ii) exponential rigidity holds on some scale $M_\gamma N^{-1}$,

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \geq CM_\gamma N^{-1} + uK^\xi N^{-1}) \leq Ce^{-cu^2}, \qquad k \in I,\ u \geq 0; \tag{7.5}$$

(iii) we have the entropy bound

$$S(\mu_{\mathbf{y}}|\omega_\gamma) \leq Ce^{-cM_\gamma^2 K^{-5\eta}}, \tag{7.6}$$

then (i)–(iii) also hold with γ replaced by $\gamma - 1$ as long as $M_{\gamma-1} \geq K^\xi$. The iteration can be started from $\gamma = A$, since (7.4) and (7.5) were proven in (7.2) and in (7.3) (even with a better bound), and (7.6) is trivial for $\gamma = A$ since $\omega_A = \mu_{\mathbf{y}}$.

We first notice that on any scale M_γ, the bound (7.4) implies (7.5) by the same argument used to deduce (7.3) for any $k \in I$ from (7.2). So we can focus on proving (7.4) and (7.6) on scale $M_{\gamma-1}$.

To prove (7.6) on scale $M_{\gamma-1}$, notice that (7.5) with $u = M_\gamma K^{-\xi}$ implies

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \geq CM_\gamma N^{-1}) \leq Ce^{-cM_\gamma^2 K^{-2\xi}}, \qquad k \in I. \tag{7.7}$$

Since

$$\theta\left(\frac{N(x_i - x_j)}{\mathcal{M}_{\gamma-1}K^{2\eta}}\right) = 0$$

unless $|x_i - x_j| \geq CM_{\gamma-1}N^{-1}K^{2\eta} \geq CM_\gamma N^{-1}K^\eta$, the scale $CM_\gamma N^{-1}$ is by a factor of K^η smaller than the scale of $x_i - x_j$ built into the definition of $\phi^{(k,M_{\gamma-1})}$ (see (7.1)). But for $i, j \in I^{(k,M_{\gamma-1})}$ we have $|x_i - x_j| \leq |x_i - \alpha_i| + |x_j - \alpha_j| + CM_{\gamma-1}N^{-1}$. Thus $\phi^{(k,M_{\gamma-1})} = 0$ unless we are on the event described in (7.7) at least for one k. Moreover, $|\nabla\phi^{(k,M_{\gamma-1})}(\mathbf{x})| \leq N^C$ for any configuration $\mathbf{x}$ in J. Thus, following the argument in [8, Lemma 3.15], via the logarithmic Sobolev inequality for $\mu_{\mathbf{y}}$, we get

$$S(\mu_{\mathbf{y}}|\omega_{\gamma-1}) \leq CKN^{-1}\mathbb{E}^{\mu_{\mathbf{y}}}|\nabla\phi^{(k,M_{\gamma-1})}|^2 \leq CN^C e^{-cM_\gamma^2 K^{-2\xi}} \leq Ce^{-cM_{\gamma-1}^2 K^{-5\eta}}.$$

Here we have used the fact that the prefactor N^C can be absorbed in the exponent by using $M_\gamma^2 K^{-2\xi} - M_{\gamma-1}^2 K^{-5\eta} \geq K^{2\xi-5\eta} = K^\eta \geq N^{\eta\delta}$ with $\xi = 3\eta$, since $M_{\gamma-1} \geq K^\xi$. We will not need it here, but we note that the same bound on the opposite relative entropy,

$$S(\omega_{\gamma-1}|\mu_{\mathbf{y}}) \leq Ce^{-cM_{\gamma-1}^2 K^{-5\eta}},$$

is also correct. Thus (7.6) for $\gamma - 1$ is proved.

Now we focus on proving (7.4) on scale $M_{\gamma-1}$. Set $1 \leq M' \leq M \leq K$ and fix $k \in I$ such that $|k - L| \leq K - M$. We state the following slightly generalized version of [8, Lemma 3.14].

Lemma 7.2. *For any integers* $1 \le M' \le M \le K$, $k \in [\![L-K+M, L+K-M]\!]$ *and* $k' \in [\![k-M+M', k+M-M']\!]$, *we have*

$$\mathbb{P}^{\omega^{(k,M)}}\left(|\lambda_{k'}^{[M']} - \lambda_k^{[M]} - \mathbb{E}^{\omega^{(k,M)}}(\lambda_{k'}^{[M']} - \lambda_k^{[M]})| > \frac{uK^{2\eta}}{N}\sqrt{\frac{M}{M'}}\right) \le Ce^{-cu^2}.$$

Compared with [8, Lemma 3.14], we first note that N^{ε} in [8] is changed to $K^{2\eta}$ because $\phi^{(k,M)}(\mathbf{x})$ in (7.1) is defined with a $K^{2\eta}$ factor instead of N^{ε}. Furthermore, here we have allowed the center at scale M' to be different from k. The only condition is that $[\![k'-M', k'+M']\!] \subset [\![k-M, k+M]\!]$. The proof of this lemma is identical to that of [8, Lemma 3.14].

In particular, for any $\gamma = 2, 3, \dots, A$, and with $M' = M_{\gamma-1}$ and $M = M_\gamma \le K^\eta M_{\gamma-1}$ and with any choice of $k_\gamma \in [\![L-K+M_\gamma, L+K-M_\gamma]\!]$ and $k_{\gamma-1} \in [\![L-K+M_{\gamma-1}, L+K-M_{\gamma-1}]\!]$ such that $[\![k_{\gamma-1}-M_{\gamma-1}, k_{\gamma-1}+M_{\gamma-1}]\!] \subset [\![k_\gamma - M_\gamma, k_\gamma + M_\gamma]\!]$, we get

$$\mathbb{P}^{\omega_\gamma}\left(|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - x_{k_\gamma}^{[M_\gamma]} - \mathbb{E}^{\omega_\gamma}(x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - x_{k_\gamma}^{[M_\gamma]})| > uK^{5\eta/2}/N\right) \le Ce^{-cu^2}. \tag{7.8}$$

The entropy bound (7.6) and the boundedness of x_k imply that

$$|\mathbb{E}^{\omega_\gamma} x_k - \mathbb{E}^{\mu_\mathbf{y}} x_k| \le C\sqrt{S(\mu_\mathbf{y}|\omega_\gamma)} \le Ce^{-cM_\gamma^2 K^{-5\eta}};$$

where $M_\gamma^2 K^{-5\eta} \ge K^{2\xi-5\eta} \ge K^\eta$ ($\eta = \xi/3$). We can combine it with (4.12) to have

$$|\mathbb{E}^{\omega_\gamma} x_k - \alpha_k| \le CK^\xi/N.$$

The measure ω_γ in (7.8) can also be changed to $\mu_\mathbf{y}$ at the expense of an entropy term $S(\mu_\mathbf{y}|\omega_\gamma)$. Using (7.6), we thus have

$$\begin{aligned}\mathbb{P}^{\mu_\mathbf{y}}\left(|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - x_{k_\gamma}^{[M_\gamma]} - (\alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_\gamma}^{[M_\gamma]})| \ge CK^\xi N^{-1} + uK^{5\eta/2}/N\right)& \\ \le Ce^{-cu^2} + Ce^{-cM_\gamma^2 K^{-5\eta}}&.\end{aligned}$$

Combining it with (7.4) and recalling $\xi = 3\eta$, we get

$$\mathbb{P}^{\mu_\mathbf{y}}(|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]}| \ge CK^\xi N^{-1} + uK^\xi/N) \le Ce^{-cu^2} + Ce^{-cM_\gamma^2 K^{-5\eta}}.$$

This gives (7.4) on scale $M_{\gamma-1}$ if $u \le cM_\gamma K^{-5\eta/2}$ with a small constant c. Suppose now that $u \ge cM_\gamma K^{-5\eta/2}$, which in particular means that $u \ge cK^{-\eta/2}$. Then, by (7.5),

$$\begin{aligned}&\mathbb{P}^{\mu_\mathbf{y}}(|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]}| \ge CK^\xi N^{-1} + uK^\xi/N) \\ &\quad \le \mathbb{P}^{\mu_\mathbf{y}}\left(|x_{k_{\gamma-1}}^{[M_{\gamma-1}]} - \alpha_{k_{\gamma-1}}^{[M_{\gamma-1}]}| \ge CM_\gamma N^{-1} + (1 - CK^{-\eta/2})uK^\xi/N\right) \\ &\quad \le \sum_{k \in I} \mathbb{P}^{\mu_\mathbf{y}}\left(|x_k - \alpha_k| \ge CM_\gamma N^{-1} + (1 - CK^{-\eta/2})uK^\xi/N\right) \\ &\quad \le CKe^{-c(1-CK^{-\eta/2})^2 u^2} \le Ce^{-c'u^2}.\end{aligned}$$

This proves (7.4) for $\gamma - 1$. Note that the constants slightly deterioriate at each iteration step, but the number of iterations is finite (of order $1/\eta = 3/\xi$), so eventually the constants C, c in (4.13) may depend on ξ. In fact, since the deterioriation is minor, one can also prove (4.13) with ξ-independent constants, but for simplicity of presentation we did not follow the change of these constants at each step.

After completing the iteration, from (7.5) for $\gamma = 1$, $M_1 = K^\xi$ we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}(|x_k - \alpha_k| \geq CK^\xi N^{-1} + uK^\xi N^{-1}) \leq Ce^{-cu^2}, \qquad k \in I.$$

This yields (4.13) for $u \geq 1$. Finally, (4.13) is trivial for $u \leq 1$ if the constant C is sufficiently large. This completes the proof of Theorem 4.2. □

7.2. Level repulsion estimates of $\mu_{\mathbf{y}}$: proof of Theorem 4.3

We now prove the level repulsion estimate, Theorem 4.3, for the local log-gas $\mu_{\mathbf{y}}$ with good boundary conditions $\mathbf{y}$. There are two key ideas in the following argument. We first recall the weak level repulsion estimate [8, (4.11)], which in the current notation asserts

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \leq s/N) \leq CNs$$

for any $s > 0$, and similar estimates may be deduced for internal gaps. Compared with (4.14), this estimate does not contain any β exponent; moreover, in order to obtain (4.17), the N factor has to be reduced to K^ξ (neglecting the irrelevant $\log N$ factor). Our first idea is to run this proof for a local measure with only K^ξ particles to reduce the N factor to K^ξ. The second idea involves introducing some auxiliary measures to catch some of the β-related factors. We first introduce these two auxiliary measures which are slightly modified versions of the local equilibrium measures:

$$\begin{aligned}&\mu_0 := \mu_{\mathbf{y},0} = Z_0(x_{L-K} - y_{L-K-1})^{-\beta}\mu_{\mathbf{y}}, \qquad \mu_1 := \mu_{\mathbf{y},1} = Z_1 W^{-\beta}\mu_{\mathbf{y}},\\ &W := (x_{L-K} - y_{L-K-1})(x_{L-K+1} - y_{L-K-1}),\end{aligned}$$

where Z_0, Z_1 are chosen for normalization. In other words, we drop the term $(x_{L-K} - y_{L-K-1})^\beta$ from the measure $\mu_{\mathbf{y}}$ in μ_0 and we drop W^β in μ_1. To estimate the upper gap, $y_{L+K+1} - x_{L+K}$, similar results will be needed when we drop the term $(y_{L+K+1} - x_{L+K})^\beta$, and the analogous version of W, but we will not state them explicitly.

We first prove the following results which are weaker than Theorem 4.3.

Lemma 7.3. *Let L and K satisfy* (4.1) *and consider the local equilibrium measure $\mu_{\mathbf{y}}$ defined in* (4.5).

(i) *Let ξ, α be any fixed positive constants and let $\mathbf{y} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$. Then for any $s > 0$ we have*

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \leq s/N) \leq C(Ks\log N)^{\beta+1}, \tag{7.9}$$

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K+1} - y_{L-K-1} \leq s/N) \leq C(Ks\log N)^{2\beta+1}. \tag{7.10}$$

(ii) *Let* $\mathbf{y}$ *be arbitrary with the only condition that* $|y_i| \le C$ *for all* i. *Then for any* $s > 0$ *we have the weaker estimates*

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \le s/N) \le (CsK/|J_{\mathbf{y}}|)^{\beta+1}, \tag{7.11}$$

$$\mathbb{P}^{\mu_{\mathbf{y},j}}(x_{L-K+1} - y_{L-K-1} \le s/N) \le (CsK/|J_{\mathbf{y}}|)^{2\beta+1}, \quad j = 0, 1. \tag{7.12}$$

To prove Lemma 7.3, we first prove estimates even weaker than (7.9)–(7.12) for $\mu_{\mathbf{y}}$ and $\mu_{\mathbf{y},j}$.

Lemma 7.4. *Let* L *and* K *satisfy* (4.1).

(i) *Let* ξ, α *be any fixed positive constants and let* $\mathbf{y} \in \mathcal{R}_{L.K} = \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$. *Then for any* $s > 0$,

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \le s/N) \le CKs \log N, \tag{7.13}$$

$$\mathbb{P}^{\mu_{\mathbf{y},j}}(x_{L-K} - y_{L-K-1} \le s/N) \le CKs \log N, \quad j = 0, 1. \tag{7.14}$$

(ii) *Let* $\mathbf{y}$ *be arbitrary with the only condition that* $|y_i| \le C$ *for all* i. *Then for any* $s > 0$ *we have the weaker estimates*

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \le s/N) \le CsK/|J_{\mathbf{y}}|, \tag{7.15}$$

$$\mathbb{P}^{\mu_{\mathbf{y},j}}(x_{L-K} - y_{L-K-1} \le s/N) \le CsK/|J_{\mathbf{y}}|, \quad j = 0, 1. \tag{7.16}$$

Proof. We will prove (7.13); the same proof with only change of notation works for (7.14) as well. We will comment on this at the end of the proof.

For notational simplicity, we first shift the coordinates by S so that in the new coordinates $\bar{y} = 0$, i.e. $y_{L-K-1} = -y_{L+K+1}$ and J is symmetric about the origin. With the notation $a := -y_{L-K-1}$ and $I := [\![L-K, L+K]\!]$, we first estimate the following quantity, for any $0 \le \varphi \le c$ (with a small constant):

$$\begin{aligned} Z_\varphi &:= \int \dots \int_{-a+a\varphi}^{a-a\varphi} d\mathbf{x} \prod_{\substack{i,j \in I \\ i<j}} (x_i - x_j)^\beta e^{-N\frac{\beta}{2}\sum_j V_{\mathbf{y}}(S+x_j)} \\ &= (1-\varphi)^{K+\beta K(K-1)/2} \int \dots \int_{-a}^{a} d\mathbf{w} \prod_{i<j} (w_i - w_j)^\beta e^{-N\frac{\beta}{2}\sum_j V_{\mathbf{y}}(S+(1-\varphi)w_j)}, \end{aligned}$$

where we have set

$$w_j := (1-\varphi)^{-1} x_{L+j}, \quad d\mathbf{x} = \prod_{|j| \le K} dx_{L+j}, \quad d\mathbf{w} = \prod_{|j| \le K} dw_j. \tag{7.17}$$

By definition,

$$\begin{aligned} &e^{-N\frac{\beta}{2}V_{\mathbf{y}}(S+(1-\varphi)w_j)} \\ &= e^{-N\frac{\beta}{2}V(S+(1-\varphi)w_j)} \prod_{k \le L-K-1} ((1-\varphi)w_j - y_k)^\beta \prod_{k \ge L+K+1} (y_k - (1-\varphi)w_j)^\beta. \end{aligned} \tag{7.18}$$

For the smooth potential V, we have

$$|V(S+(1-\varphi)w_j)) - V(S+w_j)| \le C|\varphi w_j| \le CK\varphi/N \tag{7.19}$$

with a constant depending on V, where we have used $|w_j| \le a \le CK/N$, which follows from $|J_{\mathbf{y}}| \le CK/N$ since $\mathbf{y} \in \mathcal{R}_{L,K}$ (see (4.22)).

Using $(1-\varphi)w_j - y_k \ge (1-\varphi)(w_j - y_k)$ for $L-2K \le k \le L-K-1$ and the identity

$$(1-\varphi)w_j - y_k = (w_j - y_k)\left[1 - \frac{\varphi w_j}{w_j - y_k}\right]$$

for any k, we have

$$\prod_{k \le L-K-1} ((1-\varphi)w_j - y_k)^\beta \ge (1-\varphi)^{\beta K} \prod_{k \le L-K-1} (w_j - y_k)^\beta \prod_{n < L-2K} \left[1 - \frac{\varphi w_j}{w_j - y_n}\right]^\beta, \tag{7.20}$$

and a similar estimate holds for $k \ge L+K+1$. After multiplying these estimates for all $j = 1, \ldots, K$, we thus have the bound

$$\frac{Z_\varphi}{Z_0} \ge \left[e^{-C\beta K\varphi}(1-\varphi)^{\beta K} \min_{|w| \le a} \left(\prod_{k<L-2K} \left[1 - \frac{\varphi w}{w - y_k}\right]^\beta \prod_{k > L+2K} \left[1 - \frac{\varphi w}{y_k - w}\right]^\beta \right)\right]^K.$$

Recall that $\mathbf{y} \in \mathcal{R}_{L,K}$, i.e. we have the rigidity bound for $\mathbf{y}$ with accuracy $N^{-1}K^\xi \ll K/N \sim a$ (see (4.7)), i.e. y_k's are regularly spaced on scale a or larger. Combining this with $|w| \le a \le CK/N$, we have

$$\sum_{k \le L-2K} \frac{\varphi w}{w - y_k} \le C\varphi K \log N. \tag{7.21}$$

Hence

$$\prod_{k<L-2K} \left[1 - \frac{\varphi w}{w - y_k}\right]^\beta \ge 1 - C\varphi K \log N, \tag{7.22}$$

and similar bounds hold for the $k \ge L+2K$ factors. Thus for any $\varphi \le c$ we get

$$Z_\varphi/Z_0 \ge 1 - C(\beta K^2 + K^2 \log N)\varphi \ge 1 - CK^2\varphi(\log N).$$

Now we choose $\varphi := s/(aN)$ and recall $a \sim K/N$. Therefore the $\mu_{\mathbf{y}}$-probability of $x_{L+1} - y_L \ge a\varphi = s/N$ can be estimated by

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \ge s/N) \ge Z_\varphi/Z_0 \ge 1 - CKs(\log N).$$

for all $sK \log N$ sufficiently small. If $sK \log N$ is large, then (7.13) is automatically satisfied. This proves (7.13).

In order to prove (7.15), we now replace the assumption $\mathbf{y} \in \mathcal{R}_{L,K}$ with $|y_i| \le C$. Instead of (7.20), we now have

$$\prod_{k \le L-K-1} ((1-\varphi)w_j - y_k)^\beta \ge (1-\varphi)^{\beta N} \prod_{k \le L-K-1} (w_j - y_k)^\beta,$$

and a similar estimate holds for $k \geq L + K + 1$. We thus have the bound

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K} - y_{L-K-1} \geq s/N) \geq Z_\varphi/Z_0 \geq [e^{-C\beta K\varphi}(1-\varphi)^{\beta N}]^K \geq 1 - C\varphi NK.$$

With the choice $\varphi := s/(|J_{\mathbf{y}}|N)$ this proves (7.15).

The proof of (7.14) and (7.16) for $\mu_{\mathbf{y},0}$ is very similar, just the $k = L - K - 1$ factor is missing from (7.18) in the case of $j = -K$. For $\mu_{\mathbf{y},1}$, two factors are missing. These modifications do not alter the basic estimates. □

Proof of Lemma 7.3. Recalling the definition of μ_0 and setting $X := x_{L-K} - y_{L-K-1}$ for brevity, we have

$$\mathbb{P}^{\mu_{\mathbf{y}}}(X \leq s/N) = \frac{\mathbb{E}^{\mu_0}[\mathbf{1}(X \leq s/N)X^\beta]}{\mathbb{E}^{\mu_0}[X^\beta]}.$$

From (7.14) we have

$$\mathbb{E}^{\mu_0}[\mathbf{1}(X \leq s/N)X^\beta] \leq C(s/N)^\beta Ks \log N$$

and with the choice $s = cK^{-1}(\log N)^{-1}$ in (7.14) we also have

$$\mathbb{P}^{\mu_0}\left(X \geq \frac{c}{NK \log N}\right) \geq 1/2$$

with some positive constant c. This implies that

$$\mathbb{E}^{\mu_0}[X^\beta] \geq \frac{1}{2}\left(\frac{c}{NK \log N}\right)^\beta.$$

We have thus proved that

$$\mathbb{P}^{\mu_{\mathbf{y}}}(X \leq s/N) \leq C(s/N)^\beta Ks \log N(NK \log N)^\beta = C(Ks \log N)^{\beta+1},$$

i.e. we have obtained (7.9).

For the proof of (7.10), we similarly use

$$\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K+1} - y_{L-K-1} \leq s/N) = \frac{\mathbb{E}^{\mu_1}[\mathbf{1}(x_{L-K+1} - y_{L-K-1} \leq s/N)W^\beta]}{\mathbb{E}^{\mu_1}[W^\beta]}.$$

From (7.14) we have

$$\begin{aligned}\mathbb{E}^{\mu_1}[\mathbf{1}(x_{L-K+1} - y_{L-K-1} \leq s/N)W^\beta] &\leq (s/N)^{2\beta}\mathbb{P}^{\mu_1}(x_{L-K} - y_{L-K-1} \leq s/N)\\ &\leq C(s/N)^{2\beta} Ks \log N.\end{aligned}$$

By the same inequality and with the choice $s = cK^{-1}(\log N)^{-1}$, we have

$$\mathbb{P}^{\mu_1}\left(W \geq \frac{c}{(NK \log N)^2}\right) \geq 1/2$$

with some positive constant c. This implies that

$$\mathbb{E}^{\mu_1}[W^\beta] \geq \frac{1}{2}\left(\frac{c}{(NK\log N)^2}\right)^\beta.$$

We have thus proved that

$$\begin{aligned}\mathbb{P}^{\mu_{\mathbf{y}}}(x_{L-K+1} - y_{L-K-1} \leq s/N) &\leq C(s/N)^{2\beta} Ks\log N((NK\log N)^2)^\beta \\ &= C(Ks\log N)^{2\beta+1},\end{aligned}$$

which proves (7.10). Finally, (7.11) and (7.12) can be proved using (7.15) and (7.16). □

Proof of Theorem 4.3. For a given i, define

$$\widetilde{I} := [\![\max(i - K^\xi, L-K-1), \min(i + K^\xi, L+K+1)]\!]$$

to be the indices in a K^ξ neighborhood of i. We further condition $\mu_{\mathbf{y}}$ on the points

$$z_j := x_j, \qquad j \in \widetilde{I}^c := I_{L,K} \setminus \widetilde{I},$$

and we let $\mu_{\mathbf{y},\mathbf{z}}$ denote the conditional measure on the remaining x variables $\{x_j : j \in \widetilde{I}\}$. Setting $L' := i$, $K' := K^\xi$, from the rigidity estimate (4.16) we have $(\mathbf{y},\mathbf{z}) \in \mathcal{R} = \mathcal{R}_{L',K'}(\xi^2\delta/2, \alpha)$ with a very high probability with respect to $\mu_{\mathbf{y}}$.

We will now apply (7.9) to the measure $\mu_{\mathbf{y},\mathbf{z}}$ with a new $\delta' = \delta\xi$ and $K' = K^\xi$. This ensures that the condition $N^{\delta'} \leq K'$ is satisfied and by the remark after (4.1), the change of δ affects only the threshold N_0. We obtain

$$\mathbb{P}^{\mu_{\mathbf{y},\mathbf{z}}}(x_i - x_{i+1} \leq s/N) \leq C(K^\xi s\log N)^{\beta+1}$$

with high probability in $\mathbf{z}$ with respect to $\mu_{\mathbf{y}}$. The subexponential lower bound on s, assumed in Theorem 4.3(ii), allows us to include the probability of the complement of $\mathcal{R}$ in the estimate, proving (4.17). A similar argument with (7.9) replaced by (7.10) yields (4.18).

To prove the weaker bounds (4.14), (4.15) for any $s > 0$, we may assume that the cases $L-K \leq i \leq L$ and $i > L$ are treated similarly. Since $\mathbf{y} \in \mathcal{R}_{L,K}$, we have $|J_{\mathbf{y}}| \geq cK/N$. We consider two cases, either $x_i - y_{L-K-1} \leq c'K/N$ or $x_i - y_{L-K-1} \geq c'K/N$, with $c' < c/2$. In the first case, we condition on $x_{L-K}, \ldots, x_i$ and we apply (7.15) to the measure $\nu_1 = \mu_{\mathbf{y},x_{L-K},\ldots,x_i}$. The configuration interval of this measure has length at least $cK/(2N)$, so we have

$$\mathbb{P}^{\nu_1}(x_{i+1} - x_i \leq s/N) \leq \frac{CKs}{cK/(2N)} \leq CNs. \tag{7.23}$$

In the second case, $x_i - y_{L-K-1} \geq c'K/N$, we condition on $x_{i+1}, x_{i+2}, \ldots, x_{L+K}$. The corresponding measure, denoted by $\nu_2 = \mu_{\mathbf{y},x_{i+1},\ldots,x_{L+K}}$, has a configuration interval of length at least $c'K/N$. We can now have the estimate (7.23) for ν_2. Putting these two estimates together proves (4.14). Finally, (4.15) can be proved in a similar way. □

8. Proof of Theorem 4.1

8.1. Comparison of the local statistics of two local measures

In this section, we start to compare the gap distributions of two local log-gases on the same configuration interval but with different external potential and boundary conditions. We will express the differences of the gap distributions between two measures in terms of random walks in time dependent random environments. From now on, we use microscopic coordinates and we relabel the indices so that the coordinates of x_j are $j \in I = \{-K, \dots, 0, 1, \dots, K\}$, i.e. we set $L = \widetilde{L} = 0$ in the earlier notation. This will have the effect that the labeling of the external points $\mathbf{y}$ will not run from 1 to N, but from some $L_- < 0$ to $L_+ > 0$ with $L_+ - L_- = N$. The important input is that the index set I of the internal points is macroscopically separated away from the edges, i.e. $|L_\pm| \geq \alpha N$.

The local equilibrium measures and their Hamiltonians will be denoted by the same symbols, $\mu_\mathbf{y}$ and $\mathcal{H}_\mathbf{y}$, as before, but with a slight abuse of notation we redefine them now to the microscopic scaling. Hence we have two measures $\mu_\mathbf{y} = e^{-\beta H_\mathbf{y}}/Z_\mathbf{y}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}} = e^{-\beta \widetilde{H}_{\widetilde{\mathbf{y}}}}/Z_{\widetilde{\mathbf{y}}}$, defined on the same configuration interval $J = J_\mathbf{y} = J_{\widetilde{\mathbf{y}}}$ with center $\bar{y}$, which, for simplicity, we assume to be $\bar{y} = 0$. The local density at the center is $\varrho(0) > 0$. The Hamiltonian is given by

$$
\begin{aligned}
\mathcal{H}_\mathbf{y}(\mathbf{x}) &:= \sum_{i\in I} \frac{1}{2} V_\mathbf{y}(x_i) - \sum_{\substack{i,j\in I\\ i<j}} \log|x_j - x_i|, \\
V_\mathbf{y}(x) &:= NV(x/N) - 2\sum_{j\notin I} \log|x - y_j|,
\end{aligned}
\tag{8.1}
$$

and $\widetilde{H}_{\widetilde{\mathbf{y}}}$ is defined in a similar way with V in (8.1) replaced with another external potential $\widetilde{V}$. Recall also the assumption that V'', $\widetilde{V}'' \geq -C$ (see (2.10)). We will need the rescaled version of the bounds (4.22)–(4.24), i.e.

$$
|J_\mathbf{y}| = \frac{\mathcal{K}}{\varrho(0)} + O(K^\xi), \tag{8.2}
$$

$$
V'_\mathbf{y}(x) = \varrho(0) \log \frac{d_+(x)}{d_-(x)} + O\left(\frac{K^\xi}{d(x)}\right), \qquad x \in J, \tag{8.3}
$$

$$
V''_\mathbf{y}(x) \geq \frac{\inf V''}{N} + \frac{c}{d(x)}, \qquad x \in J, \tag{8.4}
$$

where

$$
d(x) := \min\{|x - y_{-K-1}|, |x - y_{K+1}|\} \tag{8.5}
$$

is the distance to the boundary and we redefine $d_\pm(x)$ as

$$
d_-(x) := d(x) + \varrho(0)K^\xi, \qquad d_+(x) := \max\{|x - y_{-K-1}|, |x - y_{K+1}|\} + \varrho(0)K^\xi.
$$

The rescaled version of Lemma 4.5 states that (8.2)–(8.4) hold for any $\mathbf{y}$ in $\mathcal{R}_{L,K}(\xi\delta/2, \alpha/2)$, where the set $\mathcal{R}_{L,K}$, originally defined in (4.7), is expressed in microscopic coordinates.

We also rewrite (4.10) in microscopic coordinates as

$$|\mathbb{E}^{\mu_{\mathbf{y}}} x_j - \alpha_j| + |\mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}} x_j - \alpha_j| \le C K^{\xi}, \tag{8.6}$$

where

$$\alpha_j := \frac{j}{\mathcal{K}+1}|J| \tag{8.7}$$

is the rescaled version of the definition given in (4.4), but we keep the same notation.

The Dirichlet form is also redefined; in microscopic coordinates it is now given by

$$D^{\mu_{\mathbf{y}}}(\sqrt{g}) = \sum_{i\in I} D_i^{\mu_{\mathbf{y}}}(\sqrt{g}) = \frac{1}{2}\sum_{i\in I}\int |\partial_i \sqrt{g}|^2 \, d\mu_{\mathbf{y}}. \tag{8.8}$$

Due to the rescaling, the LSI (6.9) now takes the form, for $\mathbf{y} \in \mathcal{R}_{L,K}$,

$$S(g\mu_{\mathbf{y}}|\mu_{\mathbf{y}}) \le C K D^{\mu_{\mathbf{y}}}(\sqrt{g}).$$

Define the interpolating measures

$$\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}} = Z_r e^{-\beta r(\widetilde{V}_{\widetilde{\mathbf{y}}}(\mathbf{x}) - V_{\mathbf{y}}(\mathbf{x}))}\mu_{\mathbf{y}}, \qquad r \in [0,1], \tag{8.9}$$

so that $\omega^1_{\mathbf{y},\widetilde{\mathbf{y}}} = \widetilde{\mu}_{\widetilde{\mathbf{y}}}$ and $\omega^0_{\mathbf{y},\widetilde{\mathbf{y}}} = \mu_{\mathbf{y}}$ (Z_r is a normalization constant). This is again a local log-gas with Hamiltonian

$$\mathcal{H}^r_{\mathbf{y},\widetilde{\mathbf{y}}}(\mathbf{x}) = \frac{1}{2}\sum_{i\in I} V^r_{\mathbf{y},\widetilde{\mathbf{y}}}(x_i) - \sum_{i<j}\log|x_i - x_j| \tag{8.10}$$

and external potential

$$\begin{aligned} V^r_{\mathbf{y},\widetilde{\mathbf{y}}}(x) &:= (1-r)V_{\mathbf{y}}(x) + r\widetilde{V}_{\widetilde{\mathbf{y}}}(x), \\ V_{\mathbf{y}}(x) &:= NV(x/N) - 2\sum_{j\notin I}\log(x - y_i), \\ \widetilde{V}_{\widetilde{\mathbf{y}}}(x) &:= N\widetilde{V}(x/N) - 2\sum_{j\notin I}\log(x - \widetilde{y}_i). \end{aligned}$$

The Dirichlet form D^{ω} with respect to the measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ is defined similarly to (8.8).

For any bounded smooth function $Q(\mathbf{x})$ with compact support we can express the difference of the expectations with respect to two different measures $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$ as

$$\mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}} Q(\mathbf{x}) - \mathbb{E}^{\mu_{\mathbf{y}}} Q(\mathbf{x}) = \int_0^1 \frac{d}{dr}\mathbb{E}^{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}} Q(\mathbf{x})\, dr = \int_0^1 \beta\langle h_0(\mathbf{x}); Q(\mathbf{x})\rangle_{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}}\, dr, \tag{8.11}$$

where

$$h_0 = h_0(\mathbf{x}) := \sum_{i\in I}(V_{\mathbf{y}}(x_i) - \widetilde{V}_{\widetilde{\mathbf{y}}}(x_i)) \tag{8.12}$$

and $\langle f; g\rangle_{\omega} := \mathbb{E}^{\omega} fg - (\mathbb{E}^{\omega} f)(\mathbb{E}^{\omega} g)$ denotes the correlation. From now on, we will fix r. Our main result is the following estimate on the gap correlation function.

Theorem 8.1. *Consider two smooth potentials $V, \widetilde{V}$ with $V'', \widetilde{V}'' \geq -C$ and two boundary conditions $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L=0,K}(\xi^2\delta/2, \alpha)$, with some sufficiently small ξ, such that $J = J_{\mathbf{y}} = J_{\widetilde{\mathbf{y}}}$. Assume that (8.6) holds for both boundary conditions $\mathbf{y}, \widetilde{\mathbf{y}}$. Then, in particular, the rescaled version of the rigidity bound (4.13) and the level repulsion bounds (4.17), (4.18) hold for both $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$ by Theorems 4.2 and 4.3.*

Fix $\xi^* > 0$. Then there exist $\varepsilon, C > 0$, depending on ξ^*, such that for any sufficiently small ξ, any $0 \leq r \leq 1$ and $|p| \leq K^{1-\xi^*}$ we have

$$|\langle h_0; O(x_p - x_{p+1}, \dots, x_p - x_{p+n})\rangle_{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}}| \leq K^{C\xi} K^{-\varepsilon} \|O'\|_\infty \tag{8.13}$$

for any n-particle observable O, provided that $K \geq K_0(\xi, \xi^*, n)$ is large enough.

Notice that this theorem is formulated in terms of K being the only large parameter; N disappeared. We also remark that the restriction $|p| \leq K^{1-\xi^*}$ can be easily relaxed to $|p| \leq K - K^{1-\xi^*}$ with an additional argument conditioning on the set $\{x_i : i \in I \setminus \widetilde{I}\}$ to ensure that p is near the middle of the new index set $\widetilde{I}$. We will not need this more general form in this paper.

First we complete the proof of Theorem 4.1 assuming Theorem 8.1.

Proof of Theorem 4.1. The family of measures $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$, $0 \leq r \leq 1$, interpolate between $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$. So we can express the right hand side of (4.11), in the rescaled coordinates and with $L = \widetilde{L} = 0$, as

$$\begin{aligned}|[\mathbb{E}^{\mu_{\mathbf{y}}} - \mathbb{E}^{\widetilde{\mu}_{\widetilde{\mathbf{y}}}}] O(x_p - x_{p+1}, \dots, x_p - x_{p+n})| \\ \leq \int_0^1 dr\, \frac{d}{dr} \mathbb{E}^{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}} O(x_p - x_{p+1}, \dots, x_p - x_{p+n}).\end{aligned}$$

Using (8.11) and (8.13) we find that this difference is bounded by $K^{C\xi}K^{-\varepsilon}$. Choosing ξ sufficiently small so that $K^{C\xi}K^{-\varepsilon} \leq K^{-\varepsilon/2}$, we obtain (4.11) (with $\varepsilon/2$ instead of ε). $\square$

In the rest of the paper we will prove Theorem 8.1. The main difficulty is due to the fact that the correlation function of the points, $\langle x_i; x_j\rangle_\omega$, decays only logarithmically. In fact, for the GUE, Gustavsson [40, Theorem 1.3] proved that

$$\langle x_i; x_j\rangle_{\mathrm{GUE}} \sim \log \frac{N}{|i-j|+1},$$

and a similar formula is expected for ω. Therefore, it is very difficult to prove Theorem 8.1 based on this slow logarithmic decay. We notice that, however, the correlation function of the type

$$\langle g_1(x_i); g_2(x_j - x_{j+1})\rangle_\omega$$

decays much faster in $|i - j|$ since the second factor $g_2(x_j - x_{j+1})$ depends only on the difference. Correlations of the form $\langle g_1(x_i - x_{i+1}); g_2(x_j - x_{j+1})\rangle_\omega$ decay even faster. The fact that observables of differences of particles behave much nicer was a basic observation in our previous approach [29, 33, 34] to universality.

The measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ is closely related to $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$. Our first task in Section 8.2 is to show that both the rigidity and level repulsion estimates hold with respect to ω. Then we will rewrite the correlation functions in terms of a random walk representation in Proposition 9.1. The decay of correlation functions will be translated into a regularity property of the corresponding parabolic equation, whose proof will be the main content of Section 10. Section 9 consists of various cutoff estimates to remove the singularity of the diffusion coefficients in the random walk representations. We emphasize that these cutoffs are critical at $\beta = 1$; we do not know if our argument can be extended to $\beta < 1$.

8.2. Rigidity and level repulsion of the interpolating measure $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$

In this section we establish rigidity and level repulsion results for the interpolating measure $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$, similar to the ones established for $\mu_{\mathbf{y}}$ in Section 7 and stated in Theorems 4.2 and 4.3.

