Multilevel Dyson Brownian motions via Jack polynomials

Probab. Theory Relat. Fields (2015) 163:413–463
DOI 10.1007/s00440-014-0596-2

Multilevel Dyson Brownian motions via Jack polynomials

Vadim Gorin · Mykhaylo Shkolnikov

Received: 25 February 2014 / Revised: 4 November 2014 / Published online: 27 November 2014

Abstract We introduce multilevel versions of Dyson Brownian motions of arbitrary parameter $\beta > 0$, generalizing the interlacing reflected Brownian motions of Warren for $\beta = 2$. Such processes unify β corners processes and Dyson Brownian motions in a single object. Our approach is based on the approximation by certain multilevel discrete Markov chains of independent interest, which are defined by means of Jack symmetric polynomials. In particular, this approach allows to show that the levels in a multilevel Dyson Brownian motion are intertwined (at least for $\beta \geq 1$) and to give the corresponding link explicitly.

Mathematics Subject Classification 60B20 · 82C22 · 60H10 · 05E05

Contents

V. Gorin
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

V. Gorin
Institute for Information Transmission Problems of Russian Academy of Sciences, Moscow, Russia
e-mail: vadicgor@gmail.com

M. Shkolnikov (✉)
Department of Statistics, University of California, Berkeley, CA, USA
e-mail: mshkolni@gmail.com

1 Introduction

1.1 Preface

The Hermite general $\beta > 0$ ensemble of rank N is a probability distribution on the set of N tuples of reals $z_1 < z_2 < \cdots < z_N$ with density proportional to

$$\prod_{1 \le i < j \le N} (z_j - z_i)^\beta \prod_{i=1}^{N} \exp\left(-\frac{z_i^2}{2}\right). \tag{1.1}$$

When $\beta = 2$, that density describes the joint distribution of the eigenvalues of a random Hermitian $N \times N$ matrix M, whose diagonal entries are i.i.d. real standard normal random variables, while real and imaginary parts of its entries above the diagonal are i.i.d. normal random variables of variance $1/2$. The law of such a random matrix is referred to as the Gaussian Unitary Ensemble (GUE) (see e.g. [3,20,32]) and it has attracted much attention in the mathematical physics literature following the seminal work of Wigner in the 50s. Similarly, the case of $\beta = 1$ describes the joint distribution of eigenvalues of a real symmetric matrix sampled from the Gaussian Orthogonal Ensemble (GOE) and the case $\beta = 4$ corresponds to the Gaussian Symplectic Ensemble (GSE) (see e.g. [3,20,32] for the detailed definitions).

It is convenient to view the realizations of (1.1) as point processes on the real line and there are two well-known ways of adding a second dimension to that picture. The first one is to consider an N-dimensional diffusion known as Dyson Brownian motion (see [32, Chapter 9], [3, Section 4.3] and the references therein) which is the unique strong solution of the system of stochastic differential equations

$$\mathrm{d}X_i(t) = \frac{\beta}{2} \sum_{j \ne i} \frac{1}{X_i(t) - X_j(t)} \mathrm{d}t + \mathrm{d}W_i(t), \quad i = 1, 2, \ldots, N \tag{1.2}$$

with $W_1, W_2, \ldots, W_N$ being independent standard Brownian motions. If one solves (1.2) with zero initial condition, then the distribution of the solution at time 1 is given

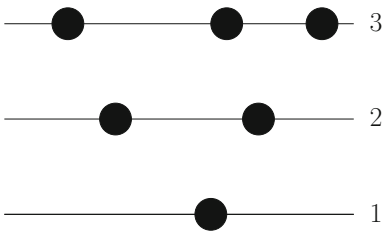

Fig. 1 Interlacing particles arising from eigenvalues of corners of a 3×3 matrix. Row number k in the picture corresponds to eigenvalues of the $k \times k$ corner

by (1.1). When $\beta = 2$ and one starts with zero initial condition, the diffusion (1.2) has two probabilistic interpretations: it can be either viewed as a system of N independent standard Brownian motions conditioned never to collide via a suitable Doob's h-transform; or it can be regarded as the evolution of the eigenvalues of a Hermitian random matrix whose elements evolve as (independent) Brownian motions.