Lemma 8.2. *Let L and K satisfy (4.1) and $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha)$. With the notation $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ there exist constants C, θ_3, C_2 and C_3 such that the following estimates hold:*

(i) [Rigidity bound]

$$\mathbb{P}^\omega(|x_i - \alpha_i| \geq CK^{C_2\xi^2}) \leq Ce^{-K^{\theta_3}}, \quad i \in I. \tag{8.14}$$

(ii) [Weak form of level repulsion] *For any $s > 0$ we have*

$$\mathbb{P}^\omega(x_{i+1} - x_i \leq s) \leq C(Ns)^{\beta+1}, \qquad i \in [\![L-K-1, L+K]\!],\ s > 0, \tag{8.15}$$

$$\mathbb{P}^\omega(x_{i+2} - x_i \leq s) \leq C(Ns)^{2\beta+1}, \qquad i \in [\![L-K-1, L+K-1]\!],\ s > 0, \tag{8.16}$$

(iii) [Strong form of level repulsion] *With some small $\theta > 0$, for any $s \geq \exp(-K^\theta)$ we have*

$$\mathbb{P}^\omega(x_{i+1} - x_i \leq s) \leq C(K^{C_3\xi}s)^{\beta+1}, \qquad i \in [\![L-K-1, L+K]\!], \tag{8.17}$$

$$\mathbb{P}^\omega(x_{i+2} - x_i \leq s) \leq C(K^{C_3\xi}s)^{2\beta+1}, \qquad i \in [\![L-K-1, L+K-1]\!]. \tag{8.18}$$

(iv) [Logarithmic Sobolev inequality]

$$S(g\omega|\omega) \leq CK D^\omega(\sqrt{g}). \tag{8.19}$$

Note that in (8.14) we state only the weaker form of the rigidity bound, similar to (4.16). It is possible to prove the strong form of rigidity with Gaussian tail (4.13) for ω, but we will not need it in this paper.

The level repulsion bounds will mostly be used in the following estimates which trivially follow from (8.15)–(8.18):

Corollary 8.3. *Under the assumptions of Lemma 8.2, for any $p < \beta + 1$ we have*

$$\mathbb{E}^{\omega} \frac{1}{|x_i - x_{i+1}|^p} \le C_p K^{C_3 \xi}, \quad i \in [\![L - K - 1, L + K]\!], \tag{8.20}$$

and for any $p < 2\beta + 1$,

$$\mathbb{E}^{\omega} \frac{1}{|x_i - x_{i+2}|^p} \le C_p K^{C_3 \xi}, \quad i \in [\![L - K - 1, L + K - 1]\!]. \qquad \square$$

The key to translate the rigidity estimate of the measures $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$ to the measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ is to show that the analogue of (8.6) holds for ω.

Lemma 8.4. *Let L and K satisfy (4.1) and $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}(\xi\delta/2, \alpha)$. Consider the local equilibrium measure $\mu_{\mathbf{y}}$ defined in (4.6) and assume that (4.10) is satisfied. Let $\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ be the measure defined in (8.9). Recall that α_k denote the equidistant points in J (see (8.7)). Then there exists a constant C, independent of ξ, such that*

$$\mathbb{E}^{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}} |x_j - \alpha_j| \le C K^{C\xi}. \tag{8.21}$$

Proof of Lemma 8.4. We first prove the following estimate on the entropy.

Lemma 8.5. *Suppose μ_1 is a probability measure and $\omega = Z^{-1} e^g d\mu_1$ for some function g and normalization Z. Then*

$$S := S(\omega|\mu_1) = \mathbb{E}^{\omega} g - \log \mathbb{E}^{\mu_1} e^g \le \mathbb{E}^{\omega} g - \mathbb{E}^{\mu_1} g. \tag{8.22}$$

Consider two probability measures $d\mu_i = Z_i^{-1} e^{-H_i} d\mathbf{x}$, $i = 1, 2$. Denote

$$g = r(H_1 - H_2), \quad 0 < r < 1,$$

and set $\omega = Z^{-1} e^g d\mu_1$ as above. Then

$$\min(S(\omega|\mu_1), S(\omega|\mu_2)) \le [\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1}](H_1 - H_2). \tag{8.23}$$

Proof. The first inequality is a trivial consequence of the Jensen inequality

$$S = \mathbb{E}^{\omega} g - \log \mathbb{E}^{\mu_1} e^g \le \mathbb{E}^{\omega} g - \mathbb{E}^{\mu_1} g.$$

The entropy inequality yields

$$\mathbb{E}^{\omega} g \le r \log \mathbb{E}^{\mu_1} e^{g/r} + rS. \tag{8.24}$$

By the definition of g, we have

$$\log \mathbb{E}^{\mu_1} e^{g/r} = -\log \int e^{-g/r} \, d\mu_2 \le \mathbb{E}^{\mu_2} g/r.$$

Using this inequality and (8.24) in (8.22), we obtain

$$S \le \frac{r}{1-r} [\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1}](H_1 - H_2).$$

We can assume that $r \leq 1/2 \leq 1-r$ since otherwise we can switch the roles of H_1 and H_2. Hence (8.23) holds, and this concludes the proof of Lemma 8.5. $\square$

We now apply this lemma with $\mu_2 = \widetilde{\mu}_{\widetilde{\mathbf{y}}}$ and $\mu_1 = \mu_{\mathbf{y}}$ to prove that

$$\min[S(\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}|\mu_{\mathbf{y}}), S(\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}|\widetilde{\mu}_{\widetilde{\mathbf{y}}})] \leq K^{C\xi}. \tag{8.25}$$

To see this, by definition of g and the rigidity estimate (4.13), we have

$$\begin{aligned}
\mathbb{E}^{\mu_2} g - \mathbb{E}^{\mu_1} g &= \frac{r}{2}[\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1}] \sum_{i\in I} [V_{\mathbf{y}}(x_i) - \widetilde{V}_{\widetilde{\mathbf{y}}}(x_i)] \\
&= \frac{r}{2}[\mathbb{E}^{\mu_2} - \mathbb{E}^{\mu_1}] \sum_{i\in I} \int_0^1 ds\, [V'_{\mathbf{y}}(s\alpha_i + (1-s)x_i) - \widetilde{V}'_{\widetilde{\mathbf{y}}}(s\alpha_i + (1-s)x_i)](x_i - \alpha_i) \\
&= [\mathbb{E}^{\mu_2} + \mathbb{E}^{\mu_1}] O\bigg(\sum_{i\in I} \sup_{s\in[0,1]} \frac{K^\xi}{d(s\alpha_i + (1-s)x_i)} |x_i - \alpha_i| \bigg) \leq K^{C\xi}.
\end{aligned}$$

In the first step we have used the fact that the leading term $V_{\mathbf{y}}(\alpha_i) - \widetilde{V}_{\widetilde{\mathbf{y}}}(\alpha_i)$ in the Taylor expansion is deterministic, so it vanishes after taking the difference of two expectations. In the last step we have used the fact that with a very high μ_1- or μ_2-probability $d(s\alpha_i + (1-s)x_i) \sim d(\alpha_i)$ are equidistant up to an additive error K^ξ if i is away from the boundary, i.e., $-K + K^{C\xi} \leq i \leq K - K^{C\xi}$ (see (4.13)). For indices near the boundary, say $-K \leq i \leq -K + K^{C\xi}$, we have used $d(s\alpha_i + (1-s)x_i) \geq c\min\{1, d(x_{-K})\}$. Noticing that $d(x_{-K}) = x_{-K} - y_{-K-1}$, the level repulsion bound (4.17) (complemented with the weaker bound (7.9) that is valid for all $s > 0$) guarantees that the short distance singularity $[d(x_{-K})]^{-1}$ has an $\mathbb{E}^{\mu_{1,2}}$ expectation bounded by $CK^{C\xi}$.

We now assume that (8.25) holds with the choice of $S(\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}|\mu_{\mathbf{y}})$ for simplicity of notation. By the entropy inequality, we have

$$\mathbb{E}^{\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}} |x_i - \alpha_i| \leq \log \mathbb{E}^{\mu_{\mathbf{y}}} e^{|x_i - \alpha_i|} + K^{C\xi}. \tag{8.26}$$

From the Gaussian tail of the rigidity estimate (4.13), we have

$$\log \mathbb{E}^{\mu_{\mathbf{y}}} e^{|x_i - \alpha_i|} \leq K^{C\xi}.$$

Using this bound in (8.26) we have proved (8.21) and this concludes the proof of Lemma 8.4. $\square$

Proof of Lemma 8.2. Given (8.21), the proof of (8.14) follows the argument in the proof of Theorem 4.2, applied to ξ^2 instead of ξ. Once the rigidity bound (8.14) is proved, we can follow the proof of Theorem 4.3 to obtain all four level repulsion estimates, (8.15)–(8.18), analogously to the proofs of (4.14), (4.15), (4.17) and (4.18), respectively. The $\log N$ factor can be incorporated into $K^{C_3\xi}$.

Finally, to prove (8.19), let $\mathcal{L}^\omega$ be the reversible generator given by the Dirichlet form

$$-\int f\mathcal{L}^\omega f\, d\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}} = \frac{1}{2} \sum_{|j|\leq K} \int (\partial_j f)^2\, d\omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}. \tag{8.27}$$

Thus for the Hamiltonian $\mathcal{H} = \mathcal{H}^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ of the measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ (see (8.10)), we have

$$\langle \mathbf{v}, \nabla^2 \mathcal{H}(\mathbf{x})\mathbf{v}\rangle = \frac{1}{2}\sum_i [(1-r)V''_{\mathbf{y}}(x_i) + r\widetilde{V}''_{\widetilde{\mathbf{y}}}(x_i)]v_i^2 + \sum_{i<j} \frac{(v_i - v_j)^2}{(x_i - x_j)^2} \ge \frac{c}{K}\sum_i v_i^2, \tag{8.28}$$

by using (8.4) and $d(x) \le CK$ for good boundary conditions. Thus LSI takes the form

$$S(g\omega|\omega) \le CKD^\omega(\sqrt{g}).$$

This completes the proof of Lemma 8.2. □

The dynamics given by the generator $\mathcal{L}^\omega$ with respect to the interpolating measure $\omega = \omega^r_{\mathbf{y},\widetilde{\mathbf{y}}}$ can also be characterized by the following SDE:

$$dx_i = dB_i + \beta\left[-\frac{1}{2}(V^r_{\mathbf{y},\widetilde{\mathbf{y}}})'(x_i) + \frac{1}{2}\sum_{j\ne i}\frac{1}{(x_i - x_j)}\right]dt, \tag{8.29}$$

where $(B_{-K}, B_{-K+1}, \dots, B_K)$ is a family of independent standard Brownian motions. With a slight abuse of notation, when we talk about the DBM process, we will use $\mathbb{P}^\omega$ and $\mathbb{E}^\omega$ to denote the probability and expectation with respect to this dynamics with initial data ω, i.e., in equilibrium. This dynamical point of view gives rise to a representation for the correlation (8.13) in terms of random walks in random environment.

Starting from Section 9 we will focus on proving Theorem 8.1. The proof is based on a dynamical idea and it will be completed in Section 9.7.

9. Local statistics of interpolating measures: Proof of Theorem 8.1

9.1. Outline of the proof of Theorem 8.1

Theorem 8.1 will be proved by the following main steps. We remind the readers that the boundary conditions $\mathbf{y}, \widetilde{\mathbf{y}}$ are in the good sets and we have chosen $L = 0$ for convenience. For simplicity, we assume that $n = 1$, i.e. we consider a single gap observable $O(x_p - x_{p+1})$.

Step 1. Random walk representation. The starting point is a representation formula for the correlation $\langle h_0, O(x_p - x_{p+1})\rangle_\omega$. For any smooth observables $F(\mathbf{x})$ and $Q(\mathbf{x})$ and any time $T > 0$ we have the following representation formula for the time dependent correlation function (see (9.14) for the precise statement):

$$\begin{aligned}\mathbb{E}^\omega Q(\mathbf{x})F(\mathbf{x}) &- \mathbb{E}^\omega Q(\mathbf{x}(0))F(\mathbf{x}(T)) \\ &= \frac{1}{2}\int_0^T dS\, \mathbb{E}^\omega \sum_{b\in I} \partial_b Q(\mathbf{x}(0))\langle \nabla F(\mathbf{x}(S)), \mathbf{v}^b(S, \mathbf{x}(\cdot))\rangle.\end{aligned}$$

Here the path $\mathbf{x}(\cdot)$ is the solution of the reversible stochastic dynamics (8.29) with equilibrium measure ω. We use the notation $\mathbb{E}^\omega$ also for the expectation with respect to the path measure starting from the initial distribution ω, and $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^{\mathcal{K}}$, recalling that $|I| = 2K + 1 = \mathcal{K}$. Furthermore, for any $b \in I$ and any fixed path $\mathbf{x}(\cdot)$, the vector $\mathbf{v}^b(t) = \mathbf{v}^b(t, \mathbf{x}(\cdot)) \in \mathbb{R}^{\mathcal{K}}$ is the solution to the equation

$$\partial_t \mathbf{v}^b(t) = -\mathcal{A}(t)\mathbf{v}^b(t), \qquad t \geq 0, \qquad v_j^b(0) = \delta_{bj}. \tag{9.1}$$

The matrix $\mathcal{A}(t)$ depends on time through the path $\mathbf{x}(t)$ and it is given by

$$\mathcal{A}(t) := \beta \nabla^2 \mathcal{H}^r_{\mathbf{y}, \widetilde{\mathbf{y}}}(\mathbf{x}(t)).$$

From (8.10), it is of the form $\mathcal{A}(t) = \widetilde{\mathcal{A}}(\mathbf{x}(t)) = \widetilde{\mathcal{B}}(\mathbf{x}(t)) + \widetilde{\mathcal{W}}(\mathbf{x}(t))$ with $\widetilde{\mathcal{W}}(\mathbf{x}(t)) \geq 0$. The matrix elements of $\widetilde{\mathcal{B}}$ are given by

$$[\widetilde{\mathcal{B}}(\mathbf{x})\mathbf{v}]_j = -\sum_{k \neq j} \widetilde{B}_{jk}(\mathbf{x})(v_k - v_j), \qquad \widetilde{B}_{jk}(\mathbf{x}) = \beta/(x_j - x_k)^2, \qquad j \neq k.$$

Furthermore, $\mathcal{A}(t) \geq CK^{-1}$ (see (8.28)), and the time to equilibrium for the $\mathbf{x}(t)$ process is of order K (Corollary 9.2). Applying this representation to $O(x_p - x_{p+1})$ and cutting off the time integration at $C_1 K \log K$ with some large constant C_1, we will have (see (9.25))

$$\begin{aligned}&\langle h_0; O(x_p - x_{p+1})\rangle_\omega \\ &= \frac{1}{2}\int_0^{C_1 K \log K} d\sigma \sum_b \mathbb{E}^\omega[\partial_b h_0(\mathbf{x}) O'(x_p - x_{p+1})(v_p^b(\sigma) - v_{p+1}^b(\sigma))] + O(\|O'\|_\infty K^{-2}),\end{aligned} \tag{9.2}$$

It is easy to check that $\partial_b h_0$ satisfies, with some small ξ', the estimate (see (9.39))

$$|\partial_b h_0(\mathbf{x})| \leq \frac{K^{\xi'}}{\min(|x_b - K|, |x_b + K|) + 1}. \tag{9.3}$$

Step 2. Cutoff of bad sets. Setting $\mathcal{T} := [0, C_1 K \log K]$, we define the "good set" of paths (see (9.26)) for which the rigidity estimate holds uniformly in time:

$$\mathcal{G} := \left\{ \sup_{s \in \mathcal{T}} \sup_{|j| \leq K} |x_j(s) - \alpha_j| \leq K^{\xi'} \right\},$$

where ξ' is s small parameter to be specified later and α_j is the classical location given by (8.7). For any $Z \in I$ and $\sigma \in \mathcal{T}$ we also define the following event that the gaps between particles near Z are not too small in an appropriate average sense:

$$\mathcal{Q}_{\sigma, Z} := \left\{ \sup_{s \in \mathcal{T}} \sup_{1 \leq M \leq K} \frac{1}{1 + |s - \sigma|} \left| \int_s^\sigma da \, \frac{1}{M} \sum_{i \in I : |i - Z| \leq M} \frac{1}{|x_i(a) - x_{i+1}(a)|^2} \right| \leq K^\rho \right\}, \tag{9.4}$$

where $\rho > 0$ is a small parameter to be specified later. By convention we set $x_i(a) = y_i$ whenever $|i| > K$. We will need that the gaps are not too small not only near Z but also near the boundary, so we define the new good set

$$\widehat{\mathcal{Q}}_{\sigma,Z} := \mathcal{Q}_{\sigma,Z} \cap \mathcal{Q}_{\sigma,-K} \cap \mathcal{Q}_{\sigma,K}.$$

Finally, we need to control the gaps not just around one time σ but around a sequence of times that dyadically accumulate at σ. The significance of this stronger condition will only be clear in the proof of our version of the De Giorgi–Nash–Moser bound in Section 10. We define

$$\widetilde{\mathcal{Q}}_{\sigma,Z} := \bigcap_{\tau \in \Xi} \widehat{\mathcal{Q}}_{\sigma+\tau,Z}, \tag{9.5}$$

where

$$\Xi := \{-K \cdot 2^{-m}(1 + 2^{-k}) : 0 \le k, m \le C \log K\}. \tag{9.6}$$

We will choose Z near the center of the interval I and show in (9.27) and (9.28) that the bad events are small in the sense that

$$\mathbb{P}^{\omega}(\mathcal{G}^c) \le Ce^{-K^{\theta}} \tag{9.7}$$

with some $\theta > 0$, and

$$\mathbb{P}^{\omega}(\widetilde{\mathcal{Q}}^c_{\sigma,Z}) \le CK^{C_4\xi-\rho} \tag{9.8}$$

for each fixed $Z \in I$ and fixed $\sigma \in \mathcal{T}$, where ξ is introduced in Theorem 8.1. Notice that while the rigidity bound (9.7) holds with a very high probability, the control on small gaps (9.8) is much weaker due to the power-law behavior of the level repulsion estimates.

Our goal is to insert the characteristic functions of the good sets into the expectation in (9.2). More precisely, we will prove in (9.41) that

$$\begin{aligned} &|\langle h_0; O(x_p - x_{p+1})\rangle_{\omega}| \\ &\quad \le \frac{1}{2}\|O'\|_{\infty} \int_0^{C_1 K \log K} \sum_{b \in I} \mathbb{E}^{\omega}[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}|\partial_b h_0(\mathbf{x})|\,|v_p^b(\sigma) - v_{p+1}^b(\sigma)|]\,d\sigma \\ &\qquad + O(\|O'\|_{\infty} K^{-\rho/6}). \end{aligned} \tag{9.9}$$

(With a slight abuse of notation we use $\mathcal{G}$ and $\widetilde{\mathcal{Q}}_{\sigma,Z}$ also to denote the characteristic function of these sets.) To prove this inequality, we note that the contribution of the bad set $\mathcal{G}^c$ can be estimated by (9.7). To bound the contribution of the bad set $\widetilde{\mathcal{Q}}^c_{\sigma,Z}$, the estimate (9.8) alone is not strong enough due to the time integration in (9.9). We will need a time-decay estimate for the solution $\mathbf{v}^b(\sigma)$. On the good set $\mathcal{G}$, the matrix element B_{jk} satisfies

$$B_{jk}(s) = \beta/(x_j(s) - x_k(s))^2 \ge b/(j-k)^2, \qquad 0 \le s \le \sigma,\ j \ne k,$$

with $b = \beta K^{-2\xi'}$. With this estimate, we will show in (9.36) that, for any $1 \le p \le q \le \infty$, the following decay estimate for the solution to (9.1) holds:

$$\|\mathbf{v}(s)\|_q \le (sb)^{-(1/p-1/q)}\|\mathbf{v}(0)\|_p, \qquad 0 < s \le \sigma. \tag{9.10}$$

This allows us to prove (9.9).

Step 3. Cutoff of the contribution from near the center. From (9.3), $\partial_b h_0(\mathbf{x})$ decays as a power law when x_b moves away from the boundary of J, i.e., when the index b moves away from $\pm K$. With the decay estimate (9.10), it is not difficult to show that the contribution of b in the interior, i.e., the terms with $|b| \le K^{1-c}$ for some $c > 0$ in the sum in (9.9), is negligible.

Step 4. Finite speed of propagation. We will prove that in the good set $\mathcal{G} \cap \widetilde{\mathcal{Q}}_{\sigma,Z}$ the dynamics (9.1) satisfies the finite speed of propagation estimate

$$|v_p^b(s)| \le \frac{CK^{c+1/2}\sqrt{s+1}}{|p-b|}$$

for some small constant c (see (9.47)). This estimate is not optimal, but it allows us to cut off the contribution in (9.9) for time $\sigma \le K^{1/4}$ for b away from the center, i.e., $K \ge |b| \ge K^{1-c}$. In this step we use $|p| \le K^{1-\xi^*}$ (ξ^* is some small constant) and the exponents are chosen such that $|p-b| \ge cK^{1-c}$.

Step 5. Parabolic regularity with singular coefficients. Finally, we have to estimate the r.h.s. of (9.9) in the regime $K^{1/4} \le \sigma \le C_1 K \log K$ and for $|p| \le K^{1-\xi^*}$ with the choice $Z = p$. This estimate will work uniformly in b. We will show that for all paths in $\mathcal{G} \cap \widetilde{\mathcal{Q}}_{\sigma,p}$, any solution to (9.1) satisfies the Hölder regularity estimate in the interior, i.e., for some constants $\alpha, \mathfrak{q} > 0$,

$$\sup_{|j-p|+|j'-p| \le \sigma^{1-\alpha}} |v_j(\sigma) - v_{j'}(\sigma)| \le CK^{\xi}\sigma^{-1-\frac{1}{2}\mathfrak{q}\alpha}. \tag{9.11}$$

Notice that the regularity depends on the time σ and that is why we need the short time cutoff in the previous step. The estimate (9.11) allows us to complete the proof that $\langle h_0; O(x_p - x_{p+1})\rangle_\omega \to 0$ as $K \to \infty$. The Hölder estimate will be stated as Theorem 9.8, and the entire Section 10 will be devoted to its proof.

9.2. *Random walk representation*

First we will recall a general formula for the correlation functions of the process (8.29) through a random walk representation (see (9.16) below). This equation in a lattice setting was given in [22, Proposition 2.2] (see also [39, Proposition 3.1]). The random walk representation already appeared in the earlier paper of Naddaf and Spencer [47], which was a probabilistic formulation of the idea of Helffer and Sjöstrand [42].

In this section we will work in a general setup. Let $J \subset \mathbb{R}$ be an interval and I an index set with cardinality $|I| = \mathcal{K}$. Consider a convex Hamilton function $\mathcal{H}(\mathbf{x})$ on $J^{\mathcal{K}}$ and let $\mathbf{x}(s)$ be the solution to

$$dx_i = dB_i + \beta \partial_i \mathcal{H}(\mathbf{x})dt, \qquad i \in I, \tag{9.12}$$

with initial condition $\mathbf{x}(0) = \mathbf{x} \in J_{\mathbf{y}}^I$, where $\{B_i : i \in I\}$ is a family of independent standard Brownian motions. The parameter $\beta > 0$ is introduced only for consistency with our applications. Let $\mathbb{E}_{\mathbf{x}}$ denote the expectation with respect to this path measure. With a slight abuse of notation, we will use $\mathbb{P}^\omega$ and $\mathbb{E}^\omega$ to denote the probability and expectation with respect to the path measure of the solution to (9.12) with initial condition $\mathbf{x}$ distributed according to ω. We assume that $\mathbb{P}^\omega(\mathbf{x}(t) \in J^{\mathcal{K}}) = 1$, i.e. the Hamiltonian confines the process to remain in the interval J. The corresponding invariant measure is $d\omega = Z_\omega^{-1} e^{-\beta\mathcal{H}(\mathbf{x})} d\mathbf{x}$ with generator $\mathcal{L}^\omega = -\frac{1}{2}\Delta + \frac{\beta}{2}\nabla\mathcal{H}\cdot\nabla$ and Dirichlet form

$$D^\omega(f) := \frac{1}{2}\int |\nabla f|^2 \, d\omega = -\int f\mathcal{L}^\omega f \, d\omega.$$

For any fixed path $\mathbf{x}(\cdot) := \{\mathbf{x}(s) : s \ge 0\}$ we define the operator ($\mathcal{K}\times\mathcal{K}$ matrix)

$$\mathcal{A}(s) := \widetilde{\mathcal{A}}(\mathbf{x}(s)),$$

where $\widetilde{\mathcal{A}} := \beta\mathcal{H}''$ and we assume that the Hessian matrix is positive definite, $\mathcal{H}''(\mathbf{x}) \ge c > 0$.

Proposition 9.1. *Assume that the Hessian matrix is positive definite,*

$$\inf_{\mathbf{x}} \mathcal{H}''(\mathbf{x}) \ge \tau^{-1} \tag{9.13}$$

with some constant $\tau > 0$. Then for any functions $F, G \in C^1(J^{\mathcal{K}}) \cap L^2(d\omega)$ and any time $T > 0$ we have

$$\begin{aligned}\mathbb{E}^\omega&[F(\mathbf{x})\,G(\mathbf{x})] - \mathbb{E}^\omega[F(\mathbf{x}(0))G(\mathbf{x}(T))]\\ &= \frac{1}{2}\int_0^T dS \int \omega(d\mathbf{x}) \sum_{a,b=1}^{\mathcal{K}} \partial_b F(\mathbf{x})\mathbb{E}_{\mathbf{x}}[\partial_a G(\mathbf{x}(S)) v_a^b(S, \mathbf{x}(\cdot))].\end{aligned} \tag{9.14}$$

Here for any $S > 0$ and for any path $\{\mathbf{x}(s) \in J^{\mathcal{K}} : s \in [0, S]\}$, we define $\mathbf{v}^b(t) = \mathbf{v}^b(t, \mathbf{x}(\cdot))$ as the solution to the equation

$$\partial_t \mathbf{v}^b(t) = -\mathcal{A}(t)\mathbf{v}^b(t), \quad t \in [0, S], \quad v_a^b(0) = \delta_{ba}. \tag{9.15}$$

The dependence of $\mathbf{v}^b$ on the path $\mathbf{x}(\cdot)$ is via the dependence $\mathcal{A}(t) = \widetilde{\mathcal{A}}(\mathbf{x}(t))$. In other words, $v_a^b(t)$ is the fundamental solution of the heat semigroup $\partial_s + \mathcal{A}(s)$.

Furthermore, for the correlation function we have

$$\langle F; G\rangle_\omega = \frac{1}{2}\int_0^\infty dS \int \omega(d\mathbf{x}) \sum_{a,b=1}^{\mathcal{K}} \partial_b F(\mathbf{x})\mathbb{E}_{\mathbf{x}}[\partial_a G(\mathbf{x}(S)) v_a^b(S, \mathbf{x}(\cdot))] \tag{9.16}$$

$$= \frac{1}{2}\int_0^{A\tau\log\mathcal{K}} dS \int \omega(d\mathbf{x}) \sum_{a,b=1}^{\mathcal{K}} \partial_b F(\mathbf{x})\mathbb{E}_{\mathbf{x}}[\partial_a G(\mathbf{x}(S)) v_a^b(S, \mathbf{x}(\cdot))] + O(\mathcal{K}^{-cA}) \tag{9.17}$$

for any constant $A > 0$.

Proof. This proposition in the lattice setting was already proved in [22, 39, 47]; we give here a proof in the continuous setting. Let $G(t, \mathbf{x})$ be the solution to the equation $\partial_t G = \mathcal{L}^\omega G$ with initial condition $G(0, \mathbf{x}) := G(\mathbf{x})$. By integrating the time derivative, we have

$$\begin{aligned} \mathbb{E}^\omega[F(\mathbf{x})\, G(\mathbf{x})] - \mathbb{E}^\omega[F(\mathbf{x}(0))G(\mathbf{x}(T))] &= -\int_0^T dS\, \frac{d}{dS}\mathbb{E}^\omega[F e^{S\mathcal{L}^\omega} G] \\ = -\int_0^T dS\, \mathbb{E}^\omega[F\mathcal{L}^\omega e^{S\mathcal{L}^\omega} G] &= \frac{1}{2}\int_0^T dS\, \mathbb{E}^\omega\langle \nabla F(\mathbf{x}), \nabla G(S, \mathbf{x})\rangle, \end{aligned} \tag{9.18}$$

where $\langle\, ,\rangle$ denotes the scalar product in $\mathbb{R}^{\mathcal{K}}$.

Taking the gradient of the equation $\partial_t G = \mathcal{L}^\omega G$ and computing the commutator $[\nabla, \mathcal{L}^\omega]$ yields the equation

$$\partial_t \nabla G(t, \mathbf{x}) = \mathcal{L}^\omega[\nabla G(t, \mathbf{x})] - \widetilde{\mathcal{A}}(\mathbf{x})[\nabla G(t, \mathbf{x})]$$

for the $\mathbf{x}$-gradient of G. Setting $\mathbf{u}(t, \mathbf{x}) := \nabla G(t, \mathbf{x})$ for brevity, we have the equation

$$\partial_t \mathbf{u}(t, \mathbf{x}) = \mathcal{L}^\omega \mathbf{u}(t, \mathbf{x}) - \widetilde{\mathcal{A}}(\mathbf{x})\mathbf{u}(t, \mathbf{x}) \tag{9.19}$$

with initial condition $\mathbf{u}(0, \mathbf{x}) = \mathbf{u}_0(\mathbf{x}) := \nabla G(\mathbf{x})$.

Notice that $\widetilde{\mathcal{A}}$ is a matrix and $\mathcal{L}^\omega$ acts on the vector $\mathbf{u}$ as a diagonal operator in the index space, i.e., $[\mathcal{L}^\omega \mathbf{u}(t, \mathbf{x})]_i = \mathcal{L}^\omega[\mathbf{u}(t, \mathbf{x})_i]$. The equation (9.19) can be solved by solving an equation (9.22) over the indices with coefficients that depend on the path generated by the operator $\mathcal{L}^\omega$ and then by taking expectation over the paths starting at $\mathbf{x}$. To obtain such a representation, we start with the time-dependent Feynman–Kac formula:

$$\mathbf{u}(\sigma, \mathbf{x}) = \mathbb{E}^{\mathbf{x}}\left[\widetilde{\mathrm{Exp}}\left(-\int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\, ds\right)\mathbf{u}_0(\mathbf{x}(\sigma))\right], \qquad \sigma > 0, \tag{9.20}$$

where

$$\begin{aligned} &\widetilde{\mathrm{Exp}}\left(-\int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\, ds\right) \\ &\quad := 1 - \int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s_1))\, ds_1 + \int_{0\le s_1 < s_2 \le \sigma} \widetilde{\mathcal{A}}(\mathbf{x}(s_1))\widetilde{\mathcal{A}}(\mathbf{x}(s_2))\, ds_1\, ds_2 + \cdots \end{aligned} \tag{9.21}$$

is the time-ordered exponential. To prove that (9.20) indeed satisfies (9.19), we notice from the definition (9.21) that

$$\begin{aligned} \mathbf{u}(\sigma, \mathbf{x}) &= \mathbb{E}_{\mathbf{x}}\widetilde{\mathrm{Exp}}\left(-\int_0^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\, ds\right)\mathbf{u}_0(\mathbf{x}(\sigma)) \\ &= \mathbb{E}_{\mathbf{x}}\mathbf{u}_0(\mathbf{x}(\sigma)) - \int_0^\sigma \mathbb{E}_{\mathbf{x}}\widetilde{\mathcal{A}}(\mathbf{x}(s_1))\mathbb{E}_{\mathbf{x}(s_1)}\widetilde{\mathrm{Exp}}\left(\int_{s_1}^\sigma \widetilde{\mathcal{A}}(\mathbf{x}(s))\, ds\right)\mathbf{u}_0(\mathbf{x}(\sigma))\, ds_1. \end{aligned}$$

Since the process is stationary in time, we have

$$\mathbf{u}(\sigma, \mathbf{x}) = \mathbb{E}_{\mathbf{x}}\mathbf{u}_0(\mathbf{x}(\sigma)) - \int_0^\sigma \mathbb{E}_{\mathbf{x}}\widetilde{\mathcal{A}}(\mathbf{x}(s_1))\mathbf{u}(\sigma - s_1, \mathbf{x}(s_1))\, ds_1$$

$$= \mathbb{E}_{\mathbf{x}} \mathbf{u}_0(\mathbf{x}(\sigma)) - \int_0^\sigma \mathbb{E}_{\mathbf{x}} \widetilde{\mathcal{A}}(\mathbf{x}(\sigma - s_1)) \mathbf{u}(s_1, \mathbf{x}(\sigma - s_1))\, ds_1$$
$$= e^{\sigma \mathcal{L}} \mathbf{u}_0(\mathbf{x}) - \int_0^\sigma [e^{(\sigma - s_1)\mathcal{L}} \widetilde{\mathcal{A}}(\cdot) \mathbf{u}(s_1, \cdot)](\mathbf{x})\, ds_1.$$

Differentiating this equation in σ we find that $\mathbf{u}$ defined in (9.20) indeed satisfies (9.19).

For any fixed path $\{\mathbf{x}(s) : s > 0\}$, the time-ordered exponential in (9.20),

$$\mathcal{U}(t) = \mathcal{U}(t; \mathbf{x}(\cdot)) := \widetilde{\mathrm{Exp}}\Big(- \int_0^t \widetilde{\mathcal{A}}(\mathbf{x}(s))\, ds \Big),$$

satisfies the matrix evolution equation

$$\partial_t \mathcal{U}(t) = -\mathcal{U}(t) \mathcal{A}(t), \qquad \mathcal{U}(0) = I,$$

which can be seen directly from (9.21). Let $\mathbf{v}^b(t)$ be the transpose of the b-th row of the matrix $\mathcal{U}(t)$. Then the equation for the column vector $\mathbf{v}^b(t)$ reads

$$\partial_t \mathbf{v}^b(t) = -\mathcal{A}(t) \mathbf{v}^b(t), \qquad \mathbf{v}^b_a(0) = \delta_{ab}. \tag{9.22}$$

Thus taking the b-th component of (9.20) we have

$$u_b(\sigma, \mathbf{x}) = \partial_b G(\sigma, \mathbf{x}) = \mathbb{E}^{\mathbf{x}} [\mathcal{U}(\sigma) \nabla G(\mathbf{x}(\sigma))]_b = \sum_a \mathbb{E}_{\mathbf{x}} [\partial_a G(\mathbf{x}(\sigma)) v^b_a(\sigma)],$$

and plugging this into (9.18), we obtain (9.14) by using $\mathbb{E}^\omega[\cdot] = \int \mathbb{E}_{\mathbf{x}}[\cdot]\, \omega(d\mathbf{x})$.

Formula (9.17) follows directly from (9.14) and from the fact that $\mathcal{H}'' \geq \tau^{-1}$ implies a spectral gap of order τ, in particular,

$$|\mathbb{E}^\omega [F(\mathbf{x}(0)) G(\mathbf{x}(T))] - \mathbb{E}^\omega[F] \mathbb{E}^\omega[G]| \leq e^{-cT/\tau} \|F\|_{L^2(\omega)} \|G\|_{L^2(\omega)}.$$

Finally, (9.16) directly follows from this, by taking the $T \to \infty$ limit. □

Now we apply our general formula to the gap correlation function on the left hand side of (8.13). To shorten formulas, we consider only the single gap case, $n = 1$; the general case is a straightforward extension. The gap index $p \in I$, $p \neq K$, is fixed; later we will impose further conditions on p to separate it from the boundary. The index set is $I = [\![-K, K]\!]$, the Hamiltonian in (9.12) is given by $\mathcal{H}^r_{\mathbf{y}, \widetilde{\mathbf{y}}}$, and (9.12) takes the form of (8.29). It is well known [2] that due to the logarithmic interaction in the Hamiltonian, $\beta \geq 1$ implies that the process $\mathbf{x}(t) = (x_{-K}(t), \ldots, x_K(t))$ preserves the initial ordering, i.e., $x_{-K}(t) \leq \cdots \leq x_K(t)$ and $x_i(t) \in J$ for every $i \in I$. The matrix $\widetilde{\mathcal{A}}$ is given by $\widetilde{\mathcal{A}} := \widetilde{\mathcal{B}} + \widetilde{\mathcal{W}}$ where $\widetilde{\mathcal{B}}$ and $\widetilde{\mathcal{W}}$ are the following $\mathbf{x}$-dependent matrices acting on vectors $\mathbf{v} \in \mathbb{R}^{\mathcal{K}}$:

$$[\widetilde{\mathcal{B}}(\mathbf{x}) \mathbf{v}]_j = -\sum_k \widetilde{B}_{jk}(\mathbf{x})(v_k - v_j), \qquad \widetilde{B}_{jk}(\mathbf{x}) = \frac{\beta}{(x_j - x_k)^2} \geq 0$$
$$[\widetilde{\mathcal{W}}(\mathbf{x}) \mathbf{v}]_j := \widetilde{W}_j(\mathbf{x}) v_j$$

with

$$\widetilde{W}_j(\mathbf{x}) := \frac{\beta}{2}\Bigg\{\sum_{|k|\geq K+1}\left[\frac{1-r}{(x_j-y_k)^2}+\frac{r}{(x_j-\widetilde{y}_k)^2}\right]+\frac{1-r}{N}V''\Big(\frac{x_j}{N}\Big)+\frac{r}{N}\widetilde{V}''\Big(\frac{x_j}{N}\Big)\Bigg\}.$$

Here $r\in[0,1]$ is a fixed parameter which we will omit from the notation of $\widetilde{\mathcal{W}}$.

For any fixed path $\mathbf{x}(\cdot)$, define the following time-dependent operators (matrices) on $\mathbb{R}^{\mathcal{K}}$:

$$\mathcal{A}(s) := \widetilde{\mathcal{A}}(\mathbf{x}(s)), \qquad \mathcal{B}(s) := \widetilde{\mathcal{B}}(\mathbf{x}(s)), \qquad \mathcal{W}(s) := \widetilde{\mathcal{W}}(\mathbf{x}(s)), \tag{9.23}$$

where $\mathcal{W}$ is a multiplication operator with the j-th diagonal $W_j(s)=\widetilde{W}_j(x_j(s))$ depending only the j-th component of the process $\mathbf{x}(s)$. Clearly $\mathcal{A}(s)=\mathcal{B}(s)+\mathcal{W}(s)$. We also define the associated (time dependent) quadratic forms which we denote by the corresponding lower case letters, in particular

$$\begin{aligned}
\mathfrak{b}(s)[\mathbf{u},\mathbf{v}] &:= \sum_{i\in I}u_i[\mathcal{B}(s)\mathbf{v}]_i=\frac{1}{2}\sum_{k,j\in I}B_{jk}(s)(u_k-u_j)(v_k-v_j),\\
\mathfrak{w}(s)[\mathbf{u},\mathbf{v}] &:= \sum_{i\in I}u_i[\mathcal{W}(s)\mathbf{v}]_i=\sum_i u_iW_i(s)v_i,\\
\mathfrak{a}(s)[\mathbf{u},\mathbf{v}] &:= \mathfrak{b}(s)[\mathbf{u},\mathbf{v}]+\mathfrak{w}(s)[\mathbf{u},\mathbf{v}].
\end{aligned}\tag{9.24}$$

With this notation we can apply Proposition 9.1 to our case and get

Corollary 9.2. *Let h_0 be given by (8.12), let $O=O_N:\mathbb{R}\to\mathbb{R}$ be an observable for $n=1$ (see (2.7)), and assume that $\mathbf{y},\widetilde{\mathbf{y}}\in\mathcal{R}_{L=0,K}(\xi^2\delta/2,\alpha)$, in particular $\mathcal{A}(s)$ given in (9.23) satisfies $\mathcal{A}(s)\geq\tau^{-1}$ with $\tau=CK$ by (8.28). Then with a large constant C_1 and for any $p\in I$, $-K\leq p\leq K-1$, we have*

$$\begin{aligned}
&\langle h_0;\,O(x_p-x_{p+1})\rangle_\omega\\
&\quad=\frac{1}{2}\int_0^{C_1K\log K}d\sigma\int\sum_{b\in I}\partial_bh_0(\mathbf{x})\mathbb{E}_{\mathbf{x}}[O'(x_p-x_{p+1})(v_p^b(\sigma)-v_{p+1}^b(\sigma))]\,\omega(d\mathbf{x})\\
&\qquad+O(\|O'\|_\infty K^{-2}),
\end{aligned}\tag{9.25}$$

where $\mathbf{v}^b(s)=\mathbf{v}^b(s,\mathbf{x}(\cdot))$ solves (9.15) with $\mathcal{A}(s)$ given in (9.23).

Proof. If h_0 were a smooth function, then (9.25) would directly follow from (9.17). The general case is a simple cutoff argument using the fact that $h_0\in L^2(d\omega)$ and

$$\begin{aligned}
\mathbb{E}^\omega|\partial_bh_0| &\leq \mathbb{E}^\omega[|(V_{\mathbf{y}})'(x_p)|+|(V_{\widetilde{\mathbf{y}}})'(x_p)|]\\
&\leq \sum_{j\notin I}\mathbb{E}^\omega\left[\frac{1}{|y_j-x_p|}+\frac{1}{|\widetilde{y}_j-x_p|}\right]+C\leq CK^{(C_3+1)\xi}.
\end{aligned}$$

Here we have used (8.20) and the fact that $\mathbf{y},\widetilde{\mathbf{y}}\in\mathcal{R}_{L=0,K}=\mathcal{R}_{L=0,K}(\xi^2\delta/2,\alpha)$ are regular on scale $K^{\xi^2}\leq K^\xi$, so the summation is effectively restricted to K^ξ terms. □

The representation (9.25) expresses the correlation function in terms of the discrete spatial derivative of the solution to (9.15). To estimate $v_p^b(\sigma, \mathbf{x}(\cdot)) - v_{p+1}^b(\sigma, \mathbf{x}(\cdot))$ in (9.25), we will now study the Hölder continuity of the solution $\mathbf{v}^b(s, \mathbf{x}(\cdot))$ to (9.15) at time $s = \sigma$ and at the spatial point p. For any fixed σ we will do so for each fixed path $\mathbf{x}(\cdot)$, with the exception of a set of "bad" paths that will have a small probability.

Notice that if all points x_i are approximately regularly spaced in the interval J, then the operator $\mathcal{B}$ has a kernel $B_{ij} \sim (i-j)^{-2}$, i.e. it is essentially a discrete version of the operator $|p| = \sqrt{-\Delta}$. Hölder continuity will thus be a consequence of the De Giorgi–Nash–Moser bound for the parabolic equation (9.15). However, we need to control the coefficients in this equation, which depend on the random walk $\mathbf{x}(\cdot)$.

For De Giorgi–Nash–Moser theory we need both upper and lower bounds on the kernel B_{ij}. The rigidity bound (8.14) guarantees a lower bound on B_{ij}, up to a factor $K^{-C_2\xi^2} \geq K^{-\xi}$. The level repulsion estimate implies certain upper bounds on B_{ij}, but only in an average sense. In the next section we define a good set of paths that satisfy both requirements.

9.3. *Sets of good paths*

From now on we assume the conditions of Theorem 8.1. In particular we are given some $\xi > 0$ and we assume that the boundary conditions satisfy $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L=0,K} = \mathcal{R}_{L=0,K}(\xi^2\delta/2, \alpha)$ and (8.6) with this ξ. We define the following "good sets":

$$\mathcal{G} := \left\{ \sup_{0 \leq s \leq C_1 K \log K} \sup_{|j| \leq K} |x_j(s) - \alpha_j| \leq K^{\xi'} \right\}, \tag{9.26}$$

where

$$\xi' := (C_2 + 1)\xi^2,$$

with C_2 being the constant in (8.14) and α_j given by (8.7). We recall the definition of the event $\widetilde{\mathcal{Q}}_{\sigma,Z}$ for any $Z \in I$ and $\sigma \in \mathcal{T} = [0, C_1 K \log K]$ from (9.5).

Lemma 9.3. *There exists a positive constant θ, depending on $\xi' = (C_2+1)\xi^2$, such that*

$$\mathbb{P}^{\omega}(\mathcal{G}^c) \leq Ce^{-K^{\theta}}. \tag{9.27}$$

Moreover, there is a constant C_4, depending on the constant C_2 in (8.14) and on C_3 in (8.17), (8.18) such that for any ξ and ρ small enough, we have

$$\mathbb{P}^{\omega}(\widetilde{\mathcal{Q}}^c_{\sigma,Z}) \leq CK^{C_4\xi - \rho} \tag{9.28}$$

for each fixed $Z \in I$ and fixed $\sigma \in \mathcal{T}$.

Proof. From the stochastic differential equation (8.29) of the dynamics we have

$$\begin{aligned} |x_i(t) - x_i(s)| \leq C|t-s| + \int_s^t \left[\sum_{\substack{j \in I \\ j \neq i}} \frac{1}{|x_j(a) - x_i(a)|} + \sum_{j \in I^c} \frac{1}{|y_j - x_i(a)|} \right] da \\ + |B_i(t) - B_i(s)|. \end{aligned} \tag{9.29}$$

Using (8.20) and the invariance of $\mathbf{x}(\cdot)$ under ω, we have the bound

$$\mathbb{E}^\omega\Big[\int_s^t \sum_{j\neq i}\frac{1}{|x_j(a)-x_i(a)|}\Big]^{3/2} \le CK^3|t-s|^{3/2}\max_{i\in I}\mathbb{E}^\omega\frac{1}{|x_i-x_{i+1}|^{3/2}}$$
$$\le CK^{3+C_3\xi}|t-s|^{3/2}.$$

This implies for any fixed $s<t\le C_1K\log K$ and for any $R>0$ that

$$\mathbb{P}^\omega\Big[\int_s^t\sum_{j\neq i}\frac{1}{|x_j(a)-x_i(a)|}\ge R\Big]\le CK^{3+C_3\xi}|t-s|^{3/2}R^{-3/2}.$$

A similar bound holds for the second summation in (9.29); the summation over large j can be performed by using the regularity of $\mathbf{y}\in\mathcal{R}_{L=0,K}$.

Set a parameter $q\le cR$ and choose a discrete set of increasing times $\{s_k : k\le (C_1K\log K)/q\}$ such that

$$0=s_0<s_1\le s_2\le\cdots\le C_1K\log K\quad\text{and}\quad |s_k-s_{k+1}|\le q.$$

From standard large deviation bounds on the Brownian motion increment $B_i(t)-B_i(s)$ and from (9.29), we have the stochastic continuity estimate

$$\mathbb{P}^\omega\Big(\sup_{s,t\in[s_k,s_{k+1}],|i|\le K}|x_i(s)-x_i(t)|\ge R\Big)\le Ke^{-CR^2/q}+CK^4q^{3/2}R^{-3/2}$$

for any fixed k. Taking sup over k, and overestimating $C_1K\log K\le K^2$, we have

$$\mathbb{P}^\omega\Big(\sup_{0\le s,t\le C_1K\log K,|t-s|\le q,|i|\le K}|x_i(s)-x_i(t)|\ge R\Big)\le K^3q^{-1}e^{-CR^2/q}+CK^6q^{1/2}R^{-3/2}$$

for any positive q and R with $q\le cR$.

From the rigidity bound (8.14) we know that for some $\theta_3>0$ and for any fixed k

$$\mathbb{P}^\omega\{|x_j(s_k)-\alpha_j|\ge CK^{C_2\xi^2}\}\le Ce^{-K^{\theta_3}},\qquad j\in I.$$

Choosing $R=K^{\xi'}/2$ and $q=\exp(-K^{\theta_3/2})$, and using $CK^{C_2\xi^2}\le K^{\xi'}/2$ with the choice of ξ', we have

$$\mathbb{P}^\omega(\mathcal{G}^c)\le Ce^{-K^{\theta_3}}K^3q^{-1}+K^3q^{-1}e^{-CR^2/q}+CK^6q^{1/2}R^{-3/2}\le C\exp(-K^{\theta_3/3})$$

for sufficiently large K, and this proves (9.27) with $\theta=\theta_3/3$.

We will now prove (9.28). The number of intersections in the definition of $\widetilde{Q}_{\sigma,Z}$ is only a $(\log K)$-power, so it will be sufficient to prove (9.28) for one set $\mathcal{Q}^c$. We will consider only the set $\mathcal{Q}^c_{\sigma,Z}$ and only for $Z=0$ and $\sigma=0$. The modification needed for the general case is only notational. We start the proof by noting that for $s>0$,

$$\frac{1}{1+s'}\int_0^{s'}da\,\frac{1}{M'}\sum_{i=-M'}^{M'}\frac{1}{|x_i(a)-x_{i+1}(a)|^2}\le C\frac{1}{1+s}\int_0^s da\,\frac{1}{M}\sum_{i=-M}^{M}\frac{1}{|x_i(a)-x_{i+1}(a)|^2}$$

holds for any $s' \in [s/2, s]$ and $M' \in [M/2, M]$. Hence it is enough to estimate

$$\mathbb{P}^{\omega}\left\{\frac{1}{1+s}\int_0^s da\,\frac{1}{M}\sum_{i=-M}^{M}\frac{1}{|x_i(a)-x_{i+1}(a)|^2}\geq K^{\rho}\right\} \tag{9.30}$$

for fixed dyadic points $(s, M) = \{(2^{-p_1}K^2, 2^{-p_2}K)\}$ in space-time for any integers $p_1, p_2 \leq C\log K$. Since the cardinality of the set of these dyadic points is just $C(\log K)^2$, it suffices to estimate (9.30) only for a fixed s, M.

The proof is different for $\beta = 1$ and $\beta > 1$. In the latter case, from (8.20) we see that the random variable in (9.30) has expectation $CK^{C_3\xi}$. Thus the probability in (9.30) is bounded by $CK^{C_3\xi-\varrho}$, so (9.28) holds in this case with C_4 slightly larger than C_3+1 to accommodate the $\log K$ factors.

In the case $\beta = 1$ the random variable in (9.30) has a logarithmically divergent expectation. To prove (9.28) for $\beta = 1$, we need to regularize the interaction on a very small scale of order K^{-C} with a large constant C. This regularization is a minor technical detail which does not affect other parts of this paper. We now explain how it is introduced, but for simplicity we will not indicate it in the notation in the subsequent sections.

For any $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}$ satisfying (4.8) and for $\varepsilon > 0$, we define the extension $\omega^{\varepsilon} := \omega^{r,\varepsilon}_{\mathbf{y},\widetilde{\mathbf{y}}}$ of the measure $\omega = \omega^{r}_{\mathbf{y},\widetilde{\mathbf{y}}}$ (see (8.9)) from the simplex $J^{\mathcal{K}} \cap \Xi^{(\mathcal{K})}$ to $\mathbb{R}^{\mathcal{K}}$ by replacing the singular logarithm with a C^2-function. For $\mathbf{x} \in \mathbb{R}^{\mathcal{K}}$ and $a := |J| \sim K$ we set

$$\begin{aligned}
\mathcal{H}_{\varepsilon}(\mathbf{x}) &:= \frac{1}{2}\sum_{i\in I}U^{\varepsilon}(x_i) - \sum_{i<j}\log_{a\varepsilon}(x_j - x_i),\\
U^{\varepsilon}(x) &:= U^{r,\varepsilon}_{\mathbf{y},\widetilde{\mathbf{y}}}(x) = (1-r)V^{\varepsilon}_{\mathbf{y}}(x) + r\widetilde{V}^{\varepsilon}_{\widetilde{\mathbf{y}}}(x),\\
V^{\varepsilon}_{\mathbf{y}}(x) &:= NV(x/N) - 2\sum_{k<-K}\log_{a\varepsilon}(x-y_k) - 2\sum_{k>K}\log_{a\varepsilon}(y_k - x),
\end{aligned}$$

where we define

$$\log_{\varepsilon}(x) := \mathbf{1}(x\geq\varepsilon)\log x + \mathbf{1}(x<\varepsilon)\left\{\log\varepsilon + \frac{x-\varepsilon}{\varepsilon} - \frac{1}{2\varepsilon^2}(x-\varepsilon)^2\right\}.$$

We remark that the same regularization for a different purpose was introduced in [25, Appendix A]. It is easy to check that $\log_{\varepsilon}$ is in $C^2(\mathbb{R})$, is concave, and satisfies

$$\lim_{\varepsilon\to 0}\log_{\varepsilon}(x) = \begin{cases}\log x & \text{if } x>0,\\ -\infty & \text{if } x\leq 0.\end{cases}$$

Furthermore, we have the lower bound

$$\partial_x^2\log_{\varepsilon}(x) \geq \begin{cases}-1/x^2 & \text{if } x>\varepsilon,\\ -1/\varepsilon^2 & \text{if } x\leq\varepsilon.\end{cases} \tag{9.31}$$

We then define

$$\omega^{\varepsilon}(d\mathbf{x}) := Z_{\varepsilon}^{-1}e^{-\beta\mathcal{H}_{\varepsilon}(\mathbf{x})}d\mathbf{x} \quad \text{on } \mathbb{R}^{\mathcal{K}}, \quad \text{where} \quad Z_{\varepsilon} := \int e^{-\beta\mathcal{H}_{\varepsilon}(\mathbf{x})}d\mathbf{x}.$$

Notice that on the support of ω^ε the particles do not necessarily keep their natural order and they are not confined to the interval J. We recall that $\omega^{r=0}_{\mathbf{y},\widetilde{\mathbf{y}}} = \mu_{\mathbf{y}}$ and $\omega^{r=1}_{\mathbf{y},\widetilde{\mathbf{y}}} = \widetilde{\mu}_{\widetilde{\mathbf{y}}}$, so these definitions also regularize the initial local measures in Theorem 4.1.

In order to apply the proof of Theorem 4.1 to ω^ε, we need two facts. First, ω and ω^ε are close in the entropy sense, i.e.

$$S(\omega|\omega^\varepsilon) \le CK^C\varepsilon^2,$$

Using this entropy bound with $\varepsilon = K^{-C'}$ for a sufficiently large C', we see that the measures $\mu_{\mathbf{y}}$ and $\widetilde{\mu}_{\widetilde{\mathbf{y}}}$ can be replaced with their regularized versions $\mu^\varepsilon_{\mathbf{y}}$, $\widetilde{\mu}^\varepsilon_{\widetilde{\mathbf{y}}}$ both in the condition (4.10) and in the statement (4.11). We can now use the argument of Section 8 with the regularized measures.

The second fact is that the rigidity and level repulsion estimates given in Lemma 8.2 also hold for the regularized measure ω^ε. In fact, apart from the rigidity in the form of (8.14), we also need the following weaker level repulsion bound:

$$\mathbb{P}^{\omega^\varepsilon}(x_{i+1} - x_i \le s) \le CK^{C\xi}s^2, \qquad i \in [\![L-K-1, L+K]\!],\ s \ge K^\xi\varepsilon.$$

Using (9.31), this bound easily implies

$$\mathbb{E}^{\omega^\varepsilon} \log''_\varepsilon(x_{i+1} - x_i) \le CK^{C\xi}|\log\varepsilon|.$$

Thus the regularized version of the random variable in (9.30) has a finite expectation and we obtain (9.28) also for $\beta = 1$.

With these comments in mind, these two facts can be proved following the same path as the corresponding results in Section 7. The only slight complication is that the particles are not ordered, but for $\varepsilon = K^{-C'}$ the regularized potential strongly suppresses switching order. More precisely, we have

$$\mathbb{P}^{\omega^\varepsilon}(x_{i+1} - x_i \le -Ma\varepsilon) \le e^{-cM^2} \tag{9.32}$$

for any $M \ge K^3$. This inequality follows from the estimate

$$\int_{-\infty}^{-Ma\varepsilon} e^{\log_{a\varepsilon} v} dv \le (a\varepsilon)^2 \int_{-\infty}^{-M} e^{-cu^2}\, du \le e^{-cM^2},$$

since for $M \ge K^3$ all other integrands in the measure ω^ε can be estimated trivially at the expense of a multiplicative error K^{CK^2} that is still negligible when compared with the factor $\exp(-cM^2)$. The estimate (9.32) allows us to restrict the analysis to $x_{i+1} \ge x_i - K^{-C''}$ with some large C''. This condition replaces the strict ordering $x_{i+1} \ge x_i$ that is present in Section 7. This replacement introduces irrelevant error factors that can be easily estimated. This completes the proof of Lemma 9.3. □

In the rest of the paper we will work with the regularized measure ω^ε but for simplicity we will not indicate this regularization in the notation.

9.4. Restrictions to the good paths

9.4.1. Restriction to the set $\mathcal{G}$. Now we show that the expectation (9.25) can be restricted to the good set $\mathcal{G}$ with a small error. We just estimate the complement as

$$\int \sum_{b\in I} |\partial_b h_0(\mathbf{x})| \mathbb{E}_{\mathbf{x}} \mathcal{G}^c [|O'(x_p - x_{p+1})|\,|v_p^b(\sigma) - v_{p+1}^b(\sigma)|]\,\omega(d\mathbf{x})$$
$$\le C\|O'\|_\infty \int \mathbb{E}^\omega \sum_b |\partial_b h_0(\mathbf{x})| \mathcal{G}^c [|v_p^b(\sigma)| + |v_{p+1}^b(\sigma)|].$$

Since $\mathcal{A} \ge 0$ as a $\mathcal{K}\times\mathcal{K}$ matrix, the equation (9.15) is a contraction in L^2. Clearly $\mathcal{A}$ is a contraction in L^1 as well, hence it is a contraction in any L^q, $1 \le q \le 2$, by interpolation. By the Hölder inequality and the L^q-contraction for some $1 < q < 2$, we see that for each fixed $b \in I$,

$$\mathbb{E}^\omega |\partial_b h_0(\mathbf{x})| \mathcal{G}^c |v_p^b(\sigma)| \le [\mathbb{E}^\omega \mathcal{G}^c]^{q/(q-1)} [\mathbb{E}^\omega |\partial_b h_0(\mathbf{x})|^q |v_p^b(\sigma)|^q]^{1/q}$$
$$\le [\mathbb{P}^\omega \mathcal{G}^c]^{q/(q-1)} \Big[\mathbb{E}^\omega |\partial_b h_0(\mathbf{x})|^q \sum_{i\in I} |v_i^p(0)|^q\Big]^{1/q} \le CK^{C_3\xi} e^{-cK^{\theta_4}} \le e^{-cK^{\theta_4}}$$

with some $\theta_4 > 0$. Here we have used (9.27) for the first factor. The second factor was estimated by (8.20) (recall the definition of h_0 from (8.12)). After summing over b, we get

$$\mathbb{E}^\omega \sum_b |\partial_b h_0(\mathbf{x})| \mathcal{G}^c [|O'(x_p - x_{p+1})|\,|v_p^b(\sigma) - v_{p+1}^b(\sigma)|] \le Ce^{-cK^{\theta_4}} \|O'\|_\infty.$$

Therefore, under the conditions of Corollary 9.2, and using the notation $\mathbb{E}^\omega$ for the process, we have

$$\begin{aligned}
&|\langle h_0; O(x_p - x_{p+1})\rangle_\omega| \\
&\qquad \le \frac{1}{2}\|O'\|_\infty \int_0^{C_1 K\log K} \sum_{b\in I} \mathbb{E}^\omega [\mathcal{G} |\partial_b h_0(\mathbf{x})|\,|(v_p^b(\sigma) - v_{p+1}^b(\sigma))|]\,d\sigma \\
&\qquad\quad + O(\|O'\|_\infty K^{-2}),
\end{aligned} \tag{9.33}$$

where $\mathbf{v}^b$ is the solution to (9.15), assuming that the constant C_1 in the upper limit of the integration is large enough.

9.4.2. Restriction to the set $\widetilde{\mathcal{Q}}$ and the decay estimates. The complement of the set $\widetilde{\mathcal{Q}}_{\sigma,Z}$ includes the "bad" paths for which the level repulsion estimate in an average sense does not hold. However, the probability of $\widetilde{\mathcal{Q}}^c_{\sigma,Z}$ is not very small, it is only a small negative power of K (see (9.28)). This estimate would not be sufficient against the time integration of order $C_1 K\log K$ in (9.33); we will have to use an L^1-L^∞ decay property of (9.15) which we now derive. Denote the L^p-norm of a vector $\mathbf{u} = \{u_j : j \in I\}$ by

$$\|\mathbf{u}\|_p = \Big(\sum_{j\in I} |u_j|^p\Big)^{1/p}.$$

Proposition 9.4. *Consider the evolution equation*

$$\partial_s \mathbf{u}(s) = -\mathcal{A}(s)\mathbf{u}(s), \qquad \mathbf{u}(s) \in \mathbb{R}^I = \mathbb{R}^{\mathcal{K}},$$

and fix $\sigma > 0$. *Suppose that for some constant* b *we have*

$$B_{jk}(s) \geq b/(j-k)^2, \qquad 0 \leq s \leq \sigma,\ j \neq k, \tag{9.34}$$

and

$$W_j(s) \geq b/d_j, \qquad d_j := \big||j| - K\big| + 1, \qquad 0 \leq s \leq \sigma. \tag{9.35}$$

Then for any $1 \leq p \leq q \leq \infty$ *we have the decay estimate*

$$\|\mathbf{u}(s)\|_q \leq (sb)^{-(1/p-1/q)}\|\mathbf{u}(0)\|_p, \qquad 0 < s \leq \sigma. \tag{9.36}$$

Proof. We consider only the case $b = 1$; the general case follows by scaling. We follow the idea of Nash and start from the L^2-identity

$$\partial_s \|\mathbf{u}(s)\|_2^2 = -2\mathfrak{a}(s)[\mathbf{u}(s), \mathbf{u}(s)].$$

For each s we can extend $\mathbf{u}(s) : I \to \mathbb{R}^{\mathcal{K}}$ to a function $\widetilde{\mathbf{u}}(s)$ on $\mathbb{Z}$ by defining $\widetilde{u}_j(s) = u_j(s)$ for $|j| \leq K$ and $\widetilde{u}_j(s) = 0$ for $j > |K|$. Dropping the time argument, we have, by the estimates (9.34) and (9.35) with $b = 1$,

$$2\mathfrak{a}(\mathbf{u}, \mathbf{u}) \geq \sum_{i,j\in\mathbb{Z}} \frac{(\widetilde{u}_i - \widetilde{u}_j)^2}{(i-j)^2} \geq c\|\widetilde{\mathbf{u}}\|_4^4\|\widetilde{\mathbf{u}}\|_2^{-2},$$

with some positive constant, where, in the second step, we used the Gagliardo–Nirenberg inequality for the discrete operator $\sqrt{-\Delta}$ (see (B.4) in Appendix B) with $p = 4$, $s = 1$. Thus we have

$$\mathfrak{a}[\mathbf{u}, \mathbf{u}] \geq c\|\mathbf{u}\|_4^4\|\mathbf{u}\|_2^{-2},$$

and the energy inequality

$$\partial_s \|\mathbf{u}\|_2^2 \leq -c\|\mathbf{u}\|_4^4\|\mathbf{u}\|_2^{-2} \leq -c\|\mathbf{u}\|_2^4\|\mathbf{u}\|_1^{-2},$$

using the Hölder estimate $\|\mathbf{u}\|_2 \leq \|\mathbf{u}\|_1^{1/3}\|\mathbf{u}\|_4^{2/3}$. Integrating this inequality from 0 to s we get

$$\|\mathbf{u}(s)\|_2 \leq Cs^{-1/2}\|\mathbf{u}(0)\|_1, \tag{9.37}$$

and similarly $\|\mathbf{u}(2s)\|_2 \leq Cs^{-1/2}\|\mathbf{u}(s)\|_1$. Since the previous proof uses only the time independent lower bounds (9.34), (9.35), we can use duality in the time interval $[s, 2s]$ to obtain

$$\|\mathbf{u}(2s)\|_\infty \leq Cs^{-1/2}\|\mathbf{u}(s)\|_2.$$

Together with (9.37) we have

$$\|\mathbf{u}(2s)\|_\infty \leq Cs^{-1}\|\mathbf{u}(0)\|_1.$$

By interpolation, we have thus proved (9.36). □

In the good set $\mathcal{G}$ (see (9.26)), the bounds (9.34) and (9.35) hold with $b = cK^{-\xi'}$. Hence from the decay estimate (9.36), for any fixed σ, Z, we can insert the other good set $\widetilde{\mathcal{Q}}_{\sigma,Z}$ into the expectation in (9.33). This is obvious since the contribution of its complement is bounded by

$$\begin{aligned}
&\int_0^{C_1 K \log K} d\sigma \sum_b \mathbb{E}^\omega \widetilde{\mathcal{Q}}^c_{\sigma,Z} \mathcal{G} |\partial_b h_0(\mathbf{x})| (v_p^b(\sigma) + v_{p+1}^b(\sigma)) \\
&\quad \le CK^{\xi'} \int_0^{C_1 K \log K} d\sigma\, \sigma^{-\frac{1}{1+\xi}} \mathbb{E}^\omega \Big[\mathcal{G} \Big(\sum_{b \in I} |\partial_b h_0(\mathbf{x})|^{1+\xi} \Big)^{\frac{1}{1+\xi}} \widetilde{\mathcal{Q}}^c_{\sigma,Z} \Big] \\
&\quad \le CK^{2\xi'} \int_0^{C_1 K \log K} d\sigma\, \sigma^{-\frac{1}{1+\xi}} \mathbb{E}^\omega \big[[1 + d(x_K(\sigma))^{-1} + d(x_{-K}(\sigma))^{-1}] \widetilde{\mathcal{Q}}^c_{\sigma,Z} \big] \\
&\quad \le CK^{2\xi'} \\
&\qquad \times \int_0^{C_1 K \log K} d\sigma\, \sigma^{-\frac{1}{1+\xi}} \big[\mathbb{E}^\omega [1 + d(x_K(\sigma))^{-1} + d(x_{-K}(\sigma))^{-1}]^{3/2} \big]^{2/3} [\mathbb{P}^\omega(\widetilde{\mathcal{Q}}^c_{\sigma,Z})]^{1/3} \\
&\quad \le CK^{2\xi'} (C_1 K \log K)^\xi K^{C_3 \xi} K^{(C_4 \xi - \rho)/3},
\end{aligned} \tag{9.38}$$

where in the first line we have used a Hölder inequality with exponents $1+\xi$ and its dual, and in the second line the decay estimate (9.36) with $q = \infty$, $p = 1+\xi$. The purpose of taking a Hölder inequality with a power slightly larger than one was to avoid the logarithmic singularity in the $d\sigma$ integration at $\sigma \sim 0$. In the third line we have split the sum into two parts and used the bound

$$|\partial_b h_0(\mathbf{x})| \le |(V_\mathbf{y})'(x_j) - (\widetilde{V}_{\widetilde{\mathbf{y}}})'(x_j)| \le K^{\xi'}/d(x_b), \tag{9.39}$$

which follows from (8.3) (with ξ replaced by ξ^2 since $\mathbf{y}, \widetilde{\mathbf{y}} \in \mathcal{R}_{L,K}(\xi^2\delta/2, \alpha/2)$). Recall that $d(x)$ is the distance to the boundary (see (8.5)). For indices away from the boundary, $|b| \le K - CK^{\xi'}$, we have $|d(x_b)| \ge K^{-\xi'} \min\{|b-K|, |b+K|\}$ on the set $\mathcal{G}$ that guarantees the finiteness of the sum. For indices near the boundary we have just estimated every term with the worst one, i.e. $b = \pm K$. We have used a Hölder inequality in the fifth line of (9.38) and computed the expectation by using (8.20) in the last line. Hence we have proved the following proposition:

Proposition 9.5. *Suppose that*

$$\rho \ge 12\xi' + 6(C_4 + C_3 + 1)\xi \tag{9.40}$$

with C_3 and C_4 defined in (8.17) and (9.28), respectively. Then for any fixed Z, $p \in I$ with $p \ne K$, we have

$$\begin{aligned}
&|\langle h_0; O(x_p - x_{p+1}) \rangle_\omega| \\
&\qquad \le \frac{1}{2} \|O'\|_\infty \int_0^{C_1 K \log K} \sum_{b \in I} \mathbb{E}^\omega [\widetilde{\mathcal{Q}}_{\sigma,Z} \mathcal{G} |\partial_b h_0(\mathbf{x})|\, |v_p^b(\sigma) - v_{p+1}^b(\sigma)|]\, d\sigma \\
&\qquad\quad + O(\|O'\|_\infty K^{-\rho/6}).
\end{aligned} \tag{9.41}$$

9.5. Short time cutoff and finite speed of propagation

The Hölder continuity of the parabolic equation (9.15) emerges only after a certain time, thus for the small σ regime in the integral (9.41) we need a different argument. Since we are interested in the Hölder continuity around the middle of the interval I (note that $|p| \le K^{1-\xi^*}$ in Theorem 8.1), and the initial condition $\partial_b h_0$ is small if b is in this region, a finite speed of propagation estimate for (9.15) will guarantee that $v_p^b(\sigma)$ is small if σ is not too large.

From now on, we fix $\sigma \le C_1 K \log K$, $|Z| \le K/2$ and a path $\mathbf{x}(\cdot)$, and assume that $\mathbf{x}(\cdot) \in \mathcal{G} \cap \widetilde{\mathcal{Q}}_{\sigma,Z}$. In particular, thanks to the definition of $\mathcal{G}$ and the regularity of the locations α_j, the time dependent coefficients $B_{ij}(s)$ and $W_i(s)$ of the equation (9.15) satisfy (9.34) and (9.35) with $b = K^{-\xi'}$.

We split the summation in (9.41). Fix a positive constant $\theta_5 > 0$. The contribution of the indices $|b| \le K^{1-\theta_5}$ to (9.41) is bounded by

$$\begin{aligned}
&\int_0^{C_1 K \log K} \mathbb{E}^\omega \widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G} \sum_{|b|\le K^{1-\theta_5}} |\partial_b h_0(\mathbf{x})|[v_p^b(\sigma) + v_{p+1}^b(\sigma)]\, d\sigma \\
&\quad \le C \int_0^{C_1 K \log K} \mathbb{E}^\omega \Big[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G} \Big[\sum_{|b|\le K^{1-\theta_5}} |\partial_b h_0(\mathbf{x})| \Big] \times \max_{|b|\le K^{1-\theta_5}} |v_p^b(\sigma)| \Big] d\sigma \\
&\quad \le C K^{\xi'-\theta_5} \int_0^{C_1 K \log K} \mathbb{E}^\omega \Big[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G} \max_{|b|\le K^{1-\theta_5}} |v_p^b(\sigma)| \Big] d\sigma \\
&\quad \le K^{\xi'-\theta_5} \int_0^{C_1 K \log K} \sigma^{-1}\, d\sigma \le K^{2\xi'-\theta_5},
\end{aligned} \tag{9.42}$$

where we neglected the v_{p+1}^b term for simplicity since it can be estimated exactly in the same way. From the second to the third line we have used the fact that

$$|\partial_b h_0(\mathbf{x})| \le \frac{K^{\xi'}}{\min\{|b-K|, |b+K|\} + 1} \le C K^{\xi'-1}, \qquad |b| \le K^{1-\theta_5},$$

holds on the set $\mathcal{G}$, from (9.39) and from the rigidity bound provided by $\mathcal{G}$. To arrive at the last line of (9.42) we have used the $L^1 \to L^\infty$ decay estimate (9.36) and we recall that the singularity $\sigma \sim 0$ can be cut off exactly as in (9.38), i.e. by considering a power slightly larger than 1 in the first line. Note that the set $\widetilde{\mathcal{Q}}_{\sigma,Z}$ played no role in this argument.

Together with (9.41) and with the choice

$$\theta_5 > \rho \tag{9.43}$$

and recalling $\rho \ge 4\xi'$ from (9.40), we have

$$\begin{aligned}
&|\langle h_0; O(x_p - x_{p+1})\rangle_\omega| \\
&\quad \le \frac{1}{2}\|O'\|_\infty \int_0^{C_1 K \log K} \sum_{|b|>K^{1-\theta_5}} \mathbb{E}^\omega[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}|\partial_b h_0(\mathbf{x})|\, |v_p^b(\sigma) - v_{p+1}^b(\sigma)|]\, d\sigma \\
&\qquad + O(\|O'\|_\infty K^{-\rho/6}).
\end{aligned} \tag{9.44}$$

The following lemma provides a finite speed of propagation estimate for the equation (9.15), which will be used to control the short time regime in (9.44). This estimate is not optimal, but it is sufficient for our purpose. The proof will be given in the next section.