An alternative way of adding a second dimension to the ensemble in (1.1) involves the so-called corner processes. For $\beta = 2$ take a $N \times N$ GUE matrix M and let $x_1^N \le x_2^N \le \cdots \le x_N^N$ be its ordered eigenvalues. More generally for every $1 \le k \le N$ let $x_1^k \le x_2^k \le \cdots \le x_k^k$ be the eigenvalues of the top-left $k \times k$ submatrix ("corner") of M. It is well-known that the eigenvalues *interlace* in the sense that $x_i^k \le x_i^{k-1} \le x_{i+1}^k$ for $i = 1, \ldots, k-1$ (see Fig. 1 for a schematic illustration of the eigenvalues).

The joint distribution of x_i^k, $1 \le i \le k \le N$ is known as the GUE-corners process (some authors also use the name "GUE-minors process") and its study was initiated in [4] and [26]. The GUE-corners process is uniquely characterized by two properties: its projection to the set of particles $x_1^N, x_2^N, \ldots, x_N^N$ is given by (1.1) with $\beta = 2$ (normalized to a probability density), and the conditional distribution of x_i^k, $1 \le i \le k \le N-1$ given $x_1^N, x_2^N, \ldots, x_N^N$ is *uniform* on the polytope defined by the interlacing conditions above, see [4,21]. Due to the combination of the gaussianity and uniformity embedded into its definition, the GUE–corners process appears as a universal scaling limit for a number of 2d models of statistical mechanics, see [22–24,26,38].

Similarly, one can construct corners processes for $\beta = 1$ and $\beta = 4$, see e.g. [33]. Extrapolating the resulting formulas for the joint density of eigenvalues to general values of $\beta > 0$ one arrives at the following definition.

Definition 1.1 The Hermite β corners process of variance $t > 0$ is the unique probability distribution on the set of reals x_i^k, $1 \le i \le k \le N$ subject to the interlacing conditions $x_i^k \le x_i^{k-1} \le x_{i+1}^k$ whose density is proportional to

$$\prod_{i<j} \left(x_j^N - x_i^N\right) \prod_{i=1}^{N} \exp\left(-\frac{(x_i^N)^2}{2t}\right) \prod_{k=1}^{N-1} \prod_{1 \le i < j \le k} \left(x_j^k - x_i^k\right)^{2-\beta} \prod_{a=1}^{k} \prod_{b=1}^{k+1} \left|x_a^k - x_b^{k+1}\right|^{\beta/2-1}. \tag{1.3}$$

The fact that the projection of the Hermite β corners process of variance 1 onto level k (that is, on the coordinates $x_1^k, x_2^k, \ldots, x_k^k$) is given by the corresponding Hermite β ensemble of (1.1) can be deduced from the Dixon-Anderson integration formula

(see [2,17]), which was studied before in the context of *Selberg integrals* (see [44], [20, Chapter 4]). One particular case of the Selberg integral is the evaluation of the normalizing constant for the probability density of the Hermite β ensemble of (1.1). We provide more details in this direction in Sect. 2.2.

The ultimate goal of the present article is to *combine* Dyson Brownian motions and corner processes in a single picture. In other words, we aim to introduce a relatively simple diffusion on interlacing particle configurations whose projection on a fixed level is given by Dyson Brownian motion of (1.2), while its fixed time distributions are given by the Hermite β corners processes of Definition 1.1.

One would think that a natural way to do this (at least for $\beta = 1, 2, 4$) is to consider an $N \times N$ matrix of suitable Brownian motions and to project it onto the (interlacing) set of eigenvalues of the matrix and its top-left $k \times k$ corners, thus generalizing the original construction of Dyson. However, the resulting stochastic process ends up being quite nasty even in the case $\beta = 2$. It is shown in [1] that already for $N = 3$ (and at least *some* initial conditions) the projection is *not* a Markov process. When one considers only two adjacent levels (that is, the projection onto $x_1^N, x_2^N, \dots, x_N^N; x_1^{N-1}, x_2^{N-1}, \dots, x_{N-1}^{N-1}$), then it can be proven (see [1]) that the projection *is* Markovian, but the corresponding SDE is very complicated.

An alternative elegant solution for the case $\beta = 2$ was given by Warren [46]. Consider the process $(Y_i^k : 1 \le i \le k \le N)$ defined through the following inductive procedure: Y_1^1 is a standard Brownian motion with zero initial condition; given Y_1^1, the processes Y_1^2 and Y_2^2 are constructed as independent standard Brownian motions started at zero and *reflected* on the trajectory of Y_1^1 in such a way that $Y_1^2(t) \le Y_1^1(t) \le Y_2^2(t)$ holds for all $t \ge 0$. More generally, having constructed the processes on the first k levels (that is, Y_i^m, $1 \le i \le m \le k$) one defines Y_i^{k+1} as an independent standard Brownian motion started at 0 and reflected on the trajectories of Y_{i-1}^k and Y_i^k in such a way that $Y_{i-1}^{k-1}(t) \le Y_i^k(t) \le Y_i^{k-1}(t)$ remains true for all $t \ge 0$ (see [46] and also [24] for more details). Warren shows that the projection of the dynamics on a level k (that is, on $Y_1^k, Y_2^k, \dots, Y_k^k$) is given by a k-dimensional Dyson Brownian Motion of (1.2) with $\beta = 2$, and that the fixed time distributions of the process $(Y_i^k : 1 \le i \le k \le N)$ are given by the Hermite β corners processes of Definition 1.1 with $\beta = 2$.

Our aim is to construct a generalization of the Warren process for general values of β. In other words, we want to answer the question "What is the general β analogue of the reflected interlacing Brownian Motions of [46]?".

1.2 Our results

Our approach to the construction of the desired general β multilevel stochastic process is based on discrete space approximation. In [24] we proved that the reflected interlacing Brownian motions of [46] can be obtained as a diffusive scaling limit for a class of stochastic dynamics on *discrete* interlacing particle configurations. The latter dynamics are constructed from independent random walks by imposing the *local* block/push interactions between particles to preserve the interlacing conditions. The special cases of such processes arise naturally in the study of two-dimensional statistical mechanics systems such as random stepped surfaces and various types of tilings (cf. [6,9,10,34]).

In Sect. 2 we introduce a deformation $X^{multi}_{disc}(t)$ of these processes depending on a positive parameter θ (which is omitted from the notation, and with $\theta = 1$ corresponding to the previously known case). The resulting discrete space dynamics is an intriguing interacting particle system with *global* interactions whose state space is given by interlacing particle configurations with integer coordinates. Computer simulations of this dynamics for $\theta = 1/2$ and $\theta = 2$ can be found at [25].

We further study the diffusive limit of $X^{multi}_{disc}(s)$ under the rescaling of time $s = \varepsilon^{-1}t$ and $\varepsilon^{-1/2}$ scaling of space as $\varepsilon \downarrow 0$. Our first result is that for any fixed $\theta > 0$, the rescaled processes are *tight* as $\varepsilon \downarrow 0$, see Theorem 5.1 for the exact statement. The continuous time, continuous space processes $Y^{mu}(t)$, defined as *subsequential limits* $\varepsilon \downarrow 0$ of the family $X^{multi}_{disc}(s)$, are our main heros and we prove a variety of results about them for different values of θ.

(1) For any $\theta \geq 2$ we show in Theorem 5.2 that $Y^{mu}(t)$ satisfies the system of SDEs (1.4) below.
(2) For any $\theta \geq 1/2$ and any $1 \leq k \leq N$ we show in Theorem 5.3 that if $Y^{mu}(t)$ is started from a θ–*Gibbs* initial condition (zero initial condition is a particular case), then the k–dimensional restriction of the $N(N-1)/2$–dimensional process $Y^{mu}(t)$ to the level k is a 2θ–Dyson Brownian motion, that is, the vector $(Y^{mu}(t)^k_1, Y^{mu}(t)^k_2, \ldots, Y^{mu}(t)^k_k)$ solves (1.2) with $\beta = 2\theta$ and suitable independent standard Brownian motions $W_1(t), W_2(t), \ldots, W_k(t)$.
(3) For any $\theta > 0$ we show that if $Y^{mu}(t)$ is started from zero initial condition, then its distribution at time t is the Hermite 2θ corners process of variance t, that is, the corresponding probability density is proportional to (1.3) with $\beta = 2\theta$. In fact, we prove a more general statement, see Theorem 5.3 and Corollary 5.4.
(4) For $\theta = 1$ the results of [24] yield that $Y^{mu}(t)$ is the collection of reflected interlacing Brownian motions of [46].