Lemma 9.6 (Finite speed of propagation estimate). *Fix $b \in I$ and $\sigma \le C_1 K \log K$. Consider $\mathbf{v}^b(s)$, the solution to (9.15), and assume that the coefficients of $\mathcal{A}$ satisfy*

$$W_i(s) \ge K^{-\xi'}/d_i, \qquad B_{ij}(s) \ge K^{-\xi'}/|i-j|^2, \qquad 0 \le s \le \sigma, \tag{9.45}$$

where $d_i := \min\{|i+K|, |i-K|\} + 1$. Assume that

$$\sup_{0\le s\le\sigma}\sup_{0\le M\le K}\frac{1}{1+s}\int_0^s \frac{1}{M}\sum_{i\in I:\,|i-Z|\le M}\;\sum_{j\in I:\,|j-Z|\le M} B_{ij}(s)\,ds \le CK^{\rho_1} \tag{9.46}$$

for some fixed Z with $|Z| \le K/2$. Then for any $s > 0$,

$$|v_p^b(s)| \le \frac{CK^{\rho_1+2\xi'+1/2}\sqrt{s+1}}{|p-b|}. \tag{9.47}$$

9.6. Proof of the finite speed of propagation estimate, Lemma 9.6

Let $1 \ll \ell \ll K$ be a parameter to be specified later. Split the time dependent operator $\mathcal{A} = \mathcal{A}(s)$ defined in (9.23) into a short range and a long range part, $\mathcal{A} = \mathcal{S} + \mathcal{R}$, with

$$(\mathcal{S}\mathbf{u})_j := -\sum_{k:\,|j-k|\le\ell} B_{jk}(u_k - u_j) + W_j u_j,$$
$$(\mathcal{R}\mathbf{u})_j := -\sum_{k:\,|j-k|>\ell} B_{jk}(u_k - u_j).$$

Note that $\mathcal{S}$ and $\mathcal{R}$ are time dependent. Denote by $U_{\mathcal{S}}(s_1, s_2)$ the semigroup associated with $\mathcal{S}$ from time s_1 to time s_2, i.e.

$$\partial_{s_2} U_{\mathcal{S}}(s_1, s_2) = -\mathcal{S}(s_2) U_{\mathcal{S}}(s_1, s_2)$$

for any $s_1 \le s_2$, and $U_{\mathcal{S}}(s_1, s_1) = I$; the notation $U_{\mathcal{A}}(s_1, s_2)$ is analogous. Then by the Duhamel formula,

$$\mathbf{v}(s) = U_{\mathcal{S}}(0, s)\mathbf{v}_0 + \int_0^s U_{\mathcal{A}}(s', s)\mathcal{R}(s')U_{\mathcal{S}}(0, s')\mathbf{v}_0\,ds'.$$

Notice that for $\ell \gg K^{\xi'}$ and for $\mathbf{x}(\cdot)$ in the good set $\mathcal{G}$ (see (9.26)), we have

$$\|\mathcal{R}\mathbf{u}\|_1 = \sum_{|j|\le K}\left|\sum_{k:\,|j-k|\ge\ell}\frac{1}{(x_j-x_k)^2}u_k\right| \le C\ell^{-1}\|\mathbf{u}\|_1,$$

or more generally,

$$\|\mathcal{R}\mathbf{u}\|_p \le C\ell^{-1}\|\mathbf{u}\|_p, \qquad 1 \le p \le \infty. \tag{9.48}$$

Recall the decay estimate (9.36) for the semigroup $U_{\mathcal{A}}$ that is applicable by (9.45). Hence we have, for $s \geq 2$,

$$\int_0^s \|U_{\mathcal{A}}(s', s)\mathcal{R}(s')U_{\mathcal{S}}(0, s')\mathbf{v}_0\|_\infty \, ds' \\ \leq K^{\xi'} \int_0^s (s - s')^{-1} \|\mathcal{R}(s')U_{\mathcal{S}}(0, s')\mathbf{v}_0\|_1 \, ds' \leq K^{\xi'} \ell^{-1} (\log s)\|\mathbf{v}_0\|_1,$$

where we have used that $U_{\mathcal{S}}$ is a contraction on L^1. The nonintegrable short time singularity for s' very close to s, $|s - s'| \leq K^{-C}$, can be removed by using the $L^p \to L^\infty$ bound (9.36) with some $p > 1$, invoking a similar argument in (9.38). In this short time cutoff argument we use the fact that $U_{\mathcal{S}}$ is an L^p-contraction for any $1 \leq p \leq 2$ by interpolation, and that the rate of the $L^p \to L^\infty$ decay of $U_{\mathcal{A}}$ is given in (9.36). Consequently,

$$\|\mathbf{v}(s) - U_{\mathcal{S}}(0, s)\mathbf{v}_0\|_\infty \leq \ell^{-1}(\log s)K^{\xi'} \leq C\ell^{-1}(\log K)K^{\xi'}, \tag{9.49}$$

where we have used that $\mathbf{x}(\cdot)$ is in the good set $\mathcal{G}$ and that $s \leq C_1 K \log K$.

We now prove a cutoff estimate for the short range dynamics. Let $\mathbf{r}(s) := U_{\mathcal{S}}(0, s)\mathbf{v}_0$ and define

$$f(s) = \sum_j \phi_j r_j^2(s), \qquad \phi_j = e^{|j-b|/\theta},$$

with some parameter $\theta \geq \ell$ to be specified later. Recall that b is the location of the initial condition, $\mathbf{v}_0 = \delta_b$. In particular, $f(0) = 1$.

Differentiating f and using $W_j \geq 0$, we have

$$\begin{aligned} f'(s) = \partial_s \sum_j \phi_j r_j^2(s) &\leq 2 \sum_j \phi_j \sum_{k:\,|j-k|\leq\ell} r_j(s) B_{kj}(s)(r_k - r_j)(s) \\ &= \sum_{|j-k|\leq\ell} B_{kj}(s)(r_k - r_j)(s)[r_j(s)\phi_j - r_k(s)\phi_k] \\ &= \sum_{|j-k|\leq\ell} B_{kj}(s)(r_k - r_j)(s)\phi_j[r_j - r_k](s) \\ &\quad + \sum_{|j-k|\leq\ell} B_{kj}(s)(r_k - r_j)(s)[\phi_j - \phi_k] r_k(s). \end{aligned}$$

In the second term we use the Schwarz inequality and absorb the quadratic term in $r_k - r_j$ into the first term that is negative. Assuming $\ell \leq \theta$, we have $\phi_k^{-2}[\phi_j - \phi_k]^2 \leq C\ell^2/\theta^2$ for $|j - k| \leq \ell$. Thus

$$\begin{aligned} f'(s) &\leq C \sum_{|j-k|\leq\ell} B_{kj}(s)\phi_k^{-1}[\phi_j - \phi_k]^2 r_k^2(s) \\ &\leq C\theta^{-2}\ell^2 \Big(\sum_{k',j:\,|j-k'|\leq\ell} B_{k'j}(s) \Big) \sum_k \phi_k r_k^2(s). \end{aligned}$$

From a Gronwall argument we have

$$f(s) \le \exp\left[C\theta^{-2}\ell^2 \int_0^s \sum_{k,j:\,|j-k|\le \ell} B_{kj}(s')\,ds'\right] f(0).$$

From the assumption (9.46) with $M = K$ and any Z, we can bound the integration in the exponent by

$$\int_0^s \sum_{k,j:\,|j-k|\le K} B_{kj}(s')\,ds' \le K^{1+\rho_1}(s+1).$$

Thus we have

$$\sum_j e^{|j-b|/\theta} r_j^2(s) = f(s) \le \exp[\theta^{-2}\ell^2 K^{1+\rho_1}(s+1)] f(0) \le C, \tag{9.50}$$

provided that we choose

$$\theta = \ell K^{(\rho_1+1)/2}\sqrt{s+1}.$$

In particular, this shows the following exponential finite speed of propagation estimate for the short range dynamics:

$$r_j(s) \le C\exp\left(-\frac{|j-b|}{\ell K^{(\rho_1+1)/2}\sqrt{s+1}}\right).$$

Now we choose $\ell = |p-b|K^{-\xi'-(\rho_1+1)/2}(s+1)^{-1/2}$ so that $e^{|p-b|/\theta} \ge \exp(K^{\xi'})$. Using this choice in (9.50) and (9.49) to estimate $v_p^b(s)$, we have thus proved that

$$|v_p^b(s)| \le \ell^{-1}(\log K)K^{\xi'} + Ce^{-K^{-\xi'}} \le \frac{K^{2\xi'+(\rho_1+1)/2}\sqrt{s+1}}{|p-b|}.$$

This concludes the proof of Lemma 9.6.

9.7. *Completing the proof of Theorem 8.1*

In this section we complete the proof of Theorem 8.1 assuming a discrete version of the De Giorgi–Nash–Moser Hölder regularity estimate for the solution (9.1) (Theorem 9.8 below).

Notice that on the set $\mathcal{G} \cap \widetilde{\mathcal{Q}}_{\sigma,Z}$ the conditions of Lemma 9.6 are satisfied, especially (9.46) with the choice

$$\rho_1 := \rho + \xi' \tag{9.51}$$

follows from the definition (9.4) since for the summands with $|i-j| \ge K^{\xi'}$ in (9.56) we can use $B_{ij} \le C|\alpha_i - \alpha_j|^{-2} \le C|i-j|^{-2}$. Thus we can use (9.47) to estimate the short time integration regime in (9.44). Setting

$$\theta_5 := \min\{\xi^*/2, 1/100\}, \tag{9.52}$$

we obtain, for any $|Z| \leq 2K^{1-\xi^*}$ and $|p| \leq K^{1-\xi^*}$,

$$\begin{aligned}
&\int_0^{K^{1/4}} \sum_{|b|>K^{1-\theta_5}} \mathbb{E}^\omega[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}|\partial_b h_0(\mathbf{x})|\,|v_p^b(\sigma) - v_{p+1}^b(\sigma)|]\,d\sigma \\
&\qquad \leq C\int_0^{K^{1/4}} \mathbb{E}^\omega\Big[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}\sum_{|b|>K^{1-\theta_5}} |\partial_b h_0(\mathbf{x})|v_p^b(\sigma)\Big]\,d\sigma \\
&\qquad \leq CK^{2\xi'+\rho_1+1/2+1/4+1/8-(1-\theta_5)}\mathbb{E}^\omega\Big[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}\sum_{|b|>K^{1-\theta_5}} |\partial_b h_0(\mathbf{x})|\Big] \\
&\qquad \leq CK^{4\xi'+\rho_1+\theta_5-1/8}\mathbb{E}^\omega\Big[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}\Big(\frac{1}{d(x_K)} + \frac{1}{d(x_{-K})}\Big)\Big] \\
&\qquad \leq CK^{4\xi'+\rho_1+C_3\xi+\theta_5-1/8} \leq K^{-1/10}
\end{aligned} \tag{9.53}$$

provided that

$$4\xi' + \rho_1 + C_3\xi \leq 1/100. \tag{9.54}$$

In the third line above we have used (9.47) together with $|p-b| \geq \frac{1}{2}K^{1-\theta_5}$. This latter bound follows from $|b| > K^{1-\theta_5}$ and $|p| \leq K^{1-\xi^*}$ and from the choice $\theta_5 < \xi^*$. In the fourth line we have used (9.39) and that on the set $\mathcal{G}$ we have

$$\sum_j \frac{1}{d(x_j)} \leq (\log K)K^{\xi'}\Big[\frac{1}{d(x_K)} + \frac{1}{d(x_{-K})}\Big].$$

Moreover, in the last step we have used (8.20). This completes the estimate for the small σ regime. Notice that the set $\widetilde{\mathcal{Q}}_{\sigma,Z}$ did not play a role in this argument.

After the short time cutoff (9.53), we finally have to control the regime of large time and large b-indices, i.e.

$$\int_{K^{1/4}}^{C_1 K\log K} \sum_{|b|>K^{1-\theta_5}} \mathbb{E}^\omega[\widetilde{\mathcal{Q}}_{\sigma,Z}\mathcal{G}|\partial_b h_0(\mathbf{x})|\,|v_p^b(\sigma) - v_{p+1}^b(\sigma)|]\,d\sigma$$

from (9.44). We will exploit the Hölder regularity of the solution $\mathbf{v}^b$ to (9.15). We will assume that the coefficients of $\mathcal{A}$ in (9.15) satisfy a certain regularity condition.

Definition 9.7. The equation

$$\partial_t \mathbf{v}(t) = -\mathcal{A}(t)\mathbf{v}(t), \qquad \mathcal{A}(t) = \mathcal{B}(t) + \mathcal{W}(t), \qquad t \in \mathcal{T}, \tag{9.55}$$

is called *regular* at the space-time point $(Z,\sigma) \in I \times \mathcal{T}$ with exponent ρ if

$$\sup_{s\in\mathcal{T}} \sup_{1\leq M\leq K} \frac{1}{1+|s-\sigma|}\left|\int_s^\sigma \frac{1}{M}\sum_{i\in I:\,|i-Z|\leq M}\;\sum_{j\in I:\,|j-Z|\leq M} B_{ij}(u)\,du\right| \leq K^\rho. \tag{9.56}$$

Furthermore, the equation is called *strongly regular* at the space-time point $(Z,\sigma) \in I\times\mathcal{T}$ with exponent ρ if it is regular at all points $\{Z\}\times\{\Xi+\sigma\}$, where we recall the definition of Ξ from (9.6):

$$\Xi = \{-K\cdot 2^{-m}(1+2^{-k}) : 0 \leq k, m \leq C\log K\}.$$

Fix a $Z \in I$ with $|Z| \le K/2$ and a $\sigma \in \mathcal{T}$. Recall that on $\mathcal{G} \cap \mathcal{Q}_{\sigma,Z}$ the regularity at $(\underset{\sim}{p}, \sigma)$ with exponent ρ_1 from (9.51) follows from (9.4). Analogously, on the event $\mathcal{G} \cap \widetilde{\mathcal{Q}}_{\sigma,Z}$, the strong regularity at (Z, σ) with a slightly increased exponent ρ_1 holds.

We formulate the partial Hölder regularity result for the equation (9.55). We collect the following facts on the coefficients $B_{ij}(s)$ and $W_i(s)$ that follow from $\mathbf{x}(\cdot) \in \mathcal{G}$:

$$B_{ij}(s) \ge K^{-\xi'}/|i-j|^2, \qquad W_i(s) \ge K^{-\xi'}/d_i \quad \text{ for any } s \in \mathcal{T},\ i, j \in I, \tag{9.57}$$

$$W_i(s) \le K^{\xi'}/d_i \quad \text{ for any } s \in \mathcal{T},\ d_i \ge K^{C\xi'}, \tag{9.58}$$

and

$$\frac{1}{C(i-j)^2} \le B_{ij}(s) \le \frac{C}{(i-j)^2} \quad \text{ for any } s \in \mathcal{T},\ |i-j| \ge \widehat{C} K^{\xi'}. \tag{9.59}$$

Theorem 9.8. *There exists a universal constant* $\mathfrak{q} > 0$ *with the following properties. Let* $\mathbf{v}(t) = \mathbf{v}^b(t)$ *be a solution to* (9.55) *for any choice of* $b \in I$*, with initial condition* $v_j^b(0) = \delta_{jb}$*. Let* $Z \in I$ *with* $|Z| \le K/2$ *and* $\sigma \in [K^{c_3}, C_1 K \log K]$ *be fixed, where* $c_3 > 0$ *is an arbitrary positive constant. There exist positive constants* ξ_0*,* ρ_0 *(depending only on* c_3*) such that if the coefficients of* $\mathcal{A}$ *satisfy* (9.57)–(9.59) *with some* $\xi' \le \xi_0$ *and the equation is strongly regular at the point* (Z, σ) *with an exponent* $\rho_1 \le \rho_0$ *then for any* $\alpha \in [0, 1/3]$ *we have*

$$\sup_{|j-Z|+|j'-Z| \le \sigma_1^{1-\alpha}} |v_j(\sigma) - v_{j'}(\sigma)| \le C K^{\xi'} \sigma^{-1-\frac{1}{2}\mathfrak{q}\alpha}, \qquad \sigma_1 := \min\{\sigma, K^{1-c_3}\}, \tag{9.60}$$

where $\mathbf{v} = \mathbf{v}^b$ *for any choice of* b*. The constant* C *in* (9.60) *depends only on* c_3*.*

Theorem 9.8 follows directly from the slightly more general Theorem 10.2 presented in Section 10 and it will be proved there.

Armed with Theorem 9.8, we now complete the proof of Theorem 8.1. As we already remarked, the conditions of Theorem 9.8 are satisfied on the set $\widetilde{\mathcal{Q}}_{\sigma,Z} \cap \mathcal{G}$ with some small universal constants ρ_0, ξ_0. For any $|p| \le K^{1-\xi^*}$ fixed, we choose $Z = p$ (in fact, we could choose any Z with $|Z - p| \le C$). Using (9.39), we have, for the large time integration regime in (9.44),

$$\begin{aligned}
&\int_{K^{1/4}}^{C_1 K \log K} \sum_{|b| > K^{1-\theta_5}} \mathbb{E}^{\omega}[\widetilde{\mathcal{Q}}_{\sigma,p} \mathcal{G} |\partial_b h_0(\mathbf{x})|\, |v_p^b(\sigma) - v_{p+1}^b(\sigma)|]\, d\sigma \\
&\quad \le C K^{\xi'} \int_{K^{1/4}}^{C_1 K \log K} \mathbb{E}^{\omega}\Bigg[\widetilde{\mathcal{Q}}_{\sigma,p} \mathcal{G} \sum_{|b| > K^{1-\theta_5}} \frac{1}{d(x_b)} |v_p^b(\sigma) - v_{p+1}^b(\sigma)|\Bigg] d\sigma \\
&\quad \le C K^{2\xi'} \int_{K^{1/4}}^{C_1 K \log K} \sigma^{-1-1/6\mathfrak{q}} \mathbb{E}^{\omega}\Bigg[\widetilde{\mathcal{Q}}_{\sigma,p} \mathcal{G} \sum_{|b| > K^{1-\theta_5}} \frac{1}{d(x_b)}\Bigg] d\sigma \\
&\quad \le C K^{3\xi' + \rho_1 + C_3\xi - \mathfrak{q}/24}.
\end{aligned} \tag{9.61}$$

In the third line we have used Theorem 9.8 with $c_3 = 1/4$ and $\alpha = 1/3$; and in the last line we have used a similar argument to the last step of (9.53).

Finally, from (9.44), (9.53) and (9.61) and $\rho_1 = \rho + \xi'$ we have

$$|\langle h_0; O(x_p - x_{p+1})\rangle_\omega| \\ \leq C\|O'\|_\infty\big(K^{4\xi'+\rho+C_3\xi-\mathfrak{q}/24} + O(K^{-1/10}) + O(K^{-\rho/6})\big). \tag{9.62}$$

For a given $\xi^* > 0$, recall that we defined $\theta_5 := \min\{\xi^*/2, 1/100\}$ and we now choose

$$\rho := \min\{\mathfrak{q}/100, \theta_5/2\} = \min\{\mathfrak{q}/100, \xi^*/4, 1/200\}, \tag{9.63}$$

in particular (9.43) is satisfied. Since $\mathfrak{q} > 0$ is a universal constant, it is then clear that for any sufficiently small ξ all conditions in (9.54) and (9.40) on the exponents $\xi, \xi' = (C_2+1)\xi^2$ and $\rho_1 = \rho + \xi'$ can be simultaneously satisfied. Therefore we can make the error term in (9.62) smaller than $K^{C\xi}K^{-\rho/6}$. With the choice of $\varepsilon = \rho/6$, where ρ is from (9.63), we have thus proved Theorem 8.1. □

Although the choices of parameters seem to be complicated, the underlying mechanism is that there is a universal positive exponent $\mathfrak{q}$ in (9.60). This exponent provides an extra smallness factor in addition to the natural size of $v_j(\sigma)$, which is σ^{-1} from the $L^1 \to L^\infty$ decay. As (9.60) indicates, this gain comes from a Hölder regularity on the relevant scale. The parameters ξ, ξ' and ξ^* can be chosen arbitrarily small (without affecting the value of $\mathfrak{q}$). These parameters govern the cutoff levels in the regularization of the coefficients of $\mathcal{A}$. There are other minor considerations due to an additional cutoff for small time where we have to use a finite speed estimate. But the arguments for this part are of technical nature and most estimates are not optimized. We have just worked out estimates sufficient to prove Theorem 8.1. The choices of exponents related to the various cutoffs do not have intrinsic meanings.

As a guide to the reader, our choices of parameters, roughly speaking, are given by the following rule: We first fix a small parameter ξ^*. Then we choose the cutoff parameter θ_5 to be slightly smaller than ξ^*, (9.52). The exponent ρ in (9.4) has a lower bound by ξ and ξ' in (9.40). On the other hand, ρ will affect the cutoff bound and so we have the condition $\rho < \theta_5$ (i.e., (9.43)). So we choose $\rho \lesssim \xi^*$ and make ξ, ξ' very small so that the lower bound requirement on ρ is satisfied. Finally, if $\xi^* \leq \mathfrak{q}/100$, we can use the gain from the Hölder continuity to compensate all the errors which depend only on ξ, ξ', ξ^*.

10. A discrete De Giorgi–Nash–Moser estimate

In this section we prove Theorem 9.8, which is a Hölder regularity estimate for the parabolic evolution equation

$$\partial_s \mathbf{u}(s) = -\mathcal{A}(s)\mathbf{u}(s), \tag{10.1}$$

where $\mathcal{A}(s) = \mathcal{B}(s) + \mathcal{W}(s)$ are symmetric matrices defined by

$$[\mathcal{B}(s)\mathbf{u}]_j = -\sum_{k\neq j\in I} B_{jk}(s)(u_k - u_j), \qquad [\mathcal{W}(s)\mathbf{u}]_i = W_i(s)u_i \tag{10.2}$$

and $B_{ij}(s) \geq 0$. Here $I = \{-K, -K+1, \ldots, K\}$ and $\mathbf{u} \in \mathbb{C}^I$. We will study this equation in a time interval $\mathcal{T} \subset \mathbb{R}$ of length $|\mathcal{T}| = \sigma$ and we will assume that $\sigma \in [K^{c_3}, CK \log K]$. The reader can safely think of $\sigma = CK \log K$. In the applications we set $\mathcal{T} = [0, \sigma]$, but we give some definitions more generally. The reason is that traditionally in the regularity theory for parabolic equations one sets the initial condition $\mathbf{u}(-\sigma)$ at some negative time $-\sigma < 0$ and one is interested in the regularity of the solution $\mathbf{u}(s)$ around $s = 0$. In this case $\mathcal{T}$ starts at $-\sigma$, so in this section $\mathcal{T} = [-\sigma, 0]$. This convention is widely used for parabolic equations and in particular in [13]. Later on in our application, we will need to make an obvious shift in time.

We will now state a general Hölder continuity result, Theorem 10.1, concerning the deterministic equation (10.1) over the finite set I and on the time interval $\mathcal{T} = [-\sigma, 0]$. Theorem 10.1 will be a local Hölder continuity result around an interior point $Z \in I$ separated away from the boundary. We recall the definition of strong regularity from Definition 9.7. The following conditions on $\mathcal{A}$ will be needed that are characterized by two parameters $\xi, \rho > 0$.

$(\mathbf{C1})_\rho$ The equation (10.1) is strongly regular with exponent ρ at the space-time point $(Z, 0)$.

$(\mathbf{C2})_\xi$ Denote by $d_i = d_i^I := \min\{|i + K + 1|, |1 + K - i|\}$ the distance of i to the boundary of I. For some large constants $C, \widehat{C} \geq 10$, the following conditions are satisfied:

$$B_{ij}(s) \geq K^{-\xi}/|i-j|^2 \quad \text{for any } s \in \mathcal{T},\ d_i \geq K/C,\ d_j \geq K/C, \tag{10.3}$$

$$W_i(s) \leq K^\xi/d_i \quad \text{if } d_i \geq K^{C\xi},\ s \in \mathcal{T}, \tag{10.4}$$

$$\frac{\mathbf{1}(\min\{d_i, d_j\} \geq K/C)}{C(i-j)^2} \leq B_{ij}(s) \leq \frac{C}{(i-j)^2} \quad \text{if } |i-j| \geq \widehat{C} K^\xi \text{ and } s \in \mathcal{T}. \tag{10.5}$$

Theorem 10.1 (Parabolic regularity with singular coefficients). *There exists a universal constant $\mathfrak{q} > 0$ such that the following holds. Consider the equation (10.1) on the time interval $\mathcal{T} = [-\sigma, 0]$ with some $\sigma \in [K^{c_3}, K^{1-c_3}]$, where $c_3 > 0$ is a positive constant. Fix $|Z| \leq K/2$ and $\alpha \in [0, 1/3]$. Suppose that $(\mathbf{C1})_\rho$ and $(\mathbf{C2})_\xi$ hold with some exponents ξ, ρ small enough depending on c_3. Then for the solution $\mathbf{u}$ to (10.1) we have*

$$\sup_{|j-Z|+|j'-Z| \leq \sigma^{1-\alpha}} |u_j(0) - u_{j'}(0)| \leq C\sigma^{-\mathfrak{q}\alpha} \|\mathbf{u}(-\sigma)\|_\infty. \tag{10.6}$$

The constant C in (10.6) depends only on c_3 and is uniform in K. The result holds for any $K \geq K_0$, where K_0 depends on c_3.

We remark that the upper bound $\sigma \leq K^{1-c_3}$ is not an important condition, it is imposed only for convenience to state (10.6) with a single scaling parameter. More generally, for any $\sigma \geq K^{c_3}$ we have

$$\sup_{|j-Z|+|j'-Z| \leq \sigma_1^{1-\alpha}} |u_j(0) - u_{j'}(0)| \leq C\sigma_1^{-\mathfrak{q}\alpha} \|\mathbf{u}(-\sigma)\|_\infty. \tag{10.7}$$

where $\sigma_1 := \min\{\sigma, K^{1-c_3}\}$. If $\sigma \geq K^{1-c_3}$, then (10.7) immediately follows by noticing that $\|u(-\sigma_1)\|_\infty \leq \|u(-\sigma)\|_\infty$ and applying (10.6) with $\sigma_1 = K^{1-c_3}$ instead of σ.

To understand why Theorem 10.1 is a Hölder regularity result, we rescale the solution so that the equation runs up to a time of order one. That is, for a given $\sigma \ll 1$ we define the rescaled solution

$$U(T, X) := u_{[\sigma X]+Z}(T\sigma), \qquad \sigma \gg 1$$

(where $[\cdot]$ denotes the integer part). Then the bound (10.6) says that

$$\sup_{|X|+|Y|\leq\varepsilon} |U(0, X) - U(0, Y)| \leq C\varepsilon^{\mathfrak{q}} \|U(-1, \cdot)\|_\infty, \qquad \varepsilon \in [\sigma^{-1/3}, 1].$$

Thus, in the macroscopic coordinates (T, X) the Hölder regularity for U holds around $(0, 0)$ from order one scales down to order $\sigma^{-1/3}$ scales. Note that Hölder regularity holds only at one space-time point, since the strong regularity condition $(\mathbf{C1})_\rho$ was centered around a given space-time point $(Z, 0)$ in miscroscopic coordinates.

Notice that by imposing the regularity condition we only require the time integration of the singularity of B_{ij} to be bounded. Thus we substantially weaken the standard assumption in parabolic regularity theory on the supremum bound on the ellipticity.

Theorem 10.1 is a Hölder regularity result with L^∞ initial data. Combining it with the decay estimate of Proposition 9.4, we get a Hölder regularity result with L^1 initial data. However, for the application of the decay estimate, we need to strengthen (10.3) to

$$B_{ij}(s) \geq K^{-\xi}/|i-j|^2, \qquad W_i(s) \geq K^{-\xi}/d_i \quad \text{for any } s \in \mathcal{T},\ i, j \in I. \tag{10.8}$$

Let $(\mathbf{C2})^*_\xi$ be the condition identical to $(\mathbf{C2})_\xi$ except that (10.3) is replaced with (10.8).

Theorem 10.2. *There exists a universal constant $\mathfrak{q} > 0$ such that the following holds. Consider the equation (10.1) on the time interval $\mathcal{T} = [-\tau - \sigma, 0]$ with some $\tau > 0$ and $\sigma \in [K^{c_3}, K^{1-c_3}]$, where $c_3 > 0$ is a positive constant. Fix $|Z| \leq K/2$ and $\alpha \in [0, 1/3]$. Suppose that $(\mathbf{C1})_\rho$ and $(\mathbf{C2})^*_\xi$ hold with some small exponents ξ, ρ depending on c_3. Then for the solution $\mathbf{u}$ to (10.1) we have*

$$\sup_{|j-Z|+|j'-Z|\leq\sigma^{1-\alpha}} |u_j(0) - u_{j'}(0)| \leq CK^\xi \sigma^{-\mathfrak{q}\alpha} \tau^{-1} \|\mathbf{u}(-\tau-\sigma)\|_1. \tag{10.9}$$

The constant C in (10.6) depends only on c_3 and is uniform in K. The result holds for any $K \geq K_0$, where K_0 depends on c_3.

Proof. We can apply Proposition 9.4 with $b = K^{-\xi}$, $p = 1$, $q = \infty$ on the time interval $[-\tau - \sigma, -\sigma]$. Then (9.36) asserts that

$$\|\mathbf{u}(-\sigma)\|_\infty \leq K^\xi \tau^{-1} \|\mathbf{u}(-\tau - \sigma)\|_1,$$

and (10.9) follows from (10.6). □

Proof of Theorem 9.8. To avoid confusion between the roles of σ, in this proof we denote the σ in the statement of Theorem 9.8 by σ'. We will apply Theorem 10.2 and we choose

σ and τ such that $\sigma' = \sigma + \tau$. We also shift the time by σ' so that the initial time is zero and the final time $\sigma + \tau = \sigma'$. Conditions $(\mathbf{C1})_\rho$ and $(\mathbf{C2})^*_\xi$ follow directly from (9.57)–(9.59) and from strong regularity at (Z, σ') but ϱ_1 and ξ' are replaced by ϱ and ξ for simplicity of notation. Given $\sigma' \in [K^{c_3}, C_1 K \log K]$, we consider two cases. If $\sigma' \le K^{1-c_3}$, we apply Theorem 10.2 with $\sigma = \tau = \sigma'/2$. Then $\|\mathbf{u}(-\tau - \sigma)\|_1$ becomes $\|\mathbf{v}^b\|_1 = 1$ on the right hand side of (10.9), and (9.60) follows. If $\sigma' \ge K^{1-c_3}$, then we apply Theorem 10.2 with $\sigma = \frac{1}{2}K^{1-c_3}$ and $\tau := \sigma' - \sigma$. In this case τ is comparable with σ' and $\sigma' \le \sigma^{3/2}$, and (9.60) again follows. □

The rest of this section is devoted to the proof of Theorem 10.1. Our strategy follows the approach of [13]; the multiscale iteration scheme and the key cutoff functions (10.20, 10.21) are also the same as in [13]. The main new feature of our argument is the derivation of the local energy estimate, Lemma 10.6, for a parabolic equation with singular coefficients satisfying $(\mathbf{C1})_\rho$ and $(\mathbf{C2})_\xi$. The proof of Lemma 10.6 will proceed in two steps. We first use condition $(\mathbf{C1})_\rho$ and the argument of the energy estimate in [13] to provide a bound in $L^\infty_t(L^2(\mathbb{Z}))$ on the solution to (10.1) (part (i) of Lemma 10.6). Along this proof we also prove an energy dissipation estimate which can be translated into the statement that the energy is small for most of the time. Using a new Sobolev type inequality (Proposition B.4) designed to deal with weak ellipticity we can improve the bound in $L^\infty_t(L^2(\mathbb{Z}))$ to an L^∞ estimate in space for most of the time to obtain part (ii) of Lemma 10.6. Finally, we run the argument again to improve the $L^\infty_t(L^2(\mathbb{Z}))$ estimate for short times (part (iii) of Lemma 10.6) that is needed to close the iteration scheme. Besides this proof, the derivation of the second De Giorgi estimate (Lemma 10.7) is also adjusted to the weaker condition $(\mathbf{C1})_\rho$.

We warn the reader that the notation for various constants in this section will follow [13] as much as possible for the sake of easy comparison. The conventions for these constants will differ from the ones in the previous sections, and, in particular, we will restate all conditions.

10.1. Hölder regularity

For any set S and any real function f define the oscillation $\mathrm{Osc}_S\, f := \sup_S f - \inf_S f$.

Theorem 10.3. *There exists a universal positive constant* $\mathfrak{q}$ *with the following property. For any two thresholds* $1 < \vartheta_1 < \vartheta_0$ *there exist two positive constants* ξ, ρ*, depending only on* ϑ_1 *and* ϑ_0*, such that the following hold:*

Set $\mathcal{M} := 2^{-\tau_0} K$ *where* $\tau_0 \in \mathbb{N}$ *is chosen such that* $\vartheta := \log K / \log \mathcal{M} \in [\vartheta_1, \vartheta_0]$*. Suppose that* (10.1) *satisfies* $(\mathbf{C1})_\rho$ *and* $(\mathbf{C2})_\xi$ *with some* $Z \in [\![-K/2, K/2]\!]$*. Suppose* $\mathbf{u}$ *is a solution to* (10.1) *in the time interval* $\mathcal{T} = [-3\mathcal{M}, 0]$*. Assume that*

$$\sup_{t\in[-3\mathcal{M},0]} \max_i |u_i(t)| \le \ell \tag{10.10}$$

for some ℓ*. Then for any* $\alpha \in [0, 1/3]$ *there is a set* $\mathcal{G} \subset [-\mathcal{M}^{1-\alpha}, 0]$ *such that*

$$\mathrm{Osc}_{Q^{(\alpha)*}}(u) \le 4\ell \mathcal{M}^{-\mathfrak{q}\alpha}, \qquad Q^{(\alpha)*} := \mathcal{G} \times [\![Z - 3\mathcal{M}^{1-\alpha}, Z + 3\mathcal{M}^{1-\alpha}]\!], \tag{10.11}$$

with

$$|[-\mathcal{M}^{1-\alpha},0]\setminus\mathcal{G}|\le\mathcal{M}^{1/4},$$

i.e. the oscillation of the solution on scale $\mathcal{M}^{1-\alpha}$ (*and away from the edges of the configuration space*) *is smaller than* $4\ell\mathcal{M}^{-\mathfrak{q}\alpha}$ *for most of the time. Moreover,*

$$\operatorname{Osc}_{\bar{Q}^{(\alpha)}}(u)\le C\ell\mathcal{M}^{-\mathfrak{q}\alpha},\qquad \bar{Q}^{(\alpha)}:=[-\mathcal{M}^{1/2},0]\times[\![Z-3\mathcal{M}^{1-\alpha},Z+3\mathcal{M}^{1-\alpha}]\!],\quad (10.12)$$

i.e. the oscillation is controlled for all times near 0.

These results hold for any $K\ge K_0$ *sufficiently large, where the threshold* K_0 *as well as the constant* C *in* (10.12) *depend only on the parameters* ϑ_0, ϑ_1.

We remark that the constant $\mathfrak{q}$ plays the role of the Hölder exponent and it depends only on ε_0 from Lemma 10.6. This will be explained after Lemma 10.8 below.

Proof of Theorem 10.1. With Theorem 10.3, we now complete the proof of Theorem 10.1. Given $\sigma\in[K^{c_3},K^{1-c_3}]$, define $\mathcal{M}:=2^{-\tau_0}K$ with some $\tau_0\in\mathbb{N}$ such that $\sigma/6\le\mathcal{M}\le\sigma/3$. Choosing $\vartheta_1:=1+\frac{1}{2}c_3$, $\vartheta_0:=2/c_3$, we clearly have $\vartheta=\log K/\log\mathcal{M}\in[\vartheta_1,\vartheta_0]$. Then (10.6) follows from (10.12) at time $t=0$ using $\sigma^{1-\alpha}\le 3\mathcal{M}^{1-\alpha}$. □

The proof of Theorem 10.3 will be a multiscale argument. On each scale $n=0,1,\dots,n_{\max}$ we define a space-time scale $M_n:=\nu^n\mathcal{M}$ and a size-scale $\ell_n:=\zeta^n\ell$ with some scaling parameters $\nu,\zeta<1$ to be chosen later. The initial scales are $M_0=\mathcal{M}$ and $\ell_0=\ell$. For notational convenience we assume that ν is of the form $\nu=2^{-j_0}$ for some integer $j_0>0$. We assume that $\nu\le\zeta^{10}/10$, and eventually ζ will be very close to 1, while ν will be very close to 0. The corresponding space-time box on scale n is given by

$$Q_n:=[-M_n,0]\times[Z-M_n,Z+M_n].$$

We will sometimes use an enlarged box

$$\widehat{Q}_n:=[-3M_n,0]\times[Z-\widehat{M}_n,Z+\widehat{M}_n],\qquad \widehat{M}_n:=LM_n,$$

with some large parameter L that will always be chosen such that $\nu\le 1/(2L)$ and thus $\widehat{Q}_n\subset Q_{n-1}$. We stress that the scaling parameters ν,ζ,L will be absolute constants, independent of any parameters in the setup of Theorem 10.3.

The smallest scale is given by the relation $M_{n_{\max}}\sim\mathcal{M}^{1-\alpha}$, i.e. $n_{\max}=\alpha\frac{\log\mathcal{M}}{|\log\nu|}$. In particular, since $\alpha\le 1/3$, all scales arising in the proofs will be between $\mathcal{M}^{2/3}$ and $\mathcal{M}$:

$$\mathcal{M}^{2/3}\le M_n\le\mathcal{M}=M_0,\qquad\forall n=0,1,\dots,n_{\max}.\quad (10.13)$$

The following statement is the main technical result that will immediately imply Theorem 10.3. In the application we will need only the second part of this technical theorem, but its formulation is tailored to its proof that will be an iterative argument from larger to smaller scales.

There will be several exponents in this theorem, but the really important one is χ: see explanation around (10.18) later. The exponents ξ and ρ can be chosen arbitrarily small and the reader can safely neglect them on a first reading.

Theorem 10.4 (Staircase estimate). *There exist positive parameters ν, ζ, L, satisfying*

$$\nu < \min\{\zeta^{10}/10, 1/(2L)\}$$

with the following property. For any two thresholds $1 < \vartheta_1 < \vartheta_0$ there exist positive constants χ, ξ, and ρ depending only on ϑ_1 and ϑ_0 (given explicitly in (10.75) and (10.76) later) such that under the setup and conditions of Theorem 10.3, for any scale $n = 0, 1, \ldots, n_{\max}$ there exists a descreasing sequence of sets $\mathcal{G}_n \subset [-3M_n, 0]$ of "good" times, $\mathcal{G}_n \subset \mathcal{G}_{n-1} \subset \ldots$, with

$$|\mathcal{G}_n^c| \le C \sum_{m=0}^{n-1} M_m^{1/4}, \qquad \mathcal{G}_n^c := [-3M_n, 0] \setminus \mathcal{G}_n, \tag{10.14}$$

such that we have the following two estimates:

(i) [Staircase estimate] *Define the constant $\bar{u}_n$ by*

$$\sup_{\widehat{Q}_n^*} |u - \bar{u}_n| = \tfrac{1}{2} \operatorname{Osc}_{\widehat{Q}_n^*}(u),$$

where

$$\widehat{Q}_n^* := ([-3M_n, 0] \cap \mathcal{G}_n) \times [Z - \widehat{M}_n, Z + \widehat{M}_n]$$

and for any $m < n$ define

$$S_{m,n} := \sum_{j=m}^{n-1} |\bar{u}_j - \bar{u}_{j+1}|.$$

Then

$$(\mathbf{ST})_n \qquad |u_i(t) - \bar{u}_n| \le \Psi_i^{(n)}(t) \quad \forall t \in [-3M_n, 0],\ \forall i, \tag{10.15}$$

where $\Psi^{(n)}$ is a function on $[-3M_n, 0] \times I$ defined by

$$\Psi_i^{(n)}(t) := \Lambda_i^{(n)} \cdot \mathbf{1}(t \in \mathcal{G}_n) + \Phi_i^{(n)}(t) \cdot \mathbf{1}(t \in \mathcal{G}_n^c)$$

with

$$\begin{aligned}\Lambda_i^{(n)} &:= \mathbf{1}(\widehat{M}_0 \le |i - Z|) \cdot \ell_0 + \sum_{m=0}^{n-1} \mathbf{1}(\widehat{M}_{m+1} \le |i - Z| \le \widehat{M}_m) \cdot [\ell_m + S_{m,n}] \\ &\quad + \mathbf{1}(|i - Z| \le \widehat{M}_n) \cdot \ell_n\end{aligned}$$

and

$$\begin{aligned}\Phi_i^{(n)}(t) &:= C_\Phi \cdot \mathbf{1}(\widehat{M}_0 \le |i - Z|) \cdot \ell_0 \\ &\quad + C_\Phi \sum_{m=0}^{n-1} \mathbf{1}(\widehat{M}_{m+1} \le |i - Z| \le \widehat{M}_m) \cdot \left[\ell_m \left(1 + \sqrt{\frac{|t| + \mathcal{M}^{1/2}}{M_m}} M_m^{\chi/2}\right) + S_{m,n}\right] \\ &\quad + C_\Phi \cdot \mathbf{1}(|i - Z| \le \widehat{M}_n) \cdot \ell_n \left(1 + \sqrt{\frac{|t| + \mathcal{M}^{1/2}}{M_n}} M_n^{\chi/2}\right)\end{aligned} \tag{10.16}$$

with some fixed constant C_Φ. The subscript Φ in C_Φ indicates that this specific constant controls the functions $\Phi^{(n)}$.

(ii) [Oscillation estimate] *For the good times we have*

$$(\mathbf{OSC})_n \qquad \tfrac{1}{2}\operatorname{Osc}_{\widehat{Q}^*_{n+1}}(u) \le \zeta\ell_n = \ell_{n+1}, \tag{10.17}$$

i.e. in the smaller box $\widehat{Q}^*_{n+1} \subset \widehat{Q}^*_n$ *the oscillation is reduced from* ℓ_n *to* ℓ_{n+1}.

All statements hold for any $K \ge K_0$ *sufficiently large, where the threshold* K_0 *as well as the constant* C_Φ *depend only on the universal constants* ν, ζ, L *and on the parameters* ϑ_0, ϑ_1, ξ, ρ.

Here the time independent profile $\Lambda^{(n)}$ is the "good" staircase function, representing the control for most of the time ("good times"). The function $i \mapsto \Lambda^{(n)}_i$ is a stepfunction that increases in $|i - Z|$ at a rate of approximately

$$\Lambda^{(n)}_i \sim \ell_n(|i-Z|/M_n)^{\mathfrak{q}}, \qquad |i-Z| \gg M_n,$$

where

$$\mathfrak{q} = |\log\zeta|/|\log\nu| \tag{10.18}$$

is a small positive exponent. Note that this exponent is the same as the final Hölder exponent in Theorems 10.3 and 10.1.

For the "bad times" (the complement of the good times), a larger control described by $\Phi^{(n)}(t)$ holds. This weaker control is time dependent and deteriorates with larger $|t|$. The exponent χ in the definition of Φ (see (10.16)), will be essentially equal to $\mathfrak{q}$ (modulo some upper cutoff, see (10.75) later). The factor $M_n^{\chi/2}$ on scale n expresses how much the estimate deteriorates for "bad times" compared with the estimate at "good times".

The bound (10.15) for good times $t \in \mathcal{G}_n$ with the control function $\Lambda^{(n)}$ directly follows from (10.17) and (10.10). The new information in (10.15) is the weaker estimate expressed by $\Phi^{(n)}$ that holds for all times. Note that $\Lambda^{(n)}_i \le \Phi^{(n)}_i(t)$, i.e. the bound

$$|u_i(t) - \bar{u}_n| \le \Phi^{(n)}_i(t), \qquad \forall t \in [-M_n, 0],\ \forall i,$$

follows from (10.15). We also remark that (10.17) implies $|\bar{u}_n - \bar{u}_{n+1}| \le \ell_n$, and thus

$$S_{m,n} = \sum_{j=m}^{n-1} |\bar{u}_j - \bar{u}_{j+1}| \le \sum_{j=m}^{n-1} \ell_j \le \frac{\ell_m}{1-\zeta} \tag{10.19}$$

gives an estimate for the effect $S_{m,n}$ of the shifts in the definition of $\Lambda^{(n)}$ and $\Phi^{(n)}$. Moreover, the uniform bound (10.10) shows that for any n,

$$|\bar{u}_n| \le \ell_0 = \ell.$$

Proof of Theorem 10.3. Without loss of generality we can assume that $\mathcal{M}^{-\alpha} \le \nu^2$, otherwise $\mathcal{M}^{-\mathfrak{q}\alpha} \ge \zeta^2 \ge 2/3$, so (10.11) immediately follows from (10.10). For $\mathcal{M}^{-\alpha} \le \nu^2$, the estimate (10.11) follow directlys from (10.17), by choosing $n \ge 1$ such that $M_{n+2} \le 3\mathcal{M}^{1-\alpha} \le M_{n+1}$, i.e. $\nu^{n+2} \le 3\mathcal{M}^{-\alpha} \le \nu^{n+1}$. Then $\ell_{n+1} = \ell\zeta^{n+1} \le 2\mathcal{M}^{-\mathfrak{q}\alpha}$ with $\mathfrak{q}$

defined in (10.18). The set $\mathcal{G}$ in Theorem 10.3 will be just $\mathcal{G}_{n+1} \cap [-\mathcal{M}^{1-\alpha}, 0]$. The estimate (10.12) follows from (10.15) by noting that for $|t| \le \mathcal{M}^{1/2} \le M_n^{3/4}$ (see (10.13)) the terms

$$\sqrt{\frac{|t| + \mathcal{M}^{1/2}}{M_m}} M_m^{\chi/2} \ll 1, \qquad m = 0, 1, \dots, n,$$

are all negligible and we simply have

$$\Phi_i^{(n)}(t) \le C_\Phi \Lambda_i^{(n)}, \qquad |t| \le \mathcal{M}^{1/2}.$$

Thus (10.12) follows exactly as (10.11). This completes the proof. □

In the rest of the section we will prove Theorem 10.4. We will iteratively check the main estimates, $(\mathbf{ST})_n$ and $(\mathbf{OSC})_n$, from scale to scale. For $n = 0$, the bound $(\mathbf{ST})_0$ is given by (10.10). In Section 10.2 we prove for any n that $(\mathbf{ST})_n$ implies $(\mathbf{OSC})_n$. In Section 10.3 we prove that $(\mathbf{ST})_n$ and $(\mathbf{OSC})_n$ imply $(\mathbf{ST})_{n+1}$. From these two statements it will follow that $(\mathbf{ST})_n$ and $(\mathbf{OSC})_n$ hold for any n. Sections 10.4 and 10.5 contain the proof of two independent results (Lemmas 10.6 and 10.7) formulated on a fixed scale, which are used in Section 10.2. These are the generalizations of the first and second De Giorgi lemmas of [13], adjusted to our situation where no supremum bound is available on the coefficients $B_{ij}(s)$, and we have control only in a certain average sense.

10.2. Proof of $(\mathbf{ST})_n \Rightarrow (\mathbf{OSC})_n$

For any real number a, we write $a_+ = \max(a, 0) \ge 0$ and $a_- = \min(a, 0) \le 0$, in particular $a = a_+ + a_-$. Fix a large integer M and a center $Z \in I$ with $d_Z \ge K/2$ (recall that d_i was defined above (10.3); it is the distance of i to the boundary). For any $\ell > 0$ and $\lambda \in (0, 1/10)$ define

$$\psi_i = \psi_i^{(M,Z,\ell)} := \ell \left(\left| \frac{i - Z}{M} \right|^{1/2} - 1 \right)_+, \tag{10.20}$$

$$\widetilde{\psi}_i = \widetilde{\psi}_i^{(M,Z,\ell,\lambda)} := \ell \left[\left(\left| \frac{i - Z}{M} \right| - \lambda^{-4} \right)_+^{1/4} - 1 \right]_+. \tag{10.21}$$

Notice that $\psi_i = 0$ if $|i - Z| \le M$ and $\widetilde{\psi}_i = 0$ if $|i - Z| \le M\lambda^{-4}$. Here ℓ will play the role of the typical size of $u - \psi$. One could scale out ℓ completely, but we keep it in. We also define the scaled versions of these functions for any $n \ge 0$:

$$\psi_i^{(n)} := \psi_i^{(M_n, Z, \ell_n)}, \qquad \widetilde{\psi}_i^{(n)} := \widetilde{\psi}_i^{(M_n, Z, \ell_n, \lambda)}.$$

Proposition 10.5. *Suppose that for some $n \ge 0$ we know $(\mathbf{ST})_j$ for any $j = 0, 1, \dots, n$. Then $(\mathbf{OSC})_n$ holds. Furthermore,*

$$\sum_i (u_i(t) - \bar{u}_n - \ell_n - \psi_i^{(n)})_+^2 \le C \left(\frac{|t| + \mathcal{M}^{1/2}}{M_n} \right) M_n^\chi \ell_n^2, \qquad t \in [-M_n, 0]. \tag{10.22}$$

Proof of Proposition 10.5. With a small constant $\lambda \in (10L^{-1/4}, 1)$ and a large integer k_0, to be specified later, define the rescaled and shifted functions

$$v_i^{(n,k)}(t) := \ell_n + \lambda^{-2k}([u_i(t) - \bar{u}_n] - \ell_n), \qquad k = 0, 1, \dots, k_0. \tag{10.23}$$

In particular, from $(\mathbf{ST})_n$ we have

$$v_i^{(n,k)}(t) \le \ell_n + \lambda^{-2k}(\Psi_i^{(n)}(t) - \ell_n), \qquad t \in [-3M_n, 0]. \tag{10.24}$$

We will show that with an appropriate choice $k = k(n)$, $v = v^{(n,k+1)}$ satisfies a better upper bound than (10.24), which then translates into a decrease in the oscillation of u on scale n. The improved upper bound on v will follow from applying two basic lemmas from parabolic regularity theory, traditionally called the first and second De Giorgi lemmas. The second De Giorgi lemma asserts that going from a larger to a smaller space-time regime, the maximum of $v_i(t)$ decreases in an average sense. The first De Giorgi lemma enhances this statement to a supremum bound for $v_i(t)$ that is strictly below the maximum of $v_i(t)$ on a larger space-time regime. This is equivalent to the reduction of the oscillation of v.

In the next section we first state these two basic lemmas, then we continue the proof of Proposition 10.5. The proofs of the De Giorgi lemmas are deferred to Sections 10.4 and 10.5.

10.2.1. Statement of the generalized De Giorgi lemmas. Both results will be formulated on a fixed space-time scale M and with a fixed size-scale ℓ. We fix a center $Z \in I$ with $|Z| \le K/2$. Recall the definition of $\psi = \psi^{(M,Z,\ell)}$ from (10.20). The first De Giorgi lemma is a local dissipation estimate:

Lemma 10.6. *There exists a small positive universal constant ε_0 with the following properties. Consider the parabolic equation (10.1) on the time interval $\mathcal{T} = [-\sigma, 0]$ with some $\sigma \in [K^{c_3}, K^{1-c_3}]$ and let $\mathbf{u}$ be a solution. Define $\mathbf{v} := \mathbf{u} - \bar{u}$ with some constant shift $\bar{u} \in \mathbb{R}$. Fix small positive constants κ, ξ, ρ, χ and a large constant ϑ_0 such that*

$$10\vartheta_0(\xi + \rho) \le \kappa \le 1/1000, \qquad \kappa + 10\vartheta_0(\xi + \rho) \le \chi \le 1/1000. \tag{10.25}$$

Let $M := K^{1/\vartheta}$ with some $\vartheta \in [1 + 2\kappa, \vartheta_0]$. Assume that the matrix elements of $\mathcal{A} = \mathcal{B} + \mathcal{W}$ satisfy (10.3)–(10.5) with exponent ξ and that (10.1) is regular with exponent ρ at the space-time points (Z, t), $t \in \Xi_0$, where

$$\Xi_0 := \{-M \cdot 2^{-m}(1 + 2^{-k}) : 0 \le m, k \le C \log M\}. \tag{10.26}$$

Assume

$$|\bar{u}| \le C\ell K^{1-\xi} M^{-1}, \tag{10.27}$$

$$\left[\frac{1}{M^2}\int_{-2M}^0 dt \sum_i (v_i(t) - \psi_i)_+^2\right]^{1/2} \le \varepsilon_0 \ell, \qquad \psi_i = \psi_i^{(M,Z,\ell)}, \tag{10.28}$$

$$\sup_{t\in[-2M,0]} \sup\{|i - Z| : v_i(t) > \psi_i\} \le M^{1+\kappa}, \tag{10.29}$$

$$\sup_{t\in[-2M,0]} \max\{v_i(t) : |i - Z| \le M^{1+\kappa}\} \le C\ell M^{\chi/2}, \tag{10.30}$$

and there exists a set $\mathcal{G}^* \subset [-2M, 0]$ *with* $|[-2M, 0] \setminus \mathcal{G}^*| \le CM^{1/4}$ *such that*

$$\sup_{t \in [-2M,0] \cap \mathcal{G}^*} \max\{v_i(t) : |i - Z| \le M^{1+\kappa}\} \le C\ell M^{\chi/10}. \tag{10.31}$$

Then, for any sufficiently large $K \ge K_0(\vartheta_0)$*, we have the following statements:*

(i) *We have*

$$\sup_{t \in [-M,0]} \sum_i (v_i(t) - \psi_i - \ell/3)_+^2 \le CM^\chi \ell^2. \tag{10.32}$$

(ii) *There exists a set* $\mathcal{G} \subset [-M, 0]$ *of "good" times such that*

$$\sup_{t \in \mathcal{G}} v_i(t) \le \ell/2 + \psi_i, \quad \forall i, \quad |[-M, 0] \setminus \mathcal{G}| \le CM^{1/4}. \tag{10.33}$$

(iii) *For any* $\widetilde{M}$ *with* $M^{2\chi} \ll \widetilde{M} \le \frac{1}{2}M$ *we have*

$$\sup_{t \in [-\widetilde{M},0]} \sum_i (v_i(t) - \psi_i - 2\ell/5)_+^2 \le C(\widetilde{M}/M)M^\chi \ell^2. \tag{10.34}$$

These results hold for any $K \ge K_0$*, where the threshold* K_0 *and the constants in* (10.32)–(10.34) *may depend on* $\chi, \kappa, \xi, \rho, \vartheta_0$ *and on the constants* C *and* $\widehat{C}$ *in* (10.3)–(10.5).

For the orientation of the reader we mention how the various exponents will be chosen in the application. The important exponents are κ and χ; they will be related by $\kappa = 3\chi/4$, in (10.69) later (actually, the really important relation is that $\kappa < \chi$). The exponents ξ, ρ will be chosen much smaller; the reader may neglect them on a first reading.

Notice that (10.32) is off from the optimal bound by a factor of M^χ. However, (10.33) shows that for most of the time, this factor is not present, while (10.34) shows that this factor is reduced if the time interval is shorter. We remark that precise coefficients of ℓ in the additive shifts appearing in (10.32)–(10.34) are not important; instead of $1/2 > 2/5 > 1/3$ essentially any three numbers between 0 and 1 with the same ordering could have been chosen.

The second De Giorgi lemma is a local descrease of oscillation on a single scale. As before, we are given three parameters, M, Z, ℓ. Define a new function F by

$$F_i = F_i^{(M,Z,\ell)} := \ell \cdot \max\left\{-1, \min\left(0, \left|\frac{i - Z}{M}\right|^2 - 81\right)\right\} \tag{10.35}$$

for any M, Z, ℓ. Notice that $-\ell \le F \le 0$, furthermore $F_i = 0$ if $|i - Z| \ge 9M$ and $F_i = -\ell$ if $|i - Z| \le 8M$. We also introduce a new parameter $\lambda \in (0, 1/10)$. Recalling the definition of $\widetilde{\psi}$ from (10.21), we also define three cutoffs, all depending on all four parameters, M, Z, ℓ, λ:

$$\varphi_i^{(0)} := \ell + \widetilde{\psi}_i + F_i,$$
$$\varphi_i^{(1)} := \ell + \widetilde{\psi}_i + \lambda F_i,$$
$$\varphi_i^{(2)} := \ell + \widetilde{\psi}_i + \lambda^2 F_i.$$

Notice that

$$\varphi_i^{(0)} \le \varphi_i^{(1)} \le \varphi_i^{(2)} \le \ell + \widetilde{\psi}_i, \tag{10.36}$$

and when $|i - Z| \ge 9M$ all inequalities become equalities. Notice that $\varphi_i^{(0)} = 0$ if $|i - Z| \le 8M$.

Lemma 10.7. *Consider the parabolic equation* (10.1) *on the time interval $\mathcal{T} = [-\sigma, 0]$ with some $\sigma \in [K^{c_3}, K^{1-c_3}]$ and let $\mathbf{u}$ be a solution. Define $\mathbf{v} := \mathbf{u} - \bar{u}$ with some constant shift $\bar{u} \in \mathbb{R}$. Fix small positive constants $\kappa_1, \kappa_2, \xi, \rho$ and a large constant ϑ_0 such that*

$$\kappa_1 + \kappa_2 + 10\vartheta_0(\xi + \rho) \le 1/1000, \tag{10.37}$$

Let $M := K^{1/\vartheta}$ with some $\vartheta \in [1 + 2\kappa_1, \vartheta_0]$. Assume that the matrix elements of $\mathcal{A} = \mathcal{B} + \mathcal{W}$ satisfy (10.3)–(10.5) *with exponent ξ and that* (10.1) *is regular with exponent ρ at (Z, t), $t \in \Xi_0$, where Ξ_0 was given in* (10.26).

For any $\delta, \mu > 0$ there exist $\gamma > 0$ and $\lambda \in (0, 1/8)$ such that whenever

$$|\bar{u}| \le C\lambda\ell K^{1-\xi} M^{-1}, \tag{10.38}$$

and the shifted solution $\mathbf{v}(t) = \mathbf{u}(t) - \bar{u}$ satisfies the following five properties;

$$\exists \mathcal{G} \subset [-3M, 0],\ |[-3M, 0] \setminus \mathcal{G}| \le CM^{1/4},\ \forall t \in \mathcal{G},\ \forall i, \qquad v_i(t) \le \ell + \widetilde{\psi}_i, \tag{10.39}$$

$$\sup_{t \in [-3M,0]} \max\{|i - Z| : v_i(t) > \ell + \widetilde{\psi}_i\} \le M^{1+\kappa_1},$$

$$\sup_{t \in [-3M,0]} \sup\{v_i(t) : |i - Z| \le M^{1+\kappa_1}\} \le \ell M^{\kappa_2}, \tag{10.40}$$

$$\frac{1}{M^2} \int_{-3M}^{-2M} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{|i - Z| \le M : v_i(t) < \varphi_i^{(0)}\}\, dt \ge \mu, \tag{10.41}$$

$$\frac{1}{M^2} \int_{-2M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i(t) > \varphi_i^{(2)}\}\, dt \ge \delta, \tag{10.42}$$

then

$$\frac{1}{M^2} \int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : \varphi_i^{(0)} < v_i(t) < \varphi_i^{(2)}\}\, dt \ge \gamma. \tag{10.43}$$

This conclusion holds for any $K \ge K_0$ where the threshold K_0 depends on all parameters ϑ_0, κ_1, κ_2, ξ, ρ, δ, μ and the constants in (10.3)–(10.5).

We remark that the choices of γ and λ are explicit: one may choose

$$\gamma := c\delta^3, \quad \lambda := c\delta^6\mu \tag{10.44}$$

with a small absolute constant c.

This lemma asserts that whenever a substantial part of the function v increases from $\varphi^{(0)}$ to $\varphi^{(2)}$ in time of order M, then there is a time interval of order M such that a substantial part of v lies between $\varphi^{(0)}$ and $\varphi^{(2)}$.

10.2.2. Verifying the assumptions of Lemma 10.7. We will apply Lemma 10.7 to the function $v = v^{(n,k)}$ given in (10.23) with the choice $M = M_n$, $\ell = \ell_n$. The following lemma collects the necessary information on $v = v^{(n,k)}$ to verify the assumptions in Lemma 10.7. The complicated relations among the parameters, listed in (10.45) and (10.46) below, can be simultaneously satisfied; their appropriate choice will be given in Section 10.2.4.

Lemma 10.8. *Assume that* $(\mathbf{ST})_n$ *holds* (*see* (10.15)). *Suppose that in addition to the previous relations* $\nu < \min\{\zeta^{10}/10, 1/(2L)\}$ *and* $\lambda \geq 10L^{-1/4}$ *among the parameters, the following further relations also hold:*

$$10 \leq (1-\zeta)\lambda^{2k_0}\zeta L^{1/4}, \quad \chi + 10\vartheta_0(\xi+\rho) \leq \frac{1}{1000}, \quad 100\vartheta_0(\xi+\rho) \leq \chi \leq \frac{|\log\zeta|}{|\log\nu|}, \tag{10.45}$$

$$\vartheta \in [1+2\chi, \vartheta_0], \qquad 1 - \tfrac{1}{2}\lambda^{2(k_0+1)} \leq \zeta < 1. \tag{10.46}$$

Then for any $v_i^{(n,k)}(t)$ *with* $k \leq k_0$, *defined in* (10.23) *and satisfying* (10.24), *we have*

$$\sup_{t\in\mathcal{G}_n}\sup_{k\leq k_0} v_i^{(n,k)}(t) \leq \ell_n + \widetilde{\psi}_i^{(n)}, \tag{10.47}$$

$$\sup_{k\leq k_0}\sup_{t\in[-3M_n,0]} \max\{|i-Z| : v_i^{(n,k)}(t) > \ell_n + \widetilde{\psi}_i^{(n)}\} \leq M_n^{1+3\chi/4}, \tag{10.48}$$

$$\sup_{k\leq k_0}\sup_{t\in[-3M_n,0]} \sup\{v_i^{(n,k)}(t) : |i-Z| \leq M_n^{1+3\chi/4}\} \leq C\ell_n M_n^{\chi/2}. \tag{10.49}$$

For the shift in (10.23) *we have*

$$|\ell_n - \lambda^{-2k}(\bar{u}_n + \ell_n)| \leq C\lambda\ell_n K^{1-\xi} M_n^{-1}. \tag{10.50}$$

The constants C may depend on all parameters in (10.45), (10.46).

We remark that the factor $3/4$ in the exponent in (10.48) can be improved to $2/3 + \varepsilon'$ for any $\varepsilon' > 0$, but what is really important for the proof is that it is *strictly smaller than* 1, since this will translate into the crucial $\kappa < \chi$ condition in (10.30).

Proof of Lemma 10.8. All four estimates follow by direct calculations from the definition of $\Psi^{(n)}(t)$ and from the relations (10.45), (10.46) among the parameters. Based upon (10.24), the estimate (10.47) amounts to checking

$$\Lambda_i^{(n)} \leq \ell_n + \lambda^{2k_0}\ell_n\left[\left(\left|\frac{i-Z}{M_n}\right| - \lambda^{-4}\right)_+^{1/4} - 1\right]_+. \tag{10.51}$$

For $|i-Z| \leq \widehat{M}_n$ we immediately have $\Lambda_i^{(n)} = \ell_n$ and thus (10.51) holds. For $\widehat{M}_{m+1} \leq |i-Z| \leq \widehat{M}_m$ (with some $m \leq n-1$) we can use (10.19) to find that $\Lambda_i^{(n)} \leq 2(1-\zeta)^{-1}\ell_m$. The right hand side of (10.51) is larger than

$$\ell_n + \lambda^{2k_0}\ell_n\left[\left(\left|\frac{\widehat{M}_{m+1}}{M_n}\right| - \lambda^{-4}\right)_+^{1/4} - 1\right]_+.$$

which is larger than $\ell_n(1+\frac{1}{2}\lambda^{2k_0}L^{1/4}\nu^{(m-n)/4})$. Now (10.51) follows from the first inequality in (10.45) and from $\nu \le \zeta^{10}/10$.

For the proof of (10.48), starting from (10.24), it is sufficient to check that

$$\Phi_i^{(n)}(t) \le \ell_n + \lambda^{2k_0}\ell_n\left[\left(\left|\frac{i-Z}{M_n}\right| - \lambda^{-4}\right)_+^{1/4} - 1\right]_+ \tag{10.52}$$

for any $|i-Z| \ge \frac{1}{2}M_n^{1+3\chi/4}$ and $t \in [-3M_n, 0]$. On the left hand side we can use the largest time $|t| = 3M_n \ge \mathcal{M}^{1/2}$. Considering the regime $\widehat{M}_m \le |i-Z| \le \widehat{M}_{m-1}$ with $M_m = M_n^{1+\beta}$ for some $0<\beta<1/2$, we see that

$$\text{l.h.s. of (10.52)} \le 2\ell_m\left(\frac{M_n}{M_m}\right)^{1/2} M_m^{\chi/2}, \qquad \text{r.h.s. of (10.52)} \ge \frac{1}{2}\lambda^{2k_0}\ell_n\left(\frac{M_m}{M_n}\right)^{1/4},$$

Using $\chi \le |\log\zeta|/|\log\nu|$ from (10.45), we have

$$\ell_m/\ell_n \le (M_m/M_n)^{\chi},$$

so (10.52) holds if

$$(M_n/M_m)^{1/2-\chi}M_m^{\chi/2} \le \tfrac{1}{4}\lambda^{2k_0}(M_m/M_n)^{1/4}. \tag{10.53}$$

Recalling that $M_m = M_n^{1+\beta}$, we see that for small χ, (10.53) is satisfied if $\beta > 2\chi/(3-6\chi)$ (and M_n is sufficiently large depending on all constants λ, ν, L, k_0, ν, ζ). This is guaranteed if $\beta \ge 3\chi/4$ since we have assumed $\chi \le 1/1000$. This proves (10.48).

For the proof of (10.49) we notice that

$$\max\{\Phi_i^{(n)}(t) : |i-Z| \le M_n^{1+3\chi/4}\} \le C_\Phi(\ell_m + M_n^{\chi/2}\ell_n) \le CM_n^{\chi/2}\ell_n \tag{10.54}$$

for any $t \in [-3M_n, 0]$, where $m<n$ is defined by $\widehat{M}_{m+1} \le M_n^{1+3\chi/4} \le \widehat{M}_m$. The first inequality in (10.54) follows from (10.16); the second one is a consequence of

$$\ell_m/\ell_n = (M_m/M_n)^{|\log\zeta|/|\log\nu|} \le (M_m/M_n)^{1/10} \le M_n^{\chi/10} \tag{10.55}$$

since $|\log\zeta| \le \frac{1}{10}|\log\nu|$. Then (10.49) directly follows from (10.24) and (10.54).

Finally, (10.50) follows from $|\bar{u}_n| \le \ell = \ell_0$, $K^{1-\xi} \ge \mathcal{M} = M_0$ (using $\vartheta \ge 1+2\xi$) and $\ell_0/\ell_n \le M_0/M_n$. This completes the proof of Lemma 10.8. □

10.2.3. Completing the proof of Proposition 10.5. We now continue the proof of Proposition 10.5. Set

$$F_i^{(n)} := F_i^{(M_n, Z, \ell_n)},$$

where F is given in (10.35), and define further cutoff functions:

$$\varphi_i^{(0),(n)} := \ell_n + \widetilde{\psi}_i^{(n)} + F_i^{(n)},$$
$$\varphi_i^{(1),(n)} := \ell_n + \widetilde{\psi}_i^{(n)} + \lambda F_i^{(n)},$$
$$\varphi_i^{(2),(n)} := \ell_n + \widetilde{\psi}_i^{(n)} + \lambda^2 F_i^{(n)}.$$

Throughout this section, n is fixed, so we will often omit it from the notation. In particular $\ell = \ell_n$, $M = M_n$, $\bar{u} = \bar{u}^{(n)}$, $v^{(k)} = v^{(n,k)}$, $F = F^{(n)}$, $\widetilde{\psi} = \widetilde{\psi}^{(n)}$, $\varphi_i^{(a)} = \varphi_i^{(a),(n)}$ for $a = 0, 1, 2$, $\mathcal{G} = \mathcal{G}_n$ etc. At the end of the proof we will add back the superscripts.

From the definitions of these cutoff functions, we have

$$\varphi_i^{(0)} \leq \varphi_i^{(1)} \leq \varphi_i^{(2)} \leq \ell + \widetilde{\psi}_i, \tag{10.56}$$

and when $|i - Z| \geq 9M$ all inequalities become equalities. Notice that $\varphi_i^{(0)} = 0$ if $|i - Z| \leq 8M$.

Choose a small constant $\mu \in (0, 1/10)$, say

$$\mu := 1/100. \tag{10.57}$$

Without loss of generality, we can assume

$$\frac{1}{M^2} \int_{-3M}^{-2M} \#\{i : |i - Z| \leq M,\ u_i(t) - \bar{u} < \varphi_i^{(0)}\}\, dt \geq \mu \tag{10.58}$$

(otherwise we can take $-u$; note that the condition in Proposition 10.5 is invariant under the $u \to -u$ sign flip).

Notice that for any $|i - Z| \leq M$ and $t \in \mathcal{G}$ the sequence $v_i^{(k)}(t)$ is decreasing in k, in particular $v_i^{(k)}(t) \leq \ell$. This follows from (10.24) and $\Psi_i^{(n)}(t) \leq \ell_n$ in this regime. From (10.58) we therefore have

$$\frac{1}{M^2} \int_{-3M}^{-2M} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : |i - Z| \leq M,\ v_i^{(k)}(t) < \varphi_i^{(0)}\}\, dt \geq \mu, \tag{10.59}$$

since the set of i indices in (10.59) is increasing in k for any $t \in \mathcal{G}$ and $v^{(0)} = u - \bar{u}$.

Assuming that the parameters satisfy (10.45) and (10.46), we can now use the conclusions (10.47)–(10.50) of Lemma 10.8. These bounds together with (10.59) allow us to apply Lemma 10.7 to $v^{(k)} = v^{(n,k)}$ with the choice

$$\kappa_1 := 3\chi/4, \qquad \kappa_2 := \chi/2, \qquad \delta := \varepsilon_0^2/100, \tag{10.60}$$

where $\varepsilon_0 > 0$ is a universal constant which was determined in Lemma 10.6. Notice that with these choices (10.37) follows from (10.45) and $\vartheta \in [1 + 2\kappa_1, \vartheta_0]$ follows from $\vartheta \in [1 + 2\chi, \vartheta_0]$. Thus the application of Lemma 10.7 shows that there exist a λ (introduced explicitly in the construction of the cutoffs $\varphi^{(a)}$ and used also in (10.21) and (10.24)) and a $\gamma > 0$ (see (10.44) for their explicit values) such that if

$$\frac{1}{M^2}\int_{-2M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(2)}\}\, dt > \delta \tag{10.61}$$

then

$$\frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : \varphi_i^{(0)} < v_i^{(k)}(t) < \varphi_i^{(2)}\}\, dt \geq \gamma.$$

Therefore

$$\begin{aligned} &\frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(2)}\}\, dt \\ &\qquad \leq \frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(0)}\}\, dt - \gamma. \end{aligned} \tag{10.62}$$

Notice that, by (10.47) and $F_i = 0$ if $|i - Z| \geq 9M$, for any $k \leq k_0$ the inequality $v_i^{(k)}(t) > \varphi_i^{(0)}$ (for $t \in \mathcal{G}$) can hold only if $|i - Z| \leq 9M$. Assuming $|i - Z| \leq 9M$, $t \in \mathcal{G}$ and $v_i^{(k)}(t) > \varphi_i^{(0)}$, we have

$$\frac{1}{\lambda^2}(v_i^{(k-1)}(t) - \ell) + \ell = v_i^{(k)}(t) > \varphi_i^{(0)}.$$

Since $|i - Z| \leq 9M \leq \lambda^{-4}M$, we have, together with (10.56) and $\widetilde{\psi}_i = 0$ in this regime,

$$v_i^{(k-1)}(t) \geq \lambda^2(\widetilde{\psi}_i + F_i) + \ell \geq \varphi_i^{(2)}.$$

Therefore, we can bound the last integral in (10.62) by

$$\begin{aligned} &\frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(0)}\}\, dt \\ &\qquad \leq \frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : |i - Z| \leq 9M,\ v_i^{(k-1)}(t) > \varphi_i^{(2)}\}\, dt. \end{aligned} \tag{10.63}$$

We have thus proved that

$$\begin{aligned} &\frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(2)}\}\, dt \\ &\qquad \leq \frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : |i - Z| \leq 9M,\ v_i^{(k-1)}(t) > \varphi_i^{(2)}\}\, dt - \gamma. \end{aligned} \tag{10.64}$$

Iterating this estimate k times, we get

$$\begin{aligned} &\frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(2)}\}\, dt \\ &\qquad \leq \frac{1}{M^2}\int_{-3M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : |i - Z| \leq 9M,\ v_i^{(0)}(t) > \varphi_i^{(2)}\}\, dt - k\gamma, \end{aligned}$$

which becomes negative if $k\gamma \geq 100$. If we set

$$k_0 := 100/\gamma, \tag{10.65}$$

then there is a $k < k_0$ such that (10.61) is violated, i.e.,

$$\frac{1}{M^2}\int_{-2M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(2)}\}\, dt \leq \delta. \tag{10.66}$$

From now on let $k = k(n)$ denote the smallest index such that (10.66) holds (recall that the underlying n dependence was omitted from the notation in most of this section). Furthermore, since $\varphi_i^{(0)} = 0$ for $|i - Z| \leq 8M$, we have

$$\begin{aligned}
&\frac{1}{M^2}\int_{-2M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : |i - Z| \leq 8M,\ v_i^{(k+1)}(t) > 0\}\, dt \\
&\qquad = \frac{1}{M^2}\int_{-2M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : |i - Z| \leq 8M,\ v_i^{(k+1)}(t) > \varphi_i^{(0)}\}\, dt \\
&\qquad \leq \frac{1}{M^2}\int_{-2M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \#\{i : v_i^{(k)}(t) > \varphi_i^{(2)}\}\, dt \leq \delta = \varepsilon_0^2/100,
\end{aligned} \tag{10.67}$$

where we have used (10.63) in the last inequality.

Armed with (10.67), our goal is to apply Lemma 10.6 with $M = M_n$ to $v = v^{(n,k(n)+1)}$ with the value $k = k(n)$ determined after (10.66). Clearly v is of the form

$$v = \lambda^{-2k-2}u + [\ell_n - \lambda^{-2k-2}(\bar{u}_n + \ell_n)], \tag{10.68}$$

i.e. it is a solution to (10.1) (namely $\lambda^{-2k-2}u$) shifted by $[\lambda_n - \lambda^{-2k-2}(\bar{u}_n + \ell_n)]$. The value κ in Lemma 10.6 will be set to

$$\kappa := 3\chi/4 \tag{10.69}$$

and the set $\mathcal{G}^*$ in Lemma 10.6 will be chosen as $\mathcal{G}^* := \mathcal{G}_n$ (for $n = 0$ we set $\mathcal{G}^* = [-3M_0, 0]$, i.e. at the zeroth step of the iteration every time is "good"; see (10.10)). The choice $\kappa = 3\chi/4$ together with the constraints on χ in (10.45) guarantee that the relations in (10.25) hold. We need to check the five conditions (10.27)–(10.31). The sixth condition, the regularity at (Z, t) for $t \in \Xi_0$, follows automatically from $(\mathbf{C1})_\rho$ since $M = M_n = \mathcal{M}\nu^n = 2^{-\tau_0}\nu^n K$ with an integer τ_0, and ν itself is a negative power of 2, thus $\Xi_0 \subset \Xi$ (see (9.6)). The first condition (10.27) for the shift in (10.68) was verified in (10.50).