The above results are complemented by the following uniqueness theorem for the system of SDEs (1.4). In particular, it implies that for $\theta \geq 2$ all the subsequential limits $Y^{mu}(t)$ of $X^{multi}_{disc}(s)$ as $\varepsilon \downarrow 0$ are the same.

Theorem 1.2 (Theorem 4.1) *For any $N \in \mathbb{N}$ and $\theta > 1$ the system of SDEs*

$$\mathrm{d}Y^k_i(t) = \left(\sum_{m \neq i} \frac{1-\theta}{Y^k_i(t) - Y^k_m(t)} - \sum_{m=1}^{k-1} \frac{1-\theta}{Y^k_i(t) - Y^{k-1}_m(t)} \right) \mathrm{d}t + \mathrm{d}W^k_i(t), \quad 1 \leq i \leq k \leq N, \tag{1.4}$$

where W^k_i, $1 \leq i \leq k \leq N$ are independent standard Brownian motions, possesses a unique weak solution taking values in the cone

$$\overline{\mathcal{G}^N} = \left\{ y = (y^k_i)_{1 \leq i \leq k \leq N} \in \mathbb{R}^{N(N+1)/2} : y^{k-1}_{i-1} \leq y^k_i \leq y^{k-1}_i \right\} \tag{1.5}$$

for any initial condition $Y(0)$ in the interior $\mathcal{G}^N$ of $\overline{\mathcal{G}^N}$.

It would be interesting to extend all the above results to general $\theta > 0$. We believe (but we do not have a proof) that the identification of $Y^{mu}(t)$ with a solution of (1.4)

is valid for all $\theta > 1$ and that the identification of the projection of $Y^{mu}(t)$ onto the N–th level with a $\beta = 2\theta$–Dyson Brownian motion is valid for any $\theta > 0$. On the other hand, $Y^{mu}(t)$ cannot be a solution to (1.4) for $\theta \leq 1$. Indeed, we know that when $\theta = 1$ the process $Y^{mu}(t)$ is a collection of reflected interlacing Brownian motions which hints that one should introduce additional local time terms in (1.4). In addition, the interpretation of the solution to (1.4) as a generalization of the one-dimensional Bessel process to a process in the Gelfand–Tseitlin cone suggests that the corresponding process for $\theta < 1$ is no longer a semimartingale and should be defined and studied along the lines of [42, Chapter XI, Exercise (1.26)].

1.3 Our methods

Our approach to the construction and study of the discrete approximating process $X^{multi}_{disc}(s)$ is related to *Jack symmetric polynomials*. Recall that Jack polynomials $J_\lambda(x_1, x_2, \ldots, x_N; \theta)$, indexed by Young diagrams λ and a positive parameter θ, are eigenfunctions of the Sekiguchi differential operators ([43], [31, Chapter VI, Section 10], [20, Chapter 12])

$$D(u;\theta) = \frac{1}{\prod_{i<j}(x_i - x_j)} \det\left[x_i^{N-j} \left(x_i \frac{\partial}{\partial x_i} + (N-j)\theta + u \right) \right]_{i,j=1,2,\ldots,N}.$$

One can also define $J_\lambda(x_1, x_2, \ldots, x_N; \theta)$ as limits of *Macdonald polynomials* $P_\lambda(\cdot; q, t)$ as $q, t \to 1$ in such a way that $t = q^\theta$ (see [31]). For the special values $\theta = 1/2, 1, 2$ these polynomials are spherical functions of Gelfand pairs $O(N) \subset U(N)$, $U(N) \subset U(N) \times U(N)$, $U(2N) \subset Sp(N)$, respectively, and are also known as *Zonal polynomials* (see e.g. [31, Chapter 7] and the references therein). It is known that spherical functions of compact type (corresponding to the above Gelfand pairs) degenerate to the spherical functions of Euclidian type, which in our case are related to real symmetric, complex Hermitian and quaternionic Hermitian matrices, respectively (see e.g. [35, Section 4] and the references therein). In particular, in the case $\theta = 1$ this is a manifestation of the fact that the tangent space to the unitary group $U(N)$ at identity can be identified with the set of Hermitian matrices. Due to all these facts it comes at no surprise that Hermite β ensembles can be obtained as limits of discrete probabilistic structures related to Jack polynomials with parameter $\theta = \beta/2$.