For the second condition (10.28), with the notation $\mathcal{G}^c := [-3M, 0] \setminus \mathcal{G}$ we write

$$\begin{aligned}
&\frac{1}{M^2}\int_{-2M}^{0} \sum_i (v_i^{(k+1)}(t) - \psi_i)_+^2\, dt \\
&\leq \frac{1}{M^2}\int_{-2M}^{0} \mathbf{1}(t \in \mathcal{G}) \cdot \sum_i (v_i^{(k+1)}(t) - \psi_i)_+^2\, dt + \frac{|\mathcal{G}^c|}{M^2} \sup_{t \in [-2M,0]} \sum_i (v_i^{(k+1)}(t) - \psi_i)_+^2 .
\end{aligned} \tag{10.70}$$

In the first term we use

$$v_i^{(n,k+1)}(t) \le \ell_n + \widetilde{\psi}_i^{(n)}, \qquad t \in \mathcal{G}_n,$$

from (10.47) (we reintroduced the superscript n). Since $\ell_n + \widetilde{\psi}_i^{(n)} \le \psi_i^{(n)}$ if $|i-Z| \ge 8M$, we see that the summation in the first term on the right hand side of (10.70) is restricted to $|i-Z| \le 8M$, and for these i's we have $v_i^{(n,k+1)}(t) \le \ell_n$ since $\widetilde{\psi}_i^{(n)} = 0$. We can therefore apply (10.67) to get

$$\frac{1}{M_n^2}\int_{-2M_n}^0 \sum_i (v_i^{(n,k+1)}(t) - \psi_i^{(n)})_+^2\, dt \le 4\delta\ell_n^2 + \frac{|\mathcal{G}_n^c|}{M_n^2} \sup_{t\in[-2M_n,0]} \sum_i (v_i^{(n,k+1)}(t) - \psi_i^{(n)})_+^2. \tag{10.71}$$

To estimate the second term, we use (10.24) and $\Psi^{(n)} \le \Phi^{(n)}$ to note that

$$\psi_i^{(n)} \le v_i^{(n,k+1)}(t) \;\Rightarrow\; \psi_i^{(n)} \le \ell_n + \lambda^{-2k-2}(\Phi_i^{(n)}(t) - \ell_n), \qquad t \in [-3M_n, 0]. \tag{10.72}$$

Suppose first that $|i-Z| \ge M_n^{1+3\chi/4}$. In this case (10.52) holds, thus (10.72) would imply

$$\psi_n^{(n)} = \ell_n\left(\left|\frac{i-Z}{M_n}\right|^{1/2} - 1\right)_+ \le \lambda^{-2k_0}\ell_n + \ell_n\left[\left(\left|\frac{i-Z}{M_n}\right| - \lambda^{-4}\right)_+^{1/4} - 1\right]_+;$$

but this is impossible for $|i-Z| \ge M_n^{1+3\chi/4}$ if M_n is large enough. In particular, this verifies (10.29). We therefore conclude that the summation in the second term on the right hand side of (10.71) is restricted to $|i-Z| \le M^{1+3\chi/4}$. For these values we have

$$v_i^{(n,k+1)}(t) \le \ell_n + \lambda^{-2k}(\Phi_i^{(n)}(t) - \ell_n) \le C\lambda^{-2k}\ell_n M_n^{\chi/2}$$

(the first inequality is from (10.24), the second from (10.54)). This verifies (10.30), recalling the choice of $\kappa = 3\chi/4$.

Inserting this information into (10.71), we have

$$4\delta\ell_n^2 + \frac{|\mathcal{G}_n^c|}{M_n^2} \sup_{t\in[-2M_n,0]} \sum_i (v_i^{(n,k+1)}(t) - \psi_i^{(n)})_+^2 \le 4\delta\ell_n^2 + C\lambda^{-2k}M_n^{-1/2+2\chi}\ell_n^2 \le \varepsilon_0^2\ell_n^2,$$

where we have used $|\mathcal{G}_n^c| \le CM_0^{1/4} \le M_n^{1/2}$ from (10.14) and (10.13). In the last step we have used the choice $\delta = \varepsilon_0^2/100$. This verifies (10.28).

Finally, we verify (10.31) with the previously mentioned choice $\mathcal{G}^* := \mathcal{G}_n$. Let i be such that $|i-Z| \le M_n^{1+3\chi/4}$ and $t \in \mathcal{G}_n$. Then from (10.24) we have

$$v_i^{(n,k+1)}(t) = \lambda^{-2k-2}\Lambda_i^{(n)} + \ell_n(1-\lambda^{-2k-2}) \le C\ell_m \le CM_n^{\chi/10}\ell_n$$

where m is chosen such that $\widehat{M}_{m+1} \le M_n^{1+3\chi/4} \le \widehat{M}_m$ and in the last step we have used (10.55).

Thus we can apply Lemma 10.6 to $v = v^{(n,k+1)}$ and from (10.33) we get the existence of a set of times, denoted by $\mathcal{G}_n' \subset [-M_n, 0]$, such that

$$\sup_{t\in\mathcal{G}_n'} v_i^{(n,k+1)}(t) \le \ell_n/2, \qquad |i-Z| \le M_n, \qquad |[-M_n,0]\setminus\mathcal{G}_n'| \le CM_n^{1/4}.$$

Defining $\mathcal{G}_{n+1} := \mathcal{G}_n \cap \mathcal{G}_n' \cap [-3M_{n+1}, 0]$ and using $\widehat{M}_{n+1} \le M_n$, we obtain

$$\sup_{t\in\mathcal{G}_{n+1}} v_i^{(n,k+1)}(t) \le \ell_n/2, \qquad |i-Z| \le \widehat{M}_{n+1}, \tag{10.73}$$

and

$$\left|\mathcal{G}_{n+1}^c\right| \le CM_n^{1/4} + |\mathcal{G}_n^c| \le C\sum_{m=0}^{n} M_m^{1/4},$$

where we have used the measure of $\mathcal{G}_n^c$ from (10.14).

Recalling the definition (10.23), from (10.73) we have

$$u_i(t) - \bar{u}_n \le \ell_n\big(1 - \tfrac{1}{2}\lambda^{2(k+1)}\big) \le \ell_n\big(1 - \tfrac{1}{2}\lambda^{2(k_0+1)}\big) \le \ell_n\zeta = \ell_{n+1}, \qquad |i-Z| \le \widehat{M}_{n+1} \text{ and } t \in \mathcal{G}_{n+1},$$

where we recall that $k \le k_0$ and (10.46). Repeating the argument for $-u$ instead of u, we obtain a similar lower bound on $u_i(t) - \bar{u}_n$. This proves (10.17) for n, i.e. $(\mathbf{OSC})_n$.

The application of Lemma 10.6 also shows (see (10.34)) that for $t \in [-M_n, 0]$ we have

$$\sum_i \big(v_i^{(n,k+1)}(t) - \big(\tfrac{2}{5}\ell_n + \psi_i^{(n)}\big)\big)_+^2 \le C\left(\frac{|t| + \mathcal{M}^{1/2}}{M_n}\right) M_n^{\chi}\ell_n^2, \tag{10.74}$$

which implies, by (10.23) and elementary algebra, the second statement in Proposition 10.5 (the constant C in (10.22) includes a factor of $\lambda^{-2k} \le \lambda^{-2k_0}$).

10.2.4. Summary of the choice of the parameters. Finally, we present a possible choice of the parameters that were used in the proof of Proposition 10.5. Especially, we need to satisfy the complicated relations (10.45), (10.46).

Lemma 10.6 gives an absolute constant ε_0. Then we choose $\delta = \varepsilon_0^2/100$, $\gamma = c\delta^3$, $\lambda = c\delta^6\mu$ (with a small constant c), $k_0 = 100/\gamma$ and $\mu = 1/100$. These choices can be found in (10.60), (10.44), (10.65) and (10.57), respectively.

Having λ, k_0 determined, we define

$$\zeta := 1 - \tfrac{1}{2}\lambda^{2(k_0+1)}, \qquad L := \lambda^{-16(k_0+1)}, \qquad \nu := 2\lambda^{16(k_0+1)}.$$

If needed, reduce λ so that $|\log\zeta|/|\log\nu| \le 1/10$. Note that the five numbers $\lambda, k_0, \zeta, \nu, L$ are absolute positive constants (meaning that they do not depend on any input parameters in Theorem 10.3). In particular, they determine the absolute constant $\mathfrak{q}$ in (10.18), which is the final Hölder exponent.

Next we set

$$\chi := \min\left\{ \frac{|\log \zeta|}{|\log \nu|}, \frac{\vartheta_1 - 1}{2}, \frac{1}{2000} \right\} \tag{10.75}$$

and then choose the exponents ξ, ρ as

$$\xi := \rho := \frac{\chi}{200\vartheta_0}. \tag{10.76}$$

Finally, $\mathcal{M} = M_0$ (or, equivalently K_0) has to be sufficiently large depending on all these exponents.

It is easy to check that this choice of the parameters satisfies all the relations that were used in the proof of Proposition 10.5. This completes the proof of that proposition. □

10.3. Proof of $(\mathbf{ST})_n + (\mathbf{OSC})_n \Rightarrow (\mathbf{ST})_{n+1}$

Proposition 10.9. *Suppose that* $(\mathbf{ST})_n$ *and* $(\mathbf{OSC})_n$ *hold for some n. Then* $(\mathbf{ST})_{n+1}$ *also holds.*

Proof. For $t \in \mathcal{G}_{n+1} \subset \mathcal{G}_n$ we have

$$|u_i(t) - \bar{u}_{n+1}| \le \ell_{n+1}, \quad |i - Z| \le \widehat{M}_{n+1},$$

by $(\mathbf{OSC})_n$. Thus we immediately get $|u_i(t) - \bar{u}_{n+1}| \le \Lambda_i^{(n+1)}$ for $|i - Z| \le \widehat{M}_{n+1}$. For $\widehat{M}_{n+1} \le |i - Z| \le \widehat{M}_1$ we just use

$$|u_i(t) - \bar{u}_{n+1}| \le |u_i(t) - \bar{u}_n| + |\bar{u}_{n+1} - \bar{u}_n| \le \Lambda_i^{(n)} + |\bar{u}_{n+1} - \bar{u}_n| \le \Lambda_i^{(n+1)},$$

where the last estimate is from the definition of Λ. For $|i - Z| \ge \widehat{M}_1$ we have the trivial bound ℓ_0.

Now we need to check the case $t \in [-3M_{n+1}, 0] \setminus \mathcal{G}_{n+1}$. For $\widehat{M}_{n+1} \le |i - Z| \le \widehat{M}_1$, from $(\mathbf{ST})_n$ we have

$$|u_i(t) - \bar{u}_{n+1}| \le |u_i(t) - \bar{u}_n| + |\bar{u}_{n+1} - \bar{u}_n| \le \Phi_i^{(n)}(t) + |\bar{u}_{n+1} - \bar{u}_n| \le \Phi_i^{(n+1)}(t),$$

where the last inequality is just from the definition of Φ. Finally, if $|i - Z| \le \widehat{M}_{n+1}$ $(\le M_n)$, we use (10.22),

$$|u_i(t) - \bar{u}_{n+1}| \le |u_i(t) - \bar{u}_n| + |\bar{u}_n - \bar{u}_{n+1}| \le 2\ell_n + \ell_n \sqrt{C \frac{|t| + \mathcal{M}^{1/2}}{M_n}} M_n^{\chi/2},$$

since in this regime $\psi_i^{(n)} = 0$. The constant C is from (10.22). The right hand side is bounded by

$$C_\Phi \ell_{n+1} \left(1 + \sqrt{\frac{|t| + \mathcal{M}^{1/2}}{M_{n+1}}} M_{n+1}^{\chi/2} \right),$$

by using $\ell_n/\ell_{n+1} = \zeta^{-1} \le \nu^{-1/10} = (M_n/M_{n+1})^{1/10}$ and choosing C_Φ large enough. This completes the proof of Proposition 10.9. □

10.4. Proof of Lemma 10.6 (first De Giorgi lemma)

Assume for notational simplicity that $Z = 0$ and set

$$\psi_i^\ell := \psi_i + \ell.$$

Since $\mathbf{v}$ solves the equation

$$\partial_s v_i(s) = -[\mathcal{A}(s)\mathbf{v}(s)]_i - W_i(s)\bar{u}, \tag{10.77}$$

by direct computation we have

$$\partial_t \frac{1}{2}\sum_i [v_i - \psi_i^\ell]_+^2 = -\sum_{ij}(v_i - \psi_i^\ell)_+ B_{ij}(v_i - v_j) - \sum_i (v_i - \psi_i^\ell)_+ W_i (v_i + \bar{u}). \tag{10.78}$$

Recall that B_{ij} depends on time, but we will omit this from the notation. Since $W_i \geq 0$, the last term can be bounded by

$$\begin{aligned} -\sum_i (v_i - \psi_i^\ell)_+ W_i(v_i + \bar{u}) &\leq -\sum_i (v_i - \psi_i^\ell)_+ W_i (v_i - \psi_i^\ell)_+ - \bar{u}\sum_i (v_i - \psi_i^\ell)_+ W_i \\ &\leq -\mathfrak{w}[(v - \psi^\ell)_+, (v - \psi^\ell)_+] + |\bar{u}| \sum_i (v_i - \psi_i^\ell)_+ W_i. \end{aligned}$$

In the first term on the right hand side of (10.78) we can symmetrize and then add and subtract ψ^ℓ to get

$$\begin{aligned} -\sum_{ij}(v_i - \psi_i^\ell)_+ B_{ij}(v_i - v_j) &= -\mathfrak{b}[(v - \psi^\ell)_+, v] \\ &= -\mathfrak{b}[(v - \psi^\ell)_+, (v - \psi^\ell)_+] - \mathfrak{b}[(v - \psi^\ell)_+, (v - \psi^\ell)_-] - \mathfrak{b}[(v - \psi^\ell)_+, \psi^\ell]. \end{aligned}$$

Since $B_{ij} \geq 0$ and $[a_+ - b_+][a_- - b_-] \geq 0$ for any real numbers a, b, for the cross-term we have $\mathfrak{b}[(v - \psi^\ell)_+, (v - \psi^\ell)_-] \geq 0$. Thus the last equation is

$$\leq -\mathfrak{b}[(v - \psi^\ell)_+, (v - \psi^\ell)_+] - \mathfrak{b}[(v - \psi^\ell)_+, \psi^\ell]. \tag{10.79}$$

Using the definition of $\mathfrak{a}$ in (9.24), we have thus proved that

$$\partial_t \frac{1}{2}\sum_i [v_i - \psi_i^\ell]_+^2 \leq -\mathfrak{a}[(v - \psi^\ell)_+, (v - \psi^\ell)_+] - \mathfrak{b}[(v - \psi^\ell)_+, \psi^\ell] + |\bar{u}| \sum_i (v_i - \psi_i^\ell)_+ W_i. \tag{10.80}$$

Decompose the first error term as $\mathfrak{b}[(v - \psi^\ell)_+, \psi^\ell] = \Omega_1 + \Omega_2 + \Omega_3$, where

$$\Omega_1 := \frac{1}{2}\sum_{|i-j| \geq M} B_{ij}[\psi_i^\ell - \psi_j^\ell]((v_i - \psi_i^\ell)_+ - (v_j - \psi_j^\ell)_+) \cdot \mathbf{1}(\max\{d_i^I, d_j^I\} \geq K/3)$$

and Ω_2 and Ω_3 are defined in the same way except that the summation is restricted to $\widehat{C}K^\xi \leq |i - j| \leq M$ for Ω_2 and $|i - j| \leq \widehat{C}K^\xi$ for Ω_3, where $\widehat{C}$ is the constant from (10.5). Notice that we have inserted the characteristic function $\mathbf{1}(\max\{d_i^I, d_j^I\} \geq K/3)$

for free, since (10.29) together with $|Z| \le K/2$ and $M^{1+\kappa} \ll K$ (from $\vartheta \ge 1 + 2\kappa$) guarantees that $(v_i - \psi_i^\ell)_+ = 0$ unless $d_i^I \ge K/3$. Thus the summation over i, j can be restricted to index pairs where at least one index is far away from the boundary. Recall from (10.5) that in the regime $|i-j| \ge M$ we have $B_{ij} \le C|i-j|^{-2}$ since $M \ge \widehat{C}K^\xi$. Moreover,

$$|\psi_i^\ell - \psi_j^\ell| \le \ell M^{-1/2}|i-j|^{1/2}. \tag{10.81}$$

Altogether we have

$$|\Omega_1| \le \ell M^{-1/2} \sum_{|i-j|\ge M} \frac{1}{|i-j|^{3/2}}[(v_i - \psi_i^\ell)_+ + (v_j - \psi_j^\ell)_+] \le \frac{\ell}{M}\sum_i (v_i - \psi_i^\ell)_+.$$

For Ω_2, by symmetry of B_{ij}, we can write

$$\begin{aligned}
-\Omega_2 := &-\sum_{\widehat{C}K^\xi \le |i-j| \le M, \psi_i^\ell \le \psi_j^\ell} B_{ij}[\psi_i^\ell - \psi_j^\ell]((v_i - \psi_i^\ell)_+ - (v_j - \psi_j^\ell)_+) \\
\le &-\sum_{\widehat{C}K^\xi \le |i-j| \le M, \psi_i^\ell \le \psi_j^\ell} B_{ij}[\psi_i^\ell - \psi_j^\ell][(v_i - \psi_i^\ell)_+ - (v_j - \psi_j^\ell)_+] \cdot \mathbf{1}(v_i - \psi_i^\ell > 0) \\
\le &\ \frac{1}{4} \sum_{\widehat{C}K^\xi \le |i-j| \le M} B_{ij}[(v_i - \psi_i^\ell)_+ - (v_j - \psi_j^\ell)_+]^2 \\
&+ 4 \sum_{\widehat{C}K^\xi \le |i-j| \le M} B_{ij}|\psi_i^\ell - \psi_j^\ell|^2 \cdot \mathbf{1}(v_i - \psi_i^\ell > 0).
\end{aligned}$$

The first term is bounded by $\frac{1}{2}\mathfrak{b}[(v - \psi^\ell)_+, (v - \psi^\ell)_+]$ and can be absorbed in the first term on the r.h.s. of (10.79). By the simple estimate $|\psi_i^\ell - \psi_j^\ell| \le C\ell|i-j|/M$ and (10.5), the second term is bounded by

$$4 \sum_{\widehat{C}K^\xi \le |i-j| \le M} B_{ij}|\psi_i^\ell - \psi_j^\ell|^2 \cdot \mathbf{1}(v_i - \psi_i^\ell > 0) \le C\ell^2 M^{-1} \sum_i \mathbf{1}(v_i - \psi_i^\ell > 0), \tag{10.82}$$

where we have again used that the summation over i is restricted to $d_i^I \ge K/3$. Thus

$$\begin{aligned}
-\Omega_2 &\le \frac{1}{2}\mathfrak{b}[(v - \psi^\ell)_+, (v - \psi^\ell)_+] + C\ell^2 M^{-1} \sum_i \mathbf{1}(v_i - \psi_i^\ell > 0) \\
&\le \frac{1}{2}\mathfrak{a}[(v - \psi^\ell)_+, (v - \psi^\ell)_+] + C\ell^2 M^{-1} \sum_i \mathbf{1}(v_i - \psi_i^\ell > 0)
\end{aligned}$$

using $\mathfrak{b} \le \mathfrak{a}$.

A similar estimate is performed for Ω_3, but in the corresponding last term we use

$$|\psi_i^\ell - \psi_j^\ell| \le CK^\xi(\ell/M)$$

for $|i-j| \le \widehat{C} K^{\xi}$. Thus we have

$$
\begin{aligned}
-\Omega_3 &\le \sum_{|i-j|\le \widehat{C}K^{\xi}} B_{ij} |\psi_i^{\ell} - \psi_j^{\ell}|^2 \cdot \mathbf{1}(v_i - \psi_i^{\ell} > 0) \\
&\le C K^{2\xi} (\ell/M)^2 \sum_{|i-j|\le \widehat{C}K^{\xi}} \mathbf{1}(v_i - \psi_i^{\ell} > 0) B_{ij} \\
&\le C \frac{K^{3\xi}\ell^2}{M^2} \sum_i \mathbf{1}(v_i - \psi_i^{\ell} > 0)[B_{i,i+1} + B_{i,i-1}].
\end{aligned}
$$

Here we have just overestimated sums by $\widehat{C} K^{\xi}$. The conclusion of the energy estimate is

$$
\begin{aligned}
\partial_t \frac{1}{2} \sum_i [v_i - \psi_i^{\ell}]_+^2 \le& -\frac{1}{2}\mathfrak{a}[(v-\psi^{\ell})_+, (v-\psi^{\ell})_+] + |\bar{u}| \sum_i (v_i - \psi_i^{\ell})_+ W_i \\
&+ \frac{C\ell}{M} \sum_i (v_i - \psi_i^{\ell})_+ + \frac{C\ell^2}{M} \sum_i \mathbf{1}(v_i - \psi_i^{\ell} > 0) + \Omega_4, \qquad (10.83) \\
\Omega_4 :=& \frac{CK^{3\xi}\ell^2}{M^2} \sum_i \mathbf{1}(v_i - \psi_i^{\ell} > 0)[B_{i,i+1} + B_{i,i-1}].
\end{aligned}
$$

Due to (10.29), we can assume that the summations in (10.83) over i are restricted to $|i| \le M^{1+\kappa}$. In this regime we have $d_i \ge cK$ thanks to $M^{1+\kappa} \le K/2$, therefore $W_i \le CK^{-1+\xi}$ by (10.4). Using the bound (10.27), we see that the error term $|\bar{u}| \sum_i (v_i - \psi_i^{\ell})_+ W_i$ can be absorbed into the first error term in line (10.83).

Let $T_k := -M(1+2^{-k})$ and $\ell_k := (\ell/3)(1-2^{-k}) \nearrow \ell/3$ where $k = 1, \ldots, C \log M$. We claim that

$$
\begin{aligned}
\int_{\tau}^{t} \Omega_4 \, ds &\le \frac{CK^{3\xi}\ell^2}{M^{1-\kappa}} \int_{\tau}^{t} ds \, \frac{1}{M^{1+\kappa}} \sum_{|i|\le M^{1+\kappa}} [B_{i,i+1} + B_{i,i-1}](s) \\
&\le C[(t-\tau)+1] K^{3\xi+\rho} \ell^2 M^{\kappa-1} \qquad (10.84)
\end{aligned}
$$

for any integer $k \le C \log M$ and for any pairs $(t,\tau) \in [T_k, 0] \times [T_{k-1}, T_{k-1} + 2^{-k-1}M]$. The estimate (10.84) holds because

$$
\int_{\tau}^{t} [\ldots] \, ds \le \int_{T_{k-1}}^{t} [\ldots] \, ds \le 8\big|t-\tau\big| + 1,
$$

where we have used the fact that the point $(T_{k-1}, Z=0)$ is regular (see (10.26)).

Define

$$
U_k = \sup_{t\in[T_k,0]} \frac{1}{M\ell_k^2} \sum_i (v_i - \psi_i^{\ell_k})_+^2(t) + \frac{1}{M\ell_k^2} \int_{T_k}^{0} \mathfrak{a}[(v-\psi^{\ell_k})_+, (v-\psi^{\ell_k})_+](s) \, ds. \qquad (10.85)
$$

Integrating (10.83) from τ to t with $\tau \in [T_{k-1}, T_{k-1}+2^{-k-1}M] = [T_{k-1}, T_k - 2^{-k-1}M]$ and $t \in [T_k, 0]$, we deduce from (10.84) that

$$\begin{aligned}
&\sum_i [v_i - \psi_i^{\ell_k}]_+^2(t) + \int_\tau^t \mathfrak{a}[(v-\psi^{\ell_k})_+, (v-\psi^{\ell_k})_+](s)\,ds \\
&\le \sum_i [v_i - \psi_i^{\ell_k}]_+^2(\tau) + C\int_\tau^t \left[\frac{\ell_k}{M}\sum_i (v_i - \psi_i^{\ell_k})_+(s) + \frac{\ell_k^2}{M}\sum_i \mathbf{1}(v_i - \psi_i^{\ell_k} > 0)(s)\right] ds \\
&\quad + C[(t-\tau)+1]K^{3\xi+\rho}\ell^2 M^{\kappa-1}. \qquad (10.86)
\end{aligned}$$

Taking the average over $\tau \in [T_{k-1}, T_{k-1}+2^{-k-1}M]$ and using the fact that in this regime $2^{-k-1}M \le t-\tau \le M$, we obtain

$$\begin{aligned}
&\sum_i [v_i - \psi_i^{\ell_k}]_+^2(t) + \int_{T_k}^t \mathfrak{a}[(v-\psi^{\ell_k})_+, (v-\psi^{\ell_k})_+](s)\,ds \\
&\quad \le C\frac{2^{k+1}}{M}\int_{T_{k-1}}^{T_k - 2^{-k-1}M} \sum_i [v_i - \psi_i^{\ell_k}]_+^2(s)\,ds \\
&\qquad + C\int_{T_{k-1}}^t \left[\frac{\ell_k}{M}\sum_i (v_i - \psi_i^{\ell_k})_+(s) + \frac{\ell_k^2}{M}\sum_i \mathbf{1}(v_i - \psi_i^{\ell_k} > 0)(s)\right] ds + CK^{3\xi+\rho}\ell^2 M^\kappa.
\end{aligned}$$

Dividing through by $M\ell_k^2$ and taking the supremum over $t \in [T_k, 0]$, for $k \ge 1$ we get

$$\begin{aligned}
U_k &\le C\frac{2^{k+1}}{M^2}\int_{T_{k-1}}^0 \sum_i \left[\frac{1}{\ell_k^2}[v_i - \psi_i^{\ell_k}]_+^2 + \frac{1}{\ell_k}(v_i - \psi_i^{\ell_k})_+ + \mathbf{1}(v_i - \psi_i^{\ell_k} > 0)\right](s)\,ds \\
&\quad + M^\kappa \frac{CK^{3\xi+\rho}}{M}. \qquad (10.87)
\end{aligned}$$

The first three integrands have the same scaling dimensions as v^2/ℓ^2. One key idea is to estimate them in terms of the L^4-norm of v and then use the Sobolev inequality. It is elementary to check these three integrands can be bounded by the L^4-norm of $(v-\psi^{\ell_k})_+$, by using the fact that if $v_i \ge \psi_i^{\ell_k}$, then $v_i - \psi_i^{\ell_{k-1}} \ge \ell_k - \ell_{k-1} = 2^{-k}\frac{\ell}{3} \ge 2^{-(k+2)}\ell$:

$$\begin{aligned}
\sum_i (v_i - \psi_i^{\ell_k})_+ &\le \sum_i (v_i - \psi_i^{\ell_k})_+ \cdot \mathbf{1}(v_i - \psi_i^{\ell_{k-1}} > 2^{-(k+2)}\ell) \\
&\le (2^{k+1})^3 \ell_k^{-3}\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^4, \\
\sum_i \mathbf{1}(v_i - \psi_i^{\ell_k} > 0) &\le (2^{k+2})^4 \ell_k^{-4}\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^4, \\
\sum_i [v_i - \psi_i^{\ell_k}]_+^2 &\le (2^{k+2})^2 \ell_k^{-2}\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^4.
\end{aligned} \qquad (10.88)$$

We now use the local version of Proposition B.4 from Appendix B; we first verify its conditions. Set

$$\mathcal{I} := [\![-2K/3, 2K/3]\!], \qquad \widehat{\mathcal{I}} := [\![-3K/4, 3K/4]\!]. \tag{10.89}$$

Clearly $f_i := (v_i - \psi_i^{\ell_{k-1}})_+$ is supported in $\mathcal{I}$; this follows from $|Z| \le K/2$, (10.29) and $M^{1+\kappa} \le M^{(\vartheta+1)/2} \ll M^\vartheta = K$. By the lower bounds on $B_{ij}(s)$ in (10.3) and (10.5) (with $C \ge 4$ in (10.5) to guarantee that the lower bound holds for any $i, j \in \widehat{\mathcal{I}}$) conditions (B.17), (B.18) hold with the choice $b = K^{-\xi}$, $a = \widehat{C}^{-1}K^{-\xi}$ and $r = C$, where C and $\widehat{C}$ are the constants from (10.5).

From (B.19) we then have

$$\begin{aligned}
\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^4 &\\
\le C\Big[\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^2\Big]&\Big[\mathfrak{a}[(v - \psi^{\ell_{k-1}})_+, (v - \psi^{\ell_{k-1}})_+] + \frac{1}{K}\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^2\Big]\\
+ CK^{4\xi}\max_i (v_i - \psi_i^{\ell_{k-1}})_+^4 &
\end{aligned} \tag{10.90}$$

(we have omitted the time variable $s \in \mathcal{T}$). The last term can be estimated by using (10.29) and (10.30) as

$$\max_i (v_i(t) - \psi_i^{\ell_{k-1}})_+ \le \max\{v_i(t) : |i - Z| \le M^{1+\kappa}\} \le C\ell M^{\chi/2}$$

for any $t \in [-2M, 0]$. For $t \in \mathcal{G}^*$ we have the stronger bound from (10.31),

$$\max_i (v_i(t) - \psi_i^{\ell_{k-1}})_+ \le \max\{v_i(t) : |i - Z| \le M^{1+\kappa}\} \le C\ell M^{\chi/8}, \qquad t \in \mathcal{G}^*.$$

Inserting these estimates, (10.88) and (10.4) into (10.87), and splitting the time integration into $\mathcal{G}^*$ and its complement, we conclude that for $k \ge 2$,

$$\begin{aligned}
U_k &\le C(2^{k+2})^5 \frac{1}{M^2\ell_k^4}\int_{T_{k-1}}^0 ds \Big[\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^2(s)\Big]\\
&\qquad \times \Big[\mathfrak{a}[(v - \psi^{\ell_{k-1}})_+, (v - \psi^{\ell_{k-1}})_+](s) + \frac{1}{K}\sum_i (v_i - \psi_i^{\ell_{k-1}})_+^2(s)\Big]\\
&\qquad + \frac{1}{M}[CM^\kappa K^{3\xi+\rho} + 32^k M^{\chi/2}K^{4\xi} + 32^k M^{2\chi-1}K^{4\xi}|[-T_{k-1}, 0] \setminus \mathcal{G}^*|]\\
&\le 32^k[C_1 U_{k-1}^2 + M^{-1+\chi}K^{-\rho}],
\end{aligned} \tag{10.91}$$

recalling that $|[-T_{k-1}, 0] \setminus \mathcal{G}^*| \le CM^{1/4}$, $K = M^\vartheta \le M^{\vartheta_0}$ and $\chi \ge \kappa + 10(\xi + \rho)\vartheta_0$. We have also used $|T_k| \le K$.

For $k = 1$, we estimate the integrands in (10.87) by L^2-norms. We have the following general estimates for any $\ell' < \ell''$:

$$\sum_i (v_i - \psi_i^{\ell''})_+ \le \sum_i (v_i - \psi_i^{\ell'})_+ \cdot \mathbf{1}(v_i - \psi_i^{\ell'} > \ell'' - \ell')$$
$$\le \frac{1}{\ell'' - \ell'} \sum_i (v_i - \psi_i^{\ell'})_+^2, \tag{10.92}$$
$$\sum_i \mathbf{1}(v_i - \psi_i^{\ell''} > 0) \le \frac{1}{(\ell'' - \ell')^2} \sum_i (v_i - \psi_i^{\ell'})_+^2.$$

We use (10.92) with $\ell'' = \ell_1$ and $\ell' = 0$ in (10.87); this implies that

$$U_1 \le \frac{C}{\ell_1^2 M^2} \int_{-2M}^0 \sum_i ds\, (v_i - \psi_i)_+^2(s) + CM^{-1+\chi}.$$

Without loss of generality, we assume that $C_1 \ge 2$, where C_1 is the constant in (10.91). Now choose the universal constant ε_0 in (10.28) so small and M large enough so that this last inequality implies

$$U_1 \le \frac{1}{32^6 C_1}. \tag{10.93}$$

Choose k_* such that $32^{k_*+2} C_1 = K^\rho$, i.e. k^* is of order $\rho \log K \ge \rho \log M$. Then from (10.91) for any $k \le k_*$ we have the recursive inequality

$$B_k \le B_{k-1}^2 + M^{-1+\chi} \quad \text{with} \quad B_k := 32^{k+2} C_1 U_k.$$

By a simple induction, this recursion implies

$$B_{k+1} \le (2B_1)^{2^k - 1} + 2M^{-1+\chi}.$$

Together with the initial estimate (10.93) we obtain $B_{k+1} \le 4M^{-1+\chi}$ for any integer k with $100 \log\log M \le k \le k_*$, in particular we can apply it to $k' = 100 \log\log M$ and obtain $U_{k'} \le CM^{-1+\chi}$. Notice that U_k is decreasing in k, as can be seen from the monotonicity in the definition (10.85) of U_k and from the fact that T_k and ℓ_k increase. Thus

$$U_k \le CM^{-1+\chi} \tag{10.94}$$

for any $k \ge 100 \log\log M$. Taking $k \to \infty$, we find from the L^2-norm term in U_k that (10.32) in Lemma 10.6 holds.

For the proof of (10.33), we notice that the estimate (10.94) together with the monotonicity also implies that

$$\frac{1}{M\ell^2} \int_{-M}^0 \mathfrak{a}[(v - \psi^{\ell/3})_+, (v - \psi^{\ell/3})_+](s)\, ds \le CM^{-1+\chi}$$

from the dissipation term in the definition of U_k.

Set

$$\mathcal{G} := \{t \in [-M, 0] : \mathfrak{a}[(v-\psi^{\ell/3})_+, (v-\psi^{\ell/3})_+](t) \le M^{\chi-1/4}\ell^2\}.$$

Then clearly

$$|[-M, 0] \setminus \mathcal{G}| \le CM^{1/4}.$$

We now use a Sobolev inequality (B.4) from Appendix B, with the choice of $p = 4$, $s = 1$ and $f_i := (v_i - \psi^{\ell/3})_+$. We recall the definitions of $\mathcal{I}$ and $\widehat{\mathcal{I}}$ from (10.89) and that $f_i = (v_i - \psi^{\ell/3})_+$ is supported in $\mathcal{I}$ by (10.29). Thus

$$\begin{aligned}\sum_i f_i^4 &\le C \sum_i f_i^2 \Big[\sum_{i \ne j \in \widehat{\mathcal{I}}} \frac{|f_i - f_j|^2}{|i-j|^2} + 2\sum_{i\in\mathcal{I}} |f_i|^2 \sum_{j \notin \widehat{\mathcal{I}}} \frac{1}{|i-j|^2}\Big] \\ &\le CK^{2\xi}\|f\|^2 \mathfrak{a}[f, f] + \frac{C}{K}\|f\|_2^4,\end{aligned}$$

where we have used the lower bound (10.3) on B_{ij}. Thus

$$\begin{aligned}\sum_i (v_i - \psi^{\ell/3})_+^4 \le CK^{2\xi} &\sum_i (v_i - \psi_i^{\ell/3})_+^2 \mathfrak{a}[(v-\psi^{\ell/3})_+, (v-\psi^{\ell/3})_+] \\ &+ \frac{C}{K}\Big[\sum_i (v_i - \psi_i^{\ell/3})_+^2\Big]^2. \end{aligned} \tag{10.95}$$

This implies that for any $t \in \mathcal{G}$ and any i,

$$\begin{aligned}(v_i(t) - \psi_i^{\ell/3})_+ &\le \|(v(t) - \psi^{\ell/3})_+\|_4 \\ &\le CK^{\xi/2}\Big(\sum_i (v_i(t) - \psi_i^{\ell/3})_+^2\Big)^{1/4}\big(M^{\chi-1/4}\ell^2\big)^{1/4} + \frac{C}{K^{1/4}}\Big[\sum_i (v_i(t) - \psi_i^{\ell/3})_+^2\Big]^{1/2} \\ &\le CM^{-1/20}\ell,\end{aligned} \tag{10.96}$$

where we have used (10.32) in the last step and the fact that $\chi \ge 10\xi\vartheta_0$ together with (10.25). This proves (10.33).

For the proof of (10.34), we first notice that it is sufficient to consider the case when $\widetilde{M}$ is of the form $\widetilde{M} = 2^{-m}M$, $m = 1, \dots, C\log M$. We now repeat the proof of (10.32) but with ℓ_k, $k \ge 1$, replaced by

$$\widehat{\ell}_k = 2\ell(1 - 2^{-k-2})/5 \tag{10.97}$$

in the definition of ψ^{ℓ_k} and working in the time interval of scale $\widetilde{M}$.

Set $\widehat{T}_k := -\widetilde{M}(1 + 2^{-k})$. Define

$$\widehat{U}_k = \sup_{t\in[\widehat{T}_k,0]} \frac{1}{M\widehat{\ell}_k^2}\sum_i (v_i - \psi_i^{\widehat{\ell}_k})_+^2(t) + \frac{1}{M\widehat{\ell}_k^2}\int_{\widehat{T}_k}^0 \mathfrak{a}[(v-\psi^{\widehat{\ell}_k})_+, (v-\psi^{\widehat{\ell}_k})_+](s)\,ds.$$

The previous proof is unchanged up to (10.84), and the integral of

$$\widehat{\Omega}_4(s) := \frac{CK^{3\xi}\widehat{\ell}^2}{M^2}\sum_i \mathbf{1}(v_i(s) - \psi_i^{\widehat{\ell}} > 0)[B_{i,i+1}(s) + B_{i,i-1}(s)]$$

is still estimated by (cf. (10.84))

$$\int_\tau^t \widehat{\Omega}_4(s)\,ds \le C[(t-\tau)+1]K^{3\xi+\rho}\widehat{\ell}^2 M^{-1+\kappa} \le C[(t-\tau)+1]\widehat{\ell}^2 K^{3\xi+\rho}M^{-1+\kappa}$$

for $\tau \in [\widehat{T}_{k-1}, \widehat{T}_{k-1} + 2^{-k-1}\widetilde{M}] = [\widehat{T}_{k-1}, \widehat{T}_k - 2^{-k-1}\widetilde{M}]$ and $t \in [\widehat{T}_k, 0]$. Here we have used (10.25).

Similarly to (10.4), we integrate (10.83) (with $\widehat{\ell}$ replacing ℓ) from τ to t,

$$\begin{aligned}
&\sum_i [v_i - \psi_i^{\widehat{\ell}_k}]_+^2(t) + \int_\tau^t \mathfrak{a}[(v-\psi^{\widehat{\ell}_k})_+, (v-\psi^{\widehat{\ell}_k})_+](s)\,ds \\
&\quad \le \sum_i [v_i - \psi_i^{\widehat{\ell}_k}]_+^2(\tau) + C\int_\tau^t \left[\frac{\ell_k}{M}\sum_i (v_i - \psi_i^{\widehat{\ell}_k})_+(s) + \frac{\widehat{\ell}_k^2}{M}\sum_i \mathbf{1}(v_i - \psi_i^{\widehat{\ell}_k} > 0)(s)\right] ds \\
&\quad\quad + C[(t-\tau)+1]\ell^2 K^{3\xi+\rho}M^{-1+\kappa}. \qquad (10.97)
\end{aligned}$$

Taking the average over $\tau \in [\widehat{T}_{k-1}, \widehat{T}_{k-1} + 2^{-k-1}\widetilde{M}] = [\widehat{T}_{k-1}, \widehat{T}_k - 2^{-k-1}\widetilde{M}]$ and using the fact that in this regime $2^{-k-1}\widetilde{M} \le t - \tau \le \widetilde{M}$, we obtain

$$\begin{aligned}
&\sum_i [v_i - \psi_i^{\widehat{\ell}_k}]_+^2(t) + \int_{\widehat{T}_k}^t \mathfrak{a}[(v-\psi^{\widehat{\ell}_k})_+, (v-\psi^{\widehat{\ell}_k})_+](s)\,ds \\
&\le C\frac{2^{k+1}}{\widetilde{M}}\int_{\widehat{T}_{k-1}}^{\widehat{T}_k - 2^{-k-1}\widetilde{M}} \sum_i [v_i - \psi_i^{\widehat{\ell}_k}]_+^2(s)\,ds \\
&+ C\int_{\widehat{T}_{k-1}}^t \left[\frac{\widehat{\ell}_k}{M}\sum_i (v_i - \psi_i^{\widehat{\ell}_k})_+(s) + \frac{\widehat{\ell}_k^2}{M}\sum_i \mathbf{1}(v_i - \psi_i^{\widehat{\ell}_k} > 0)(s)\right] ds + C\ell^2\widetilde{M}K^{3\xi+\rho}M^{-1+\kappa}.
\end{aligned}$$

Dividing through by $M\widehat{\ell}_k^2$ and taking the supremum over $t \in [\widehat{T}_k, 0]$, for $k \ge 1$ and using $\widetilde{M} \le M$, we have, as in (10.87),

$$\begin{aligned}
\widehat{U}_k &\le C\frac{2^{k+1}}{M\widetilde{M}}\int_{\widehat{T}_{k-1}}^0 \sum_i \left[\frac{1}{\widehat{\ell}_k^2}[v_i - \psi_i^{\widehat{\ell}_k}]_+^2 + \frac{1}{\widehat{\ell}_k}(v_i - \psi_i^{\widehat{\ell}_k})_+ + \mathbf{1}(v_i - \psi_i^{\widehat{\ell}_k} > 0)\right](s)\,ds \\
&\quad + C\widetilde{M}K^{3\xi+\rho}M^{-2+\kappa}. \qquad (10.98)
\end{aligned}$$

Using the bounds (10.88) and Proposition B.4 as in (10.4)–(10.91), instead of (10.91) we get

$$\widehat{U}_k \le \frac{(2^{k+2})^5}{M\widetilde{M}\widehat{\ell}_k^4} \int_{\widehat{T}_{k-1}}^0 ds \Big[\sum_i (v_i - \psi_i^{\widehat{\ell}_{k-1}})_+^2(s)\Big] \mathfrak{a}[(v-\psi^{\widehat{\ell}_{k-1}})_+, (v-\psi^{\widehat{\ell}_{k-1}})_+](s)$$
$$+ C\widetilde{M}M^{-2+\chi}K^{-\rho}$$
$$\le 32^k \Big[C_1 \frac{M}{\widetilde{M}} \widehat{U}_{k-1}^2 + C\frac{\widetilde{M}}{M} M^{-1+\chi} K^{-\rho} \Big], \quad k \ge 2.$$

Similarly to the proof of (10.94), this new recurrence inequality has the solution

$$\widehat{U}_k \le C\widetilde{M}M^{-2+\chi} \tag{10.99}$$

for any sufficiently large k, as long as the recursion can be started, i.e. if we knew

$$\widehat{U}_1 \ll \widetilde{M}/M. \tag{10.100}$$

For $k = 1$ the estimate (10.98) together with (10.92) (with $\widehat{\ell}_1$ replacing ℓ_1) becomes

$$\widehat{U}_1 \le \frac{C}{M\widetilde{M}} \int_{-2\widetilde{M}}^0 \sum_i \Big[\frac{1}{\widehat{\ell}_1^2} [v_i - \psi_i^{\widehat{\ell}_1}]_+^2 + \frac{1}{\widehat{\ell}_1} (v_i - \psi_i^{\widehat{\ell}_1})_+ + \mathbf{1}(v_i - \psi_i^{\widehat{\ell}_1} > 0) \Big](s)\, ds$$
$$+ C\widetilde{M}M^{-2+\chi}$$
$$\le \frac{C}{M\widetilde{M}} \int_{-2\widetilde{M}}^0 \sum_i \frac{1}{\ell^2} [v_i(s) - \psi_i^{\ell/3}]_+^2\, ds + C\widetilde{M}M^{-2+\chi} \le \frac{CM^\chi}{M} + C\widetilde{M}M^{-2+\chi}.$$

In the second step we have used (10.92) with $\ell'' = \widehat{\ell}_1$ and $\ell' = \ell/3$ noting that $\widehat{\ell}_1 = \frac{7}{20}\ell > \frac{1}{3}\ell$. In the last step we have used (10.32) and $2\widetilde{M} \le M$. Thus (10.100) is satisfied if $\widetilde{M} \gg M^\chi$.

Finally, taking $k \to \infty$ in (10.99) implies (10.34). This completes the proof of Lemma 10.6.

10.5. Proof of Lemma 10.7 (second De Giorgi lemma)

Set $Z = 0$ for simplicity. Since the statement is stronger if μ and δ are reduced, we can assume that they are small positive numbers, e.g. we can assume $\mu, \delta < 1/8$. We are looking for a sufficiently small λ so that there will be a positive γ with the stated properties. The key ingredient of the proof is an energy inequality (10.106) including a new dissipation term which was dropped in the proof of Lemma 10.6. Most of this section closely follows the argument in [13]; the main change is that we need to split time integrations into "good" and "bad" times. The argument [13] applies to the good times. The bad times have a small measure, so their contribution is negligible.

10.5.1. Dissipation with the good term. Let $-3M \le T_1 < T_2 < 0$. For any $t \in [-3M, 0]$, define

$$\theta_i(t) := \mathbf{1}(|i| \le 9M) \cdot \mathbf{1}(t \in \mathcal{G}) + \mathbf{1}(|i| \le M^{1+\kappa_1}) \cdot \mathbf{1}(t \notin \mathcal{G}).$$

We use the calculation (10.79)–(10.80) (with cutoff $\varphi^{(1)}$ instead of ψ^ℓ) but we keep the "good" $\mathfrak{b}[(v-\varphi^{(1)})_+, (v-\varphi^{(1)})_-] \ge 0$ term that was estimated trivially in (10.79) and we drop the (positive) potential term in $\mathfrak{a}$. We have

$$\begin{aligned}
&\frac{1}{2}\sum_i [v_i(t) - \varphi_i^{(1)}]_+^2 \Big|_{t=T_1}^{T_2} + \int_{T_1}^{T_2} \mathfrak{b}[(v(t)-\varphi^{(1)})_+, (v(t)-\varphi^{(1)})_+]\, dt \\
&\quad \le -\int_{T_1}^{T_2} \mathfrak{b}[(v(t)-\varphi^{(1)})_+, (v(t)-\varphi^{(1)})_-]\, dt - \int_{T_1}^{T_2} \mathfrak{b}[(v(t)-\varphi^{(1)})_+\theta, \varphi^{(1)}]\, dt \\
&\qquad + |\bar{u}| \int_{T_1}^{T_2} \sum_i (v_i(t) - \varphi_i^{(1)})_+ W_i \theta_i \, dt. \qquad (10.101)
\end{aligned}$$

Notice that we have inserted the characteristic function $\theta_i(t)$ using the fact that (10.39) and (10.36) imply $v_i(t) \le \varphi_i^{(1)}$ for $|i| \ge 9M$, $t \in \mathcal{G}$, and $v_i(t) \le \varphi_i^{(1)} = \widetilde{\psi}_i$ for $|i| \ge M^{1+\kappa_1}$ and $t \in [-3M, 0]$, i.e. $v_i - \varphi^{(1)} = (v_i - \varphi^{(1)})\theta_i$ for any time. Moreover, $v_i(t) - \varphi_i^{(1)} \le \lambda\ell$ for $t \in \mathcal{G}$ and $|i| \le 9M$.

The last error term in (10.101) is estimated trivially; in the regime $|i| \le M^{1+\kappa_1}$ we have $W_i \le CK^{-1+\xi}$ and then from (10.40), $|\mathcal{G}| \le CM^{1/4}$ and (10.38) we have

$$\begin{aligned}
|\bar{u}| \int_{T_1}^{T_2} \sum_i (v_i(t) - \varphi_i^{(1)})_+ W_i \theta_i \, dt &\le \lambda^2\ell^2 (T_2 - T_1) + C\lambda\ell^2 M^{\kappa_1+\kappa_2} |\mathcal{G}^c| \\
&\le C\lambda^2\ell^2[(T_2 - T_1) + \lambda^{-1} M^{1/2}] \qquad (10.102)
\end{aligned}$$

after splitting the integration regime into "good" times $\mathcal{G}$ and "bad" times $\mathcal{G}^c := [-3M, 0] \setminus \mathcal{G}$. We have also used (10.37).

The other error term in (10.101) will be estimated by a Schwarz inequality; here we use the identity

$$\mathfrak{b}(f\theta, g) = \sum_{ij} (f_i\theta_i - f_j\theta_j) B_{ij}(g_i - g_j) = \sum_{ij} (f_i\theta_i - f_j\theta_j)(\theta_i + \theta_j - \theta_i\theta_j) B_{ij}(g_i - g_j)$$

for any functions f and g, so

$$|\mathfrak{b}(f\theta, g)| \le \frac{1}{2}\sum_{ij} (f_i\theta_i - f_j\theta_j)^2 B_{ij} + 2\sum_{ij} \theta_i B_{ij}(g_i - g_j)^2,$$

i.e.

$$\begin{aligned}
&|\mathfrak{b}[(v(t)-\varphi^{(1)})_+\theta, \varphi^{(1)}]| \\
&\qquad \le \frac{1}{2}\mathfrak{b}[(v(t)-\varphi^{(1)})_+\theta, (v(t)-\varphi^{(1)})_+\theta] + 2\sum_{ij} \theta_i B_{ij}(\varphi_i^{(1)} - \varphi_j^{(1)})^2.
\end{aligned}$$

The first term will be absorbed in the quadratic term on the left of (10.101). By definition of $\varphi^{(1)}$, for the second term we have to control

$$\int_{T_1}^{T_2}\Big[\lambda^2\sum_{i,j}(F_i-F_j)^2B_{ij}+\sum_{i,j}(\widetilde{\psi}_i-\widetilde{\psi}_j)^2B_{ij}\theta_i\Big](t)\,dt. \tag{10.103}$$

Since $|F_i-F_j|\le C\ell M^{-1}|i-j|$ and F_i-F_j is supported on $|i|,|j|\le 9M$, by splitting the summation into $|i-j|\le K^{\xi}$ and its complement, we can bound the first term by

$$\begin{aligned}
&\int_{T_1}^{T_2}\lambda^2\sum_{i,j}(F_i-F_j)^2B_{ij}(t)\,dt\le\lambda^2\ell^2M^{-2}\int_{T_1}^{T_2}\sum_{|i|,|j|\le 9M}|i-j|^2B_{ij}(t)\\
&\quad\le\lambda^2\ell^2M^{-2}K^{3\xi}\int_{T_1}^{T_2}\sum_{|i|\le 9M}B_{i,i+1}(t)\,dt+C\lambda^2\ell^2M^{-2}\int_{T_1}^{T_2}\sum_{\substack{|i|,|j|\le 9M\\|i-j|\ge K^{\xi}}}\frac{|i-j|^2}{|i-j|^2}\\
&\quad\le\lambda^2\ell^2M^{-2}K^{3\xi}\int_{-3M}^{0}\sum_{|i|\le 9M}B_{i,i+1}(t)\,dt+C\lambda^2\ell^2(T_2-T_1),
\end{aligned}$$

where we have used $B_{i,j}\le B_{i,i+1}$ in the first regime and the upper bound in (10.5) in the other regime. By the regularity at $(Z,0)=(0,0)$ we can bound the last line by

$$C\lambda^2\ell^2K^{3\xi+\rho}+C\lambda^2\ell^2(T_2-T_1)\le C\lambda^2\ell^2[(T_2-T_1)+M^{1/2}]$$

(we have also used (10.37) and $K\le M^{\vartheta_0}$).

For the second term in (10.103) and for $t\in\mathcal{G}$ we use the fact that $\widetilde{\psi}_i\theta_i(t)=0$ and that the supports of θ_i and $\widetilde{\psi}_j$ are separated by a distance of order $M\gg K^{\xi}$. Thus we can use the upper bound in (10.5) to estimate the kernel:

$$\begin{aligned}
\int_{T_1}^{T_2}\mathbf{1}(t\in\mathcal{G})\sum_{i,j}(\widetilde{\psi}_i-\widetilde{\psi}_j)^2B_{ij}(t)\theta_i(t)\,dt&\le C\int_{T_1}^{T_2}\sum_{|i|\le 9M}\sum_{|j|\ge M\lambda^{-4}}\frac{\widetilde{\psi}_j^2}{|i-j|^2}\,dt\\
&\le CM(T_2-T_1)\sum_{|j|\ge M\lambda^{-4}}\frac{\widetilde{\psi}_j^2}{|j|^2}\le C\ell^2\lambda^2(T_2-T_1),
\end{aligned} \tag{10.104}$$

where we have used $\widetilde{\psi}_j\sim\ell(j/M)^{1/4}$ for large j. For times $t\notin\mathcal{G}$, we use

$$(\widetilde{\psi}_i-\widetilde{\psi}_j)^2\le\frac{C\ell^2}{M^{1/2}}\frac{(i-j)^2}{|i|^{3/2}+|j|^{3/2}}$$

to get

$$
\begin{aligned}
&\int_{T_1}^{T_2} \mathbf{1}(t\notin\mathcal{G})\sum_{i,j}(\widetilde{\psi}_i-\widetilde{\psi}_j)^2 B_{ij}(t)\theta_i(t)\,dt\\
&\quad\le \int_{T_1}^{T_2}\mathbf{1}(t\notin\mathcal{G})\frac{C\ell^2}{M^{1/2}}\sum_{|i|\le M^{1+\kappa_1}}\sum_{|j|\ge M\lambda^{-4}}\frac{1}{|i|^{3/2}+|j|^{3/2}}\,dt\\
&\qquad+\int_{T_1}^{T_2}\mathbf{1}(t\notin\mathcal{G})\frac{C\ell^2}{M^{1/2}}\sum_{|i|\le M^{1+\kappa_1}}\sum_{|j|\ge M\lambda^{-4}}B_{ij}(t)\frac{|i-j|^2\cdot\mathbf{1}(|i-j|\le K^{\xi})}{|i|^{3/2}+|j|^{3/2}}\,dt\\
&\quad\le CM^{1+\kappa_1}|\mathcal{G}^c|\frac{\ell^2}{M^{1/2}}\lambda^2M^{-1/2}\\
&\qquad+CK^{2\xi}\frac{\ell^2}{M^{1/2}}\int_{-3M}^{0}\sum_{|i|\le M^{1+\kappa_1}}\sum_{|j|\ge M\lambda^{-4}}B_{ij}(t)\frac{\mathbf{1}(|i-j|\le K^{\xi})}{|i|^{3/2}+|j|^{3/2}}\,dt\\
&\quad\le C\lambda^2\ell^2M^{\kappa_1+1/4}+CK^{3\xi}\frac{\ell^2}{M^{1/2}}\frac{1}{(M\lambda^{-4})^{3/2}}\int_{-3M}^{0}\sum_{|i|\le M^{1+\kappa_1}}B_{i,i+1}(t)\,dt\\
&\quad\le C\lambda^2\ell^2M^{1/2}.
\end{aligned}\tag{10.105}
$$

Here we have first separated the summations over i, j into $|i-j|\ge K^{\xi}$ and its complement. Then in the first regime we have used the upper bound in (10.5) and that the measure of the bad times is small, i.e., (10.39), to estimate the time integral; in the second regime we used regularity at $(Z,0)$ and the fact that $K^{\xi}\ll M^{1/10}$ by (10.37). Inserting the error estimates (10.102), (10.104) and (10.105) into (10.101), we obtain

$$
\begin{aligned}
&\frac{1}{2}\sum_i[v_i(t)-\varphi_i^{(1)}]_+^2\Big|_{t=T_1}^{T_2}+\frac{1}{2}\int_{T_1}^{T_2}\mathfrak{b}[(v(t)-\varphi^{(1)})_+,(v(t)-\varphi^{(1)})_+]\,dt\\
&\qquad\le-\int_{T_1}^{T_2}\mathfrak{b}[(v(t)-\varphi^{(1)})_+,(v(t)-\varphi^{(1)})_-]\,dt+C\ell^2\lambda^2[(T_2-T_1)+M^{1/2}].
\end{aligned}
$$

Define $H(t)=\sum_i(v_i(t)-\varphi_i^{(1)})_+^2$. We have

$$
H(T_2)+\int_{T_1}^{T_2}\mathfrak{b}[(v(t)-\varphi^{(1)})_+,(v(t)-\varphi^{(1)})_-]\,dt\le H(T_1)+C\ell^2\lambda^2[(T_2-T_1)+M^{1/2}]\tag{10.106}
$$

for any $-3M\le T_1<T_2<0$. Notice that $\mathfrak{b}(f_+,f_-)\ge 0$ for any function f. Since $|v_i(t)-\varphi_i^{(1)}|\le\lambda\ell\theta_i$ for all $t\in\mathcal{G}$, we also have

$$
H(t)\le C\lambda^2\ell^2M,\qquad t\in\mathcal{G}.\tag{10.107}
$$

10.5.2. Time slices when the good term helps. Let $\Sigma\subset\mathcal{G}$ be the set of times that $v(T)$ is substantially below $\varphi^{(0)}$, i.e.,

$$
\Sigma:=\big\{T\in(-3M,-2M)\cap\mathcal{G}:\#\{|i|\le M: v_i(T)\le\varphi_i^{(0)}\}\ge\tfrac{1}{4}\mu M\big\}.
$$

We see from (10.39) and (10.41) that

$$|\Sigma| \geq \tfrac{1}{4} M\mu - C M^{1/4} \geq \tfrac{1}{5} M\mu. \tag{10.108}$$

By (10.106) (applied to $T_1 = \min \Sigma$, $T_2 = -2M$) and (10.107) (applied to $t = T_1$),

$$\begin{aligned} C\lambda^2 \ell^2 M &\geq \int_\Sigma \mathfrak{b}[(v(t) - \varphi^{(1)})_+, (v(t) - \varphi^{(1)})_-]\, dt \\ &\geq -\int_\Sigma \sum_{ij} (v_i(t) - \varphi_i^{(1)})_+ B_{ij}(t)(v_j(t) - \varphi_j^{(1)})_-\, dt \\ &\geq -cM^{-2} \int_\Sigma \sum_{ij} (v_i(t) - \varphi_i^{(1)})_+ (v_j(t) - \varphi_j^{(1)})_-\, dt, \end{aligned} \tag{10.109}$$

where we have used the fact that $v_i(t) - \varphi_i^{(1)}$ is supported on $|i| \leq 9M$ (for $t \in \mathcal{G}$) and

$$B_{ij}(t) \geq \bar{c} M^{-2}, \qquad |i|, |j| \leq 9M, \tag{10.110}$$

with some positive constant $\bar{c}$ (this follows from the lower bound in (10.5), where $|i| \leq 9M$ and $M \leq K/10$ guarantee that $d_i \geq K/C$ holds, and $K^\xi \ll M$ guarantees that (10.5) can be used for the extreme points $i = -9M$, $j = 9M$, and finally we have used monotonicity $B_{ij} \geq B_{-9M,9M}$ for any $|i|, |j| \leq 9M$). For $t \in \Sigma$ the number of j's with $|j| \leq M$ such that $v_j(t) \leq \varphi_j^{(0)}$ is at least $\frac{1}{5}\mu M$; for such j's we have

$$-(v_j(t) - \varphi_j^{(1)})_- \geq \varphi_j^{(1)} - \varphi_j^{(0)} \geq (1-\lambda)\ell \geq \ell/2.$$

Thus we can bound (10.109) by

$$\geq c\ell M^{-1} \frac{\mu}{10} \int_\Sigma \sum_i (v_i(t) - \varphi_i^{(1)})_+\, dt \geq cM^{-1} \frac{\mu}{10\lambda} \int_\Sigma \sum_i (v_i(t) - \varphi_i^{(1)})_+^2\, dt,$$

where we have used $(v_i(t) - \varphi_i^{(1)})_+ \leq \lambda\ell$ for $t \in \mathcal{G}$.

Altogether we have proved

$$\int_\Sigma \sum_i (v_i(t) - \varphi_i^{(1)})_+^2\, dt \leq C\lambda^3 \mu^{-1} \ell^2 M^2 \leq \lambda^{3-1/8} \ell^2 M^2$$

if λ is sufficiently small (depending on μ). Thus there exists a subset $\Theta \subset \Sigma$ such that $|\Theta| \leq \lambda^{1/8} M$, and we have

$$\sum_i (v_i(t) - \varphi_i^{(1)})_+^2 \leq \lambda^{3-1/4} \ell^2 M, \qquad \forall t \in \Sigma \setminus \Theta.$$

Choosing λ small and recalling (10.108) we see that

$$\sum_i (v_i(t) - \varphi_i^{(1)})_+^2 \leq \lambda^{3-1/4} \ell^2 M \tag{10.111}$$

on a set of times t in $\Sigma \subset [-3M, -2M] \cap \mathcal{G}$ of measure at least $M\mu/8$. In particular this set of times is nonempty.

10.5.3. Finding the intermediate set. Since (10.42) is satisfied, there is a $T_0 \in (-2M, 0) \cap \mathcal{G}$ such that

$$\#\{i : (v_i(T_0) - \varphi_i^{(2)})_+ > 0\} \geq \tfrac{1}{2} M\delta - CM^{1/4}, \tag{10.112}$$

and we can choose a $T_1 \in \Sigma$ (then $T_1 < T_0$) such that

$$H(T_1) = \sum_i (v_i(T_1) - \varphi_i^{(1)})_+^2 \leq \lambda^{3-1/4} \ell^2 M$$

(such a T_1 exists by the conclusion of the previous section, (10.111)).

We also have

$$\begin{aligned} H(T_0) &= \sum_i (v_i(T_0) - \varphi_i^{(1)})_+^2 \geq \sum_i (\varphi_i^{(2)}(T_0) - \varphi_i^{(1)})_+^2 \cdot \mathbf{1}\big((v_i(T_0) - \varphi_i^{(2)})_+ > 0\big) \\ &\geq \sum_i \ell^2 (\lambda - \lambda^2)^2 F_i^2 \cdot \mathbf{1}\big((v_i(T_0) - \varphi_i^{(2)})_+ > 0\big) \geq C_F \frac{\lambda^2}{4} \ell^2 \delta^3 M \end{aligned} \tag{10.113}$$

with some positive constant C_F. This follows from (10.112); notice first that the set in (10.112) must lie in $[-9M, 9M]$ (see (10.36) and (10.39)), and even if the whole set (10.112) is near the "corner" (i.e. close to $i \sim \pm 9M$), still the sum of these F_i's is of order $\delta^3 M$ since F_i is linear near the endpoints $i = \pm 9M$.

Choose now λ small enough (depending on the fixed δ) such that

$$\lambda^{3-1/4} \ell^2 M \leq \tfrac{1}{16} C_F \lambda^2 \ell^2 \delta^3 M.$$

Since $H(T)$ is continuous and goes from a small value $H(T_1) \leq \frac{1}{16} C_F \lambda^2 \ell^2 \delta^3 M$ to a large value $H(T_0) \geq \frac{1}{4} C_F \lambda^2 \ell^2 \delta^3 M$, the set of intermediate times

$$D := \left\{ t \in (T_1, T_0) : \tfrac{1}{16} C_F \lambda^2 \ell^2 \delta^3 M < H(t) < \tfrac{1}{4} C_F \lambda^2 \ell^2 \delta^3 M \right\}$$

is nonempty.

Lemma 10.10. *The set D contains an interval of size at least $c\delta^3 M$ with some positive constant $c > 0$. Moreover, for any $t \in D \cap \mathcal{G}$, we have*

$$\#\{i : \varphi_i^{(2)} \leq v_i(t)\} \leq \tfrac{1}{2} \delta M. \tag{10.114}$$

Proof. By continuity, there is an intermediate time $T' \in D \subset [T_1, T_0]$ such that $H(T') = \frac{1}{8} C_F \lambda^2 \ell^2 \delta^3 M$. We can assume that T' is the largest such time, i.e.

$$H(t) > \tfrac{1}{8} C_F \lambda^2 \ell^2 \delta^3 M \quad \text{ for any } t \in [T', T_0] \cap D. \tag{10.115}$$

Let $T'' = T' + c\delta^3 M$ with a small $c > 0$. We claim that $[T', T''] \subset D$. For any $t \in [T', T'']$ we can use (10.106):

$$\begin{aligned} H(t) &\leq H(T') + C\ell^2\lambda^2[(t - T') + M^{1/2}] \leq \tfrac{1}{8} C_F \lambda^2 \ell^2 \delta^3 M + Cc\ell^2 \lambda^2 \delta^3 M \\ &< \tfrac{1}{4} C_F \lambda^2 \ell^2 \delta^3 M \end{aligned} \tag{10.116}$$

if c is sufficiently small. This means that as t runs through $[T', T'']$, $H(t)$ does not reach $\frac{1}{4}C_F\lambda^2\ell^2\delta^3 M$, in particular $[T', T''] \subset (T_1, T_0)$ since $H(T_0)$ is already above this threshold. Combining then (10.116) with (10.115), we get $[T', T''] \subset D$. This proves the first statement of the lemma.

For the second statement, suppose for contradiction that $\#\{i : \varphi_i^{(2)} \le v_i(\tau)\} > \frac{1}{2}\delta M$ for some $\tau \in D \cap \mathcal{G}$. Going through the estimate (10.113) but with T_0 replaced with τ, we would get $H(\tau) \ge C_F\frac{\lambda^2}{4}\ell^2\delta^3 M$, which contradicts $\tau \in D$. $\square$

Define the exceptional set $\mathcal{F} \subset D \cap \mathcal{G}$ of times where v is below $\varphi^{(0)}$, i.e.

$$\mathcal{F} := \Big\{t \in D \cap \mathcal{G} : \#\{|j| \le 8M : v_j(t) - \varphi_j^{(0)} \le 0\} \ge \mu M\Big\}.$$

This set is very small, since from (10.107) (applied to $t_{\max} := \sup \mathcal{F} \in \bar{\mathcal{G}}$) we have

$$\begin{aligned}
C\lambda^2\ell^2 M &\ge -\int_{-3M}^{t_{\max}} \sum_{ij} (v_i(t) - \varphi_i^{(1)})_+ B_{ij}(t)(v_j(t) - \varphi_j^{(1)})_- \, dt \\
&\ge -\int_{\mathcal{F}} \sum_{|i|,|j| \le 9M} (v_i(t) - \varphi_i^{(1)})_+ B_{ij}(t)(v_j(t) - \varphi_j^{(1)})_- \, dt \\
&\ge -\bar{c}M^{-2} \int_{\mathcal{F}} \sum_{|i|,|j| \le 9M} (v_i(t) - \varphi_i^{(1)})_+ (v_j(t) - \varphi_j^{(1)})_- \, dt \\
&\ge \frac{\bar{c}}{2M}\ell\mu \int_{\mathcal{F}} \sum_{|i| \le 9M} (v_i(t) - \varphi_i^{(1)})_+ \, dt,
\end{aligned}$$

where we have restricted the time integration to $\mathcal{F}$ in the first step, and then used (10.110) in the second step. In the third step we have used the fact that whenever $v_j(t) - \varphi_j^{(0)} \le 0$ (see the definition of $\mathcal{F}$), then $-(v_j(t) - \varphi_j^{(1)})_- \ge \ell(1-\lambda) \ge \ell/2$.

By (10.39), $(v_i(t) - \varphi_i^{(1)})_+ \le \ell\lambda$ and $(v_i(t) - \varphi_i^{(1)})_+ = 0$ if $|i| \ge 9M$ and $t \in \mathcal{G}$. Hence we can continue the above estimate:

$$C\lambda^2\ell^2 M \ge \frac{\bar{c}\mu}{2M\lambda}\int_{\mathcal{F}} \sum_i (v_i(t) - \varphi_i^{(1)})_+^2 \, dt = \frac{\bar{c}\mu}{2M\lambda}\int_{\mathcal{F}} H(t)dt \ge \frac{\bar{c}C_F}{32}\lambda\ell^2\delta^3\mu|\mathcal{F}|.$$

Here we have used $\mathcal{F} \subset D$ and the fact that in D we have a lower bound on $H(t)$. The conclusion is that

$$|\mathcal{F}| \le \frac{C\lambda}{\delta^3\mu}M$$

with some fixed constant $C > 0$. Using $|D| \ge c\delta^3 M$ from Lemma 10.10 and the smallness of $|\mathcal{G}^c|$, we thus have

$$|\mathcal{F}| \le |D \cap \mathcal{G}|/2, \qquad |D \cap \mathcal{G}| \ge \tfrac{1}{2}c\delta^3 M$$

if λ is sufficiently small, like $\lambda \le c\delta^6\mu$.

This means that $|D \setminus \mathcal{F}| \geq \frac{c}{2}\delta^3 M$. Now we claim that for $t \in (D \cap \mathcal{G}) \setminus \mathcal{F}$ we have

$$A(t) := \#\{i : \varphi_i^{(0)} < v_i(t) < \varphi_i^{(2)}\} \geq M/2. \tag{10.117}$$

This is because $t \notin \mathcal{F}$ guarantees that the lower bound $\varphi_i^{(0)} \leq v_i(t)$ is violated not more than $\mu M \leq M/4$ times among the indices $|i| \leq 8M$. By (10.114), the upper bound $v_i(t) \leq \varphi_i^{(2)}$ is violated not more than $\frac{1}{2}\delta M \leq M/4$ times.

Finally, integrating (10.117) gives

$$\int_{-3M}^{0} \#\{i : \varphi_i^{(0)} < v_i(t) < \varphi_i^{(2)}\}\, dt = \int_{-3M}^{0} A(t)\, dt \geq \frac{M}{2}|(D \cap \mathcal{G}) \setminus \mathcal{F}| \geq c\delta^3 M^2$$

with some small $c > 0$, which implies (10.43) with $\gamma := c\delta^3$. This proves Lemma 10.7. □

Appendix A. Proof of Lemma 4.5

First we show that on the set $\mathcal{R}_{L,K}$, the length of the interval $J = J_{\mathbf{y}} = (y_{L-K-1}, y_{L+K+1})$ satisfies (4.22). We first write

$$|J| = |y_{L+K+1} - y_{L-K-1}| = |\gamma_{L+K+1} - \gamma_{L-K-1}| + O(N^{-1+\xi\delta/2}).$$

Then we use the Taylor expansion $\varrho(x) = \varrho(\bar{y}) + O(x - \bar{y})$ around the midpoint $\bar{y}$ of J. Here we have used the fact that $\varrho \in C^1$ away from the edge. Thus from (2.14),

$$\mathcal{K}+1 = N\int_{\gamma_{L-K-1}}^{\gamma_{L+K+1}} \varrho = N\int_{y_{L-K-1}}^{y_{L+K+1}} \varrho + O(N^{\xi\delta/2}) = N|J|\varrho(\bar{y}) + O(N|J|^2) + O(N^{\xi\delta/2}),$$

since the contribution of the second order term in the Taylor expansion is of order $N|J|^2$. Expressing $|J|$ from this equation and using (4.1), we arrive at (4.22).

Now we prove (4.23). We set

$$U(x) := V(x) - \frac{2}{N}\sum_{j:\,|j-L|\geq K+K^{\xi}} \log|x - \gamma_j|.$$

The potential U is similar to $V_{\mathbf{y}}$, but the interactions with the external points near the edges of J (y_j's with $|j - L| < K + K^{\xi}$) have been removed and the external points y_j away from the edges have been replaced by their classical value γ_j. In proving (4.23), we will first compare $V_{\mathbf{y}}$ with an auxiliary potential U and then we compute U'.

First we estimate the difference $V_{\mathbf{y}}'(x) - U'(x)$. We fix $x \in J$, and for definiteness, we assume that $d(x) = x - y_{L-K-1}$, i.e. x is closer to the lower endpoint of J; the other case is analogous. We get (explanations will be given after the equation)

$$
\begin{aligned}
|V_{\mathbf{y}}'(x) - U'(x)| &\le \frac{1}{N}\sum_{K<|j-L|<K+K^\xi}\frac{1}{|x-y_j|} + \frac{1}{N}\sum_{|j-L|\ge K+K^\xi}\frac{|y_j-\gamma_j|}{|x-y_j|\,|x-\gamma_j|}\\
&\le \frac{CK^\xi}{Nd(x)} + \frac{N^{-1+\delta\xi/2}}{d(x)}\frac{1}{N}\Big[\sum_{j=\alpha N/2}^{L-K-K^\xi} + \sum_{j=L+K+K^\xi}^{N(1-\alpha/2)}\Big]\frac{1}{|x-\gamma_j|}\\
&\quad + \frac{CN^{-4/15+\varepsilon}}{N}\Big[\sum_{j=N^{3/5+\varepsilon}}^{\alpha N/2} 1 + \sum_{j=N(1-\alpha/2)}^{N-N^{3/5+\varepsilon}} 1\Big] + \frac{C}{N}\Big[\sum_{j=1}^{N^{3/5+\varepsilon}} 1 + \sum_{j=N-N^{3/5+\varepsilon}}^{N} 1\Big]\\
&\le \frac{CK^\xi}{Nd(x)} + \frac{CN^{-1+\delta\xi/2}\log N}{d(x)} + CN^{-4/15+\varepsilon} \le \frac{CK^\xi}{Nd(x)}.
\end{aligned}
\tag{A.1}
$$

Here for the first bulk sum, $j \in [\![N\alpha/2, L-K-K^\xi]\!]$, we have used $|y_j-\gamma_j| \le N^{-1+\xi\delta/2}$ from the definition of $\mathcal{R}_{L,K}$ and the fact that for $j \le L-K-K^\xi$ we have

$$
\begin{aligned}
x-\gamma_j &\ge y_{L-K-1}-\gamma_j \ge \gamma_{L-K-1}-\gamma_j - |y_{L-K-1}-\gamma_{L-K-1}|\\
&\ge cN^{-1}(L-K-1-j) - CN^{-1+\xi\delta/2} \ge c'N^{-1}(L-K-1-j)
\end{aligned}
$$

with some positive constants c, c'. This estimate allows one to sum up $|x-\gamma_j|^{-1}$ at the expense of a $\log N$ factor. A similar estimate holds for $j \ge L+K+K^\xi$. In the intermediate sum, $j \in [\![N^{3/5+\varepsilon}, N\alpha/2]\!]$, we have used $|y_j-\gamma_j| \le CN^{-4/15+\varepsilon}$ and the fact that $|x-y_j|$ and $|x-\gamma_j|$ are bounded from below by a positive constant since

$$
x-y_j \ge y_{L-K-1}-y_j \ge y_{\alpha N}-y_j \ge \gamma_{N\alpha}-\gamma_{N\alpha/2} + O(N^{-1+\xi\delta/2}) \ge c,
$$

and similarly for $x-\gamma_j$. Finally, very near the edge, e.g. for $j \le N^{3/5+\varepsilon}$, we have just estimated $|y_j-\gamma_j|$ by a constant. This explains (A.1).

Now we estimate $U'(x)$. We use the fact that the equilibrium measure $\varrho = \varrho_V$ satisfies the identity

$$
\frac{1}{2}V'(x) = \int \frac{\varrho(y)}{x-y}\,dy
$$

from the Euler–Lagrange equation for (2.13) (see [1, 12]). Thus $\frac{1}{2}|U'(x)| \le |\Omega_1| + |\Omega_2| + |\Omega_3|$ with

$$
\begin{aligned}
\Omega_1 &:= \int_{\gamma_{L-K-K^\xi}}^{\gamma_{L+K+K^\xi}} \frac{\varrho(y)}{x-y}\,dy,\\
\Omega_2 &:= \int_A^{\gamma_{L-K-K^\xi}} \frac{\varrho(y)}{x-y}\,dy - \frac{1}{N}\sum_{j=1}^{L-K-K^\xi}\frac{1}{x-\gamma_j},\\
\Omega_3 &:= \int_{\gamma_{L+K+K^\xi}}^{B} \frac{\varrho(y)}{x-y}\,dy - \frac{1}{N}\sum_{j=L+K+K^\xi}^{N}\frac{1}{x-\gamma_j},
\end{aligned}
$$

where $[A, B]$ is the support of the density ρ.

To estimate Ω_1, we use the Taylor expansion $\varrho(y) = \varrho(x) + O(|x-y|)$. For definiteness we again assume that $d(x) = x - y_{L-K-1}$, and use the fact that on $\mathcal{R}_{L,K}$ we have

$$\gamma_{L-K-K^\xi} \le y_{L-K-1} \le x \le y_{L+K+1} \le \gamma_{L+K+K^\xi}.$$

We thus obtain

$$\begin{aligned}\Omega_1 &= \int_{\gamma_{L-K-K^\xi}}^{\gamma_{L+K+K^\xi}} \frac{\varrho(x) + O(|x-y|)}{x-y}\,dy = \varrho(x)\log\frac{\gamma_{L+K+K^\xi} - x}{x - \gamma_{L-K-K^\xi}} + O(K/N)\\ &= \varrho(\bar{y})\log\frac{d_+(x)}{d_-(x)} + O(KN^{-1+\xi}).\end{aligned} \tag{A.2}$$

In the first step above we have computed the leading term of the integral, while the other term was estimated trivially using the fact that the integration length is $\gamma_{L+K+K^\xi} - \gamma_{L-K-K^\xi} = O(K/N)$. In the second step we have used the fact that $\varrho \in C^1$ away the edge, i.e. $\varrho(x) = \varrho(\bar{y}) + O(K/N)$. To estimate the logarithm, we have used

$$\begin{aligned}\gamma_{L+K+K^\xi} - x &= (\gamma_{L+K+K^\xi} - \gamma_{L+K+1}) + (\gamma_{L+K+1} - y_{L+K+1}) + (y_{L+K+1} - x)\\ &= \varrho(\bar{y})N^{-1}K^\xi + O(N^{-1+\xi\delta/2}) + (y_{L+K+1} - x) = d_+(x) + O(N^{-1+\xi\delta/2})\end{aligned}$$

and the similar relation

$$x - \gamma_{L-K-K^\xi} = d_-(x) + O(N^{-1+\xi\delta/2}).$$

Notice that the error term in (A.2) is smaller than the target estimate $K^\xi/(Nd(x))$ since $d(x) \le K/N \ll K^{-1+\xi}N^{-\xi}$.

Now we estimate the Ω_2 term; Ω_3 can be treated analogously. We can write (with the convention $\gamma_0 = A$)

$$\begin{aligned}|\Omega_2| &= \left|\sum_{j=1}^{L-K-K^\xi} \int_{\gamma_{j-1}}^{\gamma_j} \varrho(y)\left[\frac{1}{x-y} - \frac{1}{x-\gamma_j}\right]dy\right|\\ &\le C\sum_{j=1}^{L-K-K^\xi} (\gamma_j - \gamma_{j-1})\int_{\gamma_{j-1}}^{\gamma_j} \frac{\varrho(y)}{(x-y)^2}\,dy\\ &\le CN^{-1}\int_{A+\kappa}^{\gamma_{L-K-K^\xi}} \frac{dy}{(x-y)^2} + CN^{-2/3}\int_A^{A+\kappa} \frac{dy}{(x-y)^2} \le \frac{CN^{-1}}{d(x)}.\end{aligned}$$

In the first step we have used

$$\int_{\gamma_{j-1}}^{\gamma_j} \varrho(y) = \frac{1}{N}$$

from (2.14). In the second step we have used the fact that $\gamma_j - \gamma_{j-1} = O_\kappa(N^{-1})$ in the bulk, i.e. for $\gamma_j \ge A + \kappa$, and $\max_j(\gamma_j - \gamma_{j-1}) = O(N^{-2/3})$ (the order $N^{-2/3}$ comes from the fact that the density ρ vanishes as a square root at the endpoints). The parameter $\kappa = \kappa(\alpha)$ is chosen such that $A + 2\kappa \le y_{L-K-1}$, which can be achieved since $L \ge \alpha N$

and y_{L-K-1} is close to γ_{L-K-1}. In the very last step we have absorbed the $N^{-2/3}$ error term into $(Nd(x))^{-1} \geq K^{-1} \gg N^{-2/3}$.

Finally, we prove (4.24). Since $|y_j - \gamma_j| \leq K^{\xi}/N$, it follows that $|x - y_j| \sim |x - \gamma_j|$ for $|x - \gamma_j| \geq K^{\xi}/N$. Thus we have

$$V_{\mathbf{y}}''(x) = V''(x) + \frac{2}{N}\sum_{j\notin I}\frac{1}{(x-y_j)^2} \geq \inf V'' + \frac{c}{N}\sum_{j\notin I}\frac{1}{(x-\gamma_j)^2} \geq \inf V'' + \frac{c}{d(x)}$$

with some positive constant c (depending only on α). In estimating the summation, we have used the fact that the sequence γ_k is regularly spaced with gaps of order $1/N$. This completes the proof of Lemma 4.5. □

Appendix B. Discrete Gagliardo–Nirenberg inequalities

Recall the integral formula for the quadratic form of the operator $(-\Delta)^{s/2}$ in $\mathbb{R}$ for any $s \in (0, 2)$:

$$\int_{\mathbb{R}} \phi(x)((-\Delta)^{s/2}\phi)(x)\,dx = C(s)\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|\phi(x)-\phi(y)|^2}{|x-y|^{1+s}}\,dx\,dy,$$

where $C(s)$ is an explicit positive constant, $C(1) = (2\pi)^{-1}$ and $\phi \in H^{s/4}(\mathbb{R})$. We have the following Gagliardo–Nirenberg type inequality in the critical case (see [48, (1.4)] with $n = 1$, $p = 4$)

$$\|\phi\|_4^4 \leq C\|\phi\|_2^2 \int_{\mathbb{R}} \phi(x)(\sqrt{-\Delta}\,\phi)(x)\,dx, \qquad \phi : \mathbb{R} \to \mathbb{R}. \tag{B.1}$$

We first give a slight generalization of this inequality:

Proposition B.1. *Let $p \in (2, \infty)$ and $s \in (1 - 2/p, 2)$. Then*

$$\|\phi\|_p \leq C_{p,s}\|\phi\|_2^{1-\frac{p-2}{sp}}\left[\int_{\mathbb{R}} \phi(x)((-\Delta)^{s/2}\phi)(x)\,dx\right]^{\frac{p-2}{2sp}} \tag{B.2}$$

with some constant $C_{p,s}$ with $\|\cdot\|_p = \|\cdot\|_{L^p(\mathbb{R})}$.

Proof. We follow the proof of [48, Theorem 2]. Setting $q = p/(p-1)$ and using the Hausdorff–Young and Hölder inequalities, for any $\lambda > 0$ and $\alpha > 1 - q/2$ we obtain

$$\begin{aligned}\|\phi\|_p &\leq C_p\|\widehat{\phi}\|_q \leq C_p\|\widehat{\phi}(\xi)(\lambda + |\xi|)^{\alpha/q}\|_2\|(\lambda+|\xi|)^{-\alpha/q}\|_{2q/(2-q)}\\ &\leq C_{p,\alpha}(\lambda^{\alpha/q}\|\phi\|_2 + \langle\phi, (-\Delta)^{\alpha/q}\phi)^{1/2}\rangle\lambda^{(1-\alpha)/q-1/2}\\ &\leq C_{p,\alpha}\|\phi\|_2^{1-\frac{2-q}{2\alpha}}\langle\phi, (-\Delta)^{\alpha/q}\phi)^{\frac{2-q}{4\alpha}},\end{aligned} \tag{B.3}$$

where in the last step we have chosen $\lambda = (\phi, |p|^{2\alpha/q}\rangle^{q/(2\alpha)}\|\phi\|^{-q/\alpha}$. We use $\langle\cdot,\cdot\rangle$ to denote the inner product in $L^2(\mathbb{R})$. Setting $s = 2\alpha/q$, we obtain (B.2). □

Now we derive the discrete version of this inequality.

Proposition B.2. *Let $p \in (2, \infty)$ and $s \in (1 - 2/p, 2)$. Then there exists a positive constant $C_{p,s}$ such that*

$$\|f\|_p \le C_{p,s} \|f\|_2^{1-\frac{p-2}{sp}} \left[\sum_{i \ne j \in \mathbb{Z}} \frac{|f_i - f_j|^2}{|i - j|^{1+s}} \right]^{\frac{p-2}{2sp}} \tag{B.4}$$

for any function $f : \mathbb{Z} \to \mathbb{R}$, where $\|f\|_p = \|f\|_{L^p(\mathbb{Z})} = (\sum_i |f_i|^p)^{1/p}$.

Proof. Given $f : \mathbb{Z} \to \mathbb{R}$, let $\phi : \mathbb{R} \to \mathbb{R}$ be its linear interpolation, i.e. $\phi(i) := f_i$ for $i \in \mathbb{Z}$ and

$$\phi(x) = f_i + (f_{i+1} - f_i)(x - i) = f_{i+1} - (f_{i+1} - f_i)(i + 1 - x), \qquad x \in [i, i+1]. \tag{B.5}$$

It is easy to see that

$$C_p^{-1} \|\phi\|_{L^p(\mathbb{R})} \le \|f\|_{L^p(\mathbb{Z})} \le C_p \|\phi\|_{L^p(\mathbb{R})}, \qquad 2 \le p \le \infty, \tag{B.6}$$

with some constant C_p. We claim that

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^{1+s}}\, dx\, dy \le C_s \sum_{i \ne j \in \mathbb{Z}} \frac{|f_i - f_j|^2}{|i - j|^{1+s}} \tag{B.7}$$

with some constant C_s; then (B.6) and (B.7) will yield (B.4) from (B.1).

To prove (B.7), we can write

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^{1+s}}\, dx\, dy = \sum_{i,j} \int_i^{i+1} \int_j^{j+1} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^{1+s}}\, dx\, dy. \tag{B.8}$$

Using the explicit formula (B.5), we first compute the $i = j$ terms in (B.8):

$$\begin{aligned} &\sum_i \int_i^{i+1} \int_i^{i+1} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^{1+s}}\, dx\, dy \\ &\qquad = \sum_i |f_i - f_{i+1}|^2 \int_i^{i+1} \int_i^{i+1} \frac{dx\, dy}{|x - y|^s} = C_s \sum_i \frac{|f_i - f_{i+1}|^2}{|i - (i+1)|^{1+s}} \end{aligned} \tag{B.9}$$

with some explicit C_s. Next we compute the terms $|i - j| = 1$ in (B.8). We assume $j = i - 1$, the terms $j = i + 1$ being analogous:

$$\begin{aligned} \sum_i \int_i^{i+1} \int_{i-1}^{i} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^{1+s}}\, dx\, dy &\le \sum_i (f_{i+1} - f_i)^2 \int_i^{i+1} \int_{i-1}^{i} \frac{(x-i)^2}{(x-y)^{1+s}}\, dx\, dy \\ &\quad + \sum_i (f_i - f_{i-1})^2 \int_i^{i+1} \int_{i-1}^{i} \frac{(i-y)^2}{(x-y)^{1+s}}\, dx\, dy, \end{aligned} \tag{B.10}$$

where we have used $\phi(x) = f_i + (f_{i+1} - f_i)(x - i)$ and $\phi(y) = f_i - (f_i - f_{i-1})(i - y)$. The above integrals are finite constants C_s, so we get

$$\sum_i \int_i^{i+1} \int_{i-1}^{i} \frac{|\phi(x) - \phi(y)|^2}{|x - y|^2}\, dx\, dy \le C_s \sum_i \frac{(f_{i+1} - f_i)^2}{(i + 1 - i)^{1+s}} + \frac{(f_i - f_{i-1})^2}{(i - (i-1))^{1+s}}.$$

Finally, for the terms $|i-j|\ge 2$, we can just replace $(x-y)^{1+s}$ by $(i-j)^{1+s}$ on the right hand side of (B.8) and use simple Schwarz inequalities to get

$$\int_i^{i+1}\int_j^{j+1}\frac{|\phi(x)-\phi(y)|^2}{|x-y|^{1+s}}\,dx\,dy \le C_s\frac{|f_i-f_j|^2+|f_{i+1}-f_i|^2+|f_{j+1}-f_j|^2}{|i-j|^{1+s}}.$$

After summing up we get

$$\begin{aligned}\sum_{|i-j|\ge 2}\int_i^{i+1}\int_j^{j+1}&\frac{|\phi(x)-\phi(y)|^2}{|x-y|^{1+s}}\,dx\,dy\\ &\le C_s\sum_{|i-j|\ge 2}\frac{|f_i-f_j|^2}{|i-j|^{1+s}}+C_s\sum_i\frac{|f_{i+1}-f_i|^2}{((i+1)-i)^{1+s}}.\end{aligned} \tag{B.11}$$

The estimates (B.9), (B.10) and (B.11) together yield (B.7). □

With two fixed parameters $a,b>0$, define

$$m(\xi):=|\xi|\cdot\mathbf{1}(|\xi|\le a)+b|\xi|,\qquad \xi\in\mathbb{R}. \tag{B.12}$$

We will consider the operator $T=m(\sqrt{-\Delta})$ defined by m being its Fourier multiplier, i.e.

$$\widehat{T\phi}(\xi)=m(\xi)\widehat{\phi}(\xi).$$

Proposition B.3. *We have*

$$\|\phi\|_4^4\le C\|\phi\|_2^2\langle\phi,m(\sqrt{-\Delta})\phi\rangle+\frac{C}{ab^3}\|\phi\|_\infty^4. \tag{B.13}$$

Proof. Let $\chi\in C_0^\infty(\mathbb{R})$ be a symmetric cutoff function such that $0\le\chi\le 1$, $\chi(\xi)=1$ for $|\xi|\le 1/2$ and $\chi(\xi)=0$ for $|\xi|\ge 1$. Set $\chi_a(\xi)=\chi(\xi/a)$. Split $\phi=\phi_1+\phi_2$ into low and high Fourier modes, via the Fourier transforms:

$$\phi=\phi_1+\phi_2,\qquad \widehat{\phi_1}(\xi):=\widehat{\phi}(\xi)\chi_a(\xi),\qquad \widehat{\phi_2}(\xi):=\widehat{\phi}(\xi)(1-\chi_a(\xi)).$$

First we estimate the contribution from ϕ_1. With $p=4$, $s=1$ in (B.2) we have

$$\begin{aligned}\|\phi_1\|_4&\le C\|\phi_1\|_2^{1/2}\left[\int_{\mathbb{R}}|\widehat{\phi_1}(\xi)|^2|\xi|\,d\xi\right]^{1/4}\le C\|\phi_1\|_2^{1/2}\langle\phi_1,m(\sqrt{-\Delta})\phi_1\rangle^{1/4}\\ &\le C\|\phi\|_2^{1/2}\langle\phi,m(\sqrt{-\Delta})\phi\rangle^{1/4},\end{aligned}$$

where we have used $|\xi|\le m(\xi)$ on the support of $\widehat{\phi_1}$ and in the last step we have used $|\widehat{\phi_1}|\le|\widehat{\phi}|$ pointwise.

For the contribution of ϕ_2, with any $\delta>0$ we have

$$\|\phi_2\|_4\le\|\phi_2\|_\infty^{1/4}\|\phi_2\|_3^{3/4}\le\delta^{-4}\|\phi_2\|_\infty+\delta^{4/3}\|\phi_2\|_3.$$

In the first term we use the Littlewood–Paley inequality $\|\phi_2\|_\infty \le C\|\phi\|_\infty$, where C depends only on the choice of χ but is independent of a. In the second term we use (B.2) with $s = 2/3$, $p = 3$:

$$\begin{aligned}\|\phi_2\|_3 &\le C\|\phi_2\|^{1/2}\Big[\int |\widehat{\phi}_2(\xi)|^2|\xi|^{2/3}\, d\xi\Big]^{1/4}\\ &\le Cb^{-1/4}a^{-1/12}\|\phi_2\|^{1/2}\langle\phi_2, m(\sqrt{-\Delta})\phi_2\rangle^{1/4},\end{aligned}$$

where in the second step we have used $|\xi|^{2/3} \le 2b^{-1}a^{-1/3}m(\xi)$ for all $|\xi| \ge a/2$, i.e. on the support of $\widehat{\phi}_2$. Using $|\widehat{\phi}_2| \le |\widehat{\phi}|$, we thus have

$$\|\phi_2\|_4 \le C\delta^{-4}\|\phi\|_\infty + C\delta^{4/3}b^{-1/4}a^{-1/12}\|\phi\|^{1/2}\langle\phi, m(\sqrt{-\Delta})\phi\rangle^{1/4}.$$

Choosing $\delta = b^{3/16}a^{1/16}$, we obtain (B.13). □

Finally, we derive a discrete version and a localized discrete version of Proposition B.3.

Proposition B.4. *Let $\{B_{ij} : i \ne j \in \mathbb{Z}\}$ be a bi-infinite matrix of nonnegative numbers with $B_{ij} = B_{ji}$.*

(i) [Global version] *Assume that for some positive constants a, b, r with $b \le r \le 1$, we have*

$$B_{ij} \ge b/|i-j|^2, \quad \forall i \ne j \in \mathbb{Z}, \tag{B.14}$$

$$B_{ij} \ge r/|i-j|^2, \quad \forall i, j \in \mathbb{Z} \text{ with } |i-j| \ge a^{-1}. \tag{B.15}$$

Then for any function $f : \mathbb{Z} \to \mathbb{R}$ we have

$$\|f\|_4^4 \le \frac{C}{r}\|f\|_2^2\sum_{i\ne j} B_{ij}|f_i - f_j|^2 + \frac{C}{ab^3}\|f\|_\infty^4. \tag{B.16}$$

(ii) [Local version] *Let $\mathcal{I} = [\![Z-L, Z+L]\!] \subset \mathbb{Z}$ be a subinterval of length $|\mathcal{I}| = 2L+1$ around $Z \in \mathbb{Z}$ and let $\widehat{\mathcal{I}} := [\![Z-(1+\tau)L, Z+(1+\tau)L]\!] \subset \mathbb{Z}$ be a slightly larger interval, where $\tau > 0$. Assume that for some positive constants a, b, r with $b \le r \le 1$, we have*

$$B_{ij} \ge b/|i-j|^2, \quad \forall i, j \in \widehat{\mathcal{I}}, \tag{B.17}$$

$$B_{ij} \ge r/|i-j|^2, \quad \forall i, j \in \widehat{\mathcal{I}} \text{ with } |i-j| \ge a^{-1}. \tag{B.18}$$

Then for any function $f : \mathbb{Z} \to \mathbb{R}$ with $\mathrm{supp}(f) \subset \mathcal{I}$ we have

$$\|f\|_4^4 \le C\|f\|_2^2\Big[\frac{1}{r}\sum_{i\ne j\in\mathcal{I}} B_{ij}|f_i - f_j|^2 + \frac{1}{L\tau}\|f\|_2^2\Big] + \frac{C}{ab^3}\|f\|_\infty^4. \tag{B.19}$$

Proof. Following the proof of Proposition B.2, for any $f : \mathbb{Z} \to \mathbb{R}$ we define its continuous extension ϕ by (B.5). Then the combination of (B.6) and (B.13) yields

$$\|f\|_4^4 \le C\|f\|_2^2 \langle \phi, m(\sqrt{-\Delta})\phi\rangle + \frac{C}{ab^3}\|f\|_\infty^4,$$

where m is given in (B.12) and a, b will be determined later. We compute

$$\langle \phi, m(\sqrt{-\Delta})\phi\rangle \le b\langle \phi, \sqrt{-\Delta}\,\phi\rangle + \langle \phi, \sqrt{-\Delta}\,\chi_{2a}^2(\sqrt{-\Delta})\,\phi\rangle, \tag{B.20}$$

where we have used $\mathbf{1}(|\xi| \le a) \le \chi_{2a}^2(\xi)$ by the definition of χ at the beginning of the proof of Proposition B.3. The first term is bounded by

$$\begin{aligned} b\langle \phi, \sqrt{-\Delta}\,\phi\rangle &= b\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|\phi(x)-\phi(y)|^2}{|x-y|^2}\,dx\,dy \le Cb\sum_{i<j}\frac{|f_i-f_j|^2}{|i-j|^2} \\ &\le \sum_{i<j} B_{ij}|f_i-f_j|^2 \end{aligned} \tag{B.21}$$

using (B.7) in the first estimate and (B.14) in the second one. For the second term in (B.20) we use the trivial arithmetic inequality

$$|\xi|\chi_{2a}^2(\xi) \le Q(\xi) \quad \text{with} \quad Q(\xi) := 100a(1-e^{-|\xi|/a}).$$

Thus

$$\langle \phi, \sqrt{-\Delta}\,\chi_{2a}^2(\sqrt{-\Delta})\phi\rangle \le \int_{\mathbb{R}} |\widehat{\phi}(\xi)|^2 Q(\xi)\,d\xi = 50\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{|\phi(x)-\phi(y)|^2}{(x-y)^2+a^{-2}}\,dx\,dy.$$

Mimicking the argument leading to (B.7), we can continue this estimate as

$$\langle \phi, \sqrt{-\Delta}\,\chi_{2a}^2(\sqrt{-\Delta})\phi\rangle \le C\sum_{i\ne j\in\mathbb{Z}} \frac{|f_i-f_j|^2}{|i-j|^2+a^{-2}} \le \frac{C}{r}\sum_{i\ne j\in\mathbb{Z}} B_{ij}|f_i-f_j|^2, \tag{B.22}$$

where we have used (B.15) in the last step. This completes the proof of (B.16).

The proof of (B.19) is very similar, just in the very last estimates of (B.21) and (B.22) we use the fact that f is supported in $\mathcal{I}$. For example in (B.21) we have

$$\begin{aligned} b\sum_{i<j}\frac{|f_i-f_j|^2}{|i-j|^2} &= b\sum_{i<j\in\widehat{\mathcal{I}}}\frac{|f_i-f_j|^2}{|i-j|^2} + 2b\sum_{i\in\mathcal{I}}|f_i|^2\sum_{j\notin\widehat{\mathcal{I}}}\frac{1}{|i-j|^2} \\ &\le \sum_{i<j} B_{ij}|f_i-f_j|^2 + \frac{2}{L\tau}\|f\|_2^2, \end{aligned}$$

and the estimate in (B.22) is analogous. □

Acknowledgments. L. Erdős was partially supported by SFB-TR 12 Grant of the German Research Council and by ERC Advanced Grant, RANMAT 338804.

H.-T. Yau was partially supported by NSF grants DMS-1307444 and Simons Investigator Award.

References

[1] Albeverio, S., Pastur, L., Shcherbina, M.: On the $1/n$ expansion for some unitary invariant ensembles of random matrices. Comm. Math. Phys. **224**, 271–305 (2001) Zbl 1038.82039 MR 1869000

[2] Anderson, G., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Stud. Adv. Math. 118, Cambridge Univ. Press (2009) Zbl 1184.15023 MR 2760897

[3] Bach, V., Moller, J.-S.: Correlation at low temperature. I. Exponential decay. J. Funct. Anal. **203**, 93–148 (2003) Zbl 1031.82003 MR 1996869

[4] Bekerman, F., Figalli, A., Guionnet, A.: Transport maps for β-matrix models and universality. Comm. Math. Phys. **338**, 589–619 (2015) MR 3351052

[5] Ben Arous, G., Bourgade, P.: Extreme gaps between eigenvalues of random matrices. Ann. Probab. **41**, 2648–2681 (2013) Zbl 1282.60008 MR 3112927

[6] Ben Arous, G., Péché, S.: Universality of local eigenvalue statistics for some sample covariance matrices. Comm. Pure Appl. Math. **58**, 1316–1357 (2005) Zbl 1075.62014 MR 2162782

[7] Bleher, P., Its, A.: Semiclassical asymptotics of orthogonal polynomials, Riemann–Hilbert problem, and universality in the matrix model. Ann. of Math. **150**, 185–266 (1999) Zbl 0956.42014 MR 1715324

[8] Bourgade, P., Erdős, L., Yau, H.-T.: Universality of general β-ensembles. Duke Math. J. **163**, 1127–1190 (2014) Zbl 1298.15040 MR 3192527

[9] Bourgade, P., Erdős, L., Yau, H.-T.: Bulk universality of general β-ensembles with non-convex potential. J. Math. Phys. **53**, 095221 (2012) Zbl 1278.82032 MR 2905803

[10] Bourgade, P., Erdős, L., Yau, H.-T.: Edge universality of beta ensembles. Comm. Math. Phys. **332**, 261–354 (2014) Zbl 1306.82010 MR 3253704

[11] Bourgade, P., Erdős, Yau, H.-T., Yin, J.: Fixed energy universality for generalized Wigner matrices. arXiv:1407.5606 (2014)

[12] Boutet de Monvel, A., Pastur, L., Shcherbina, M.: On the statistical mechanics approach in the random matrix theory. Integrated density of states. J. Statist. Phys. **79**, 585–611 (1995) Zbl 1081.82569 MR 1327898

[13] Caffarelli, L., Chan, C. H., Vasseur, A.: Regularity theory for parabolic nonlinear integral operators. J. Amer. Math. Soc. **24**, 849–889 (2011) Zbl 1223.35098 MR 2784330

[14] Cotar, C., Deuschel, J. D.: Decay of covariances, uniqueness of ergodic component and scaling limit for a class of $\nabla\phi$ systems with non-convex potential. Ann. Inst. H. Poincaré Probab. Statist. **48**, 819–853 (2012) Zbl 1247.60133 MR 2976565

[15] Cotar, C., Deuschel, J. D., Müller, S.: Strict convexity of the free energy for non-convex gradient models at moderate β. Comm. Math. Phys. **286**, 359–376 (2009) Zbl 1173.82010 MR 2470934

[16] Deift, P.: Orthogonal Polynomials and Random Matrices: a Riemann–Hilbert Approach. Courant Lecture Notes in Math. 3, Amer. Math. Soc., Providence, RI (1999) MR 1677884

[17] Deift, P., Gioev, D.: Universality in random matrix theory for orthogonal and symplectic ensembles. Int. Math. Res. Papers **2007**, no. 2, art. ID rpm004, 116 pp. Zbl 1136.82021 MR 2335245

[18] Deift, P., Gioev, D.: Random Matrix Theory: Invariant Ensembles and Universality. Courant Lecture Notes in Math. 18, Amer. Math. Soc., Providence, RI (2009) Zbl 1171.15023 MR 2514781

[19] Deift, P., Kriecherbauer, T., McLaughlin, K. T.-R., Venakides, S., Zhou, X.: Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. Comm. Pure Appl. Math. **52**, 1335–1425 (1999) Zbl 0944.42013 MR 1702716

[20] Deift, P., Kriecherbauer, T., McLaughlin, K. T.-R., Venakides, S., Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights. Comm. Pure Appl. Math. **52**, 1491–1552 (1999) Zbl 1026.42024 MR 1711036

[21] Deuschel, J. D.: The random walk representation for interacting diffusion processes. In: Interacting Stochastic Systems, Springer, 377–393 (2005) Zbl 1111.82049 MR 2118583

[22] Deuschel, J.-D., Giacomin, G., Ioffe, D.: Large deviations and concentration properties for $\nabla\varphi$ interface models. Probab. Theor. Related Fields **117**, 49–111 (2000) Zbl 0988.82018 MR 1759509

[23] Dumitriu, I., Edelman, A.: Matrix models for beta ensembles. J. Math. Phys. **43**, 5830–5847 (2002) Zbl 1060.82020 MR 1936554

[24] Dyson, F. J.: A Brownian-motion model for the eigenvalues of a random matrix. J. Math. Phys. **3**, 1191–1198 (1962) Zbl 0111.32703 MR 0148397

[25] Erdős, L., Knowles, A., Yau, H.-T., Yin, J.: Spectral statistics of Erdős–Rényi graphs II: Eigenvalue spacing and the extreme eigenvalues. Comm. Math. Phys. **314**, 587–640 (2012) Zbl 1251.05162 MR 2964770

[26] Erdős, L., Péché, S., Ramirez, J., Schlein, B., Yau, H.-T.: Bulk universality for Wigner matrices. Comm. Pure Appl. Math. **63**, 895–925 (2010) Zbl 1216.15025 MR 2662426

[27] Erdős, L., Ramírez, J., Schlein, B., Tao, T., Vu, V. and Yau, H.-T.: Bulk universality for Wigner Hermitian matrices with subexponential decay. Math. Res. Lett. **17**, 667–674 (2010) Zbl 1277.15027 MR 2661171

[28] Erdős, L., Ramírez, J.-A., Schlein, B., Yau, H.-T.: Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. Electron. J. Probab. **15** (2010), 526–603 Zbl 1225.15034 MR 2639734

[29] Erdős, L., Schlein, B., Yau, H.-T.: Universality of random matrices and local relaxation flow. Invent. Math. **185**, 75–119 (2011) Zbl 1225.15033 MR 2810797

[30] Erdős, L., Schlein, B., Yau, H.-T., Yin, J.: The local relaxation flow approach to universality of the local statistics for random matrices. Ann. Inst. H. Poincaré Probab. Statist. **48**, 1–46 (2012) Zbl 1285.82029 MR 2919197

[31] Erdős, L., Yau, H.-T.: Universality of local spectral statistics of random matrices. Bull. Amer. Math. Soc. **49**, 377–414 (2012) Zbl 1263.15032 MR 2917064

[32] Erdős, L., Yau, H.-T.: A comment on the Wigner–Dyson–Mehta bulk universality conjecture for Wigner matrices. Electron. J. Probab. **17**, no. 28, 5 pp. (2012) Zbl 1245.15038 MR 2915664

[33] Erdős, L., Yau, H.-T., Yin, J.: Bulk universality for generalized Wigner matrices. Probab. Theory Related Fields, **154**, 341–407 (2012) Zbl 1277.15026 MR 2981427

[34] Erdős, L., Yau, H.-T., Yin, J.: Universality for generalized Wigner matrices with Bernoulli distribution. J. Combinatorics **1**, 15–81 (2011) Zbl 1235.15029 MR 2847916

[35] Erdős, L., Yau, H.-T., Yin, J.: Rigidity of eigenvalues of generalized Wigner matrices. Adv. Math. **229**, 1435–1515 (2012) Zbl 1238.15017 MR 2871147

[36] Felsinger, M., Kassmann, M.: Local regularity for parabolic nonlocal operators. Comm. Partial Differential Equations **38**, 1539–1573 (2013) Zbl 1277.35090 MR 3169755

[37] Fokas, A. S., Its, A. R., Kitaev, A. V.: The isomonodromy approach to matrix models in 2D quantum gravity. Comm. Math. Phys. **147**, 395–430 (1992) Zbl 0760.35051 MR 1174420

[38] Forrester, P. J.: Log-Gases and Random Matrices. London Math. Soc. Monogr. Ser. 34, Princeton Univ. Press, Princeton, NJ (2010) Zbl 1217.82003 MR 2641363

[39] Giacomin, G., Olla, S., Spohn, H.: Equilibrium fluctuations for $\nabla\varphi$ interface model. Ann. Probab. **29**, 1138–1172 (2001) Zbl 1017.60100 MR 1872740

[40] Gustavsson, J.: Gaussian fluctuations of eigenvalues in the GUE. Ann. Inst. H. Poincaré Probab. Statist. **41**, 151–178 (2005) Zbl 1073.60020 MR 2124079

[41] Helffer, B.: Semiclassical Analysis, Witten Laplacians and Statistical Mechanics. Ser. Partial Differential Equations Appl. 1, World Sci. (2002) Zbl 1046.82001 MR 1936110

[42] Helffer, B., Sjöstrand, J.: On the correlation for Kac-like models in the convex case. J. Statist. Phys. **74**, 349–409 (1994) Zbl 0946.35508 MR 1257821

[43] Johansson, K.: Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. Comm. Math. Phys. **215**, 683–705 (2001) Zbl 01631876 MR 1810949

[44] Knowles, A., Yin, J.: Eigenvector distribution of Wigner matrices. Probab. Theory Related Fields **155**, 543–582 (2013) Zbl 1268.15033 MR 3034787

[45] Kriecherbauer, T., Shcherbina, M.: Fluctuations of eigenvalues of matrix models and their applications. arXiv:1003.6121 (2010)

[46] Mehta, M. L.: Random Matrices. Academic Press, New York (1991) Zbl 0780.60014 MR 1083764

[47] Naddaf, A., Spencer, T.: On homogenization and scaling limit of some gradient perturbations of a massless free field. Comm. Math. Phys. **183**, 55–84 (1997) MR 1461951

[48] Ogawa, T., Ozawa, T.: Trudinger type inequalities and uniqueness of weak solutions for the nonlinear Schrödinger mixed problem. J. Math. Anal. Appl. **155**, 531–540 (1991) Zbl 0733.35095 MR 1097298

[49] Pastur, L., Shcherbina, M.: Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles. J. Statist. Phys. **86**, 109–147 (1997) Zbl 0916.15009 MR 1435193

[50] Pastur, L., Shcherbina M.: Bulk universality and related properties of Hermitian matrix models. J. Statist. Phys. **130**, 205–250 (2008) Zbl 1136.15015 MR 2375744

[51] Shcherbina, M.: Orthogonal and symplectic matrix models: universality and other properties. Comm. Math. Phys. **307**, 761–790 (2011) Zbl 1232.15027 MR 2842965

[52] Shcherbina, M.: Change of variables as a method to study general β-models: Bulk universality. J. Math. Phys. **55**, 043504 (2014) Zbl 1296.82032

[53] Tao, T.: The asymptotic distribution of a single eigenvalue gap of a Wigner matrix. Probab. Theory Related Fields **157**, 81–106 (2013) Zbl 1280.15023 MR 310184

[54] Tao, T., Vu, V.: Random matrices: Universality of the local eigenvalue statistics. Acta Math. **206**, 127–204 (2011) Zbl 1217.15043 MR 27846651

[55] Tracy, C., Widom, H.: Level spacing distributions and the Airy kernel. Comm. Math. Phys. **159**, 151–174 (1994) Zbl 0789.35152 MR 1257246

[56] Tracy, C., Widom, H.: On orthogonal and symplectic matrix ensembles. Comm. Math. Phys. **177**, 727–754 (1996) Zbl 0851.60101 MR 1385083

[57] Valkó, B., Virág, B.: Continuum limits of random matrices and the Brownian carousel. Invent. Math. **177**, 463–508 (2009) Zbl 1204.60012 MR 2534097

[58] Valkó, B., Virág, B.: Large gaps between random eigenvalues. Ann. Probab. **38**, 1263–1279 (2010) Zbl 1223.60009 MR 2674999

[59] Widom H.: On the relation between orthogonal, symplectic and unitary matrix ensembles. J. Statist. Phys. **94**, 347–363 (1999) Zbl 0935.60090 MR 1675356

[60] Wigner, E.: Characteristic vectors of bordered matrices with infinite dimensions. Ann. of Math. **62**, 548–564 (1955) Zbl 0067.08403 MR 0077805