On the discrete level our construction of the multilevel stochastic dynamics is based on a procedure introduced by Diaconis and Fill [16], which has recently been used extensively in the study of Markov chains on interlacing particle configurations (see e.g. [6,7,9–11]). The idea is to use *commuting Markov operators* and conditional independence to construct a multilevel Markov chain with given single level marginals. In our case these operators can be written in terms of Jack polynomials. In the limit the commutation relation we use turns into the following statement, which might be of independent interest.

Let $P_N(t;\beta)$ denote the Markov transition operators of the N-dimensional Dyson Brownian Motion of (1.2) and let $L^N_{N-1}(\beta)$ denote the Markov transition operator corresponding to conditioning the $(N-1)$-st level (that is, $x_1^{N-1}, x_2^{N-1}, \ldots, x_{N-1}^{N-1}$) on the N-th level (that is, $x_1^N, x_2^N, \ldots, x_N^N$) in the Hermite β corners process.

Proposition 1.3 (Corollary of Theorem 5.3) *For any $\beta \geq 1$ the links $L^N_{N-1}(\beta)$ given by the stochastic transition kernels*

$$\frac{\Gamma(N\beta/2)}{\Gamma(\beta/2)^N} \prod_{1\leq i<m\leq N-1} \left(x_m^{N-1} - x_i^{N-1}\right) \prod_{i=1}^{N-1} \prod_{j=1}^{N} \left|x_j^N - x_i^{N-1}\right|^{\beta/2-1} \prod_{1\leq j<n\leq N} \left(x_n^N - x_j^N\right)^{1-\beta} \tag{1.6}$$

intertwine the semigroups $P_N(t;\beta)$ and $P_{N-1}(t;\beta)$ in the sense that

$$L^N_{N-1}(\beta) P_N(t;\beta) = P_{N-1}(t;\beta) L^N_{N-1}(\beta), \quad t \geq 0. \tag{1.7}$$

The latter phenomenon can be subsumed into a general theory of intertwinings for diffusions, which for example also includes the findings in [45] and [39].

Due to the presence of singular drift terms neither existence, nor uniqueness of the solution of (1.4) is straightforward. When dealing with systems of SDEs with singular drift terms one typically shows the existence and uniqueness of strong solutions by truncating the singularity first (thus, obtaining a well-behaved system of SDEs) and by proving afterwards that the solution cannot reach the singularity in finite time using suitable Lyapunov functions, see e.g. [3, proof of Proposition 4.3.5]. However, for $1 < \theta < 2$ the solutions of (1.4) *do reach* some of the singularities. A similar phenomenon occurs in the case of the β–Dyson Brownian motion (1.2) with $0 < \beta < 1$, for which the existence and uniqueness theorem was established in [14] using the theory of multivalued SDEs; however, in the multilevel setting we lack a certain monotonicity property which plays a crucial role in [14]. In addition, due to the intrinsic asymmetry built into the drift terms the solution of (1.4) seems to be beyond the scope of the processes that can be constructed using Dirichlet forms (see e.g. [40] for Dirichlet form constructions of symmetric diffusions with a singular drift at the boundary of their domain and for the limitations of that method). Instead, by localizing in time, using appropriate Lyapunov functions, and applying the Girsanov Theorem, we are able to reduce (1.4) to a number of non-interacting Bessel processes, whose existence and uniqueness is well-known. This approach has an additional advantage over Dirichlet form type constructions, since it allows to establish convergence to the solution of (1.4) via martingale problem techniques, which is how our proof of Theorem 5.2 goes.

Note also that for $\theta = 1$ ($\beta = 2$) the interactions in the definition of $X^{multi}_{disc}(s)$ become local. We have studied the convergence of such dynamics to the process of Warren in [24], with the proof being based on the continuity of a suitable Skorokhod reflection map. For general values of $\theta > 0$ neither the discrete dynamics, nor the continuous dynamics can be obtained as the image of an explicitly known process under the Skorokhod reflection map of [24].

1.4 Further developments and open problems

It would be interesting to study the asymptotic behavior of both the discrete and the continuous dynamics as the number of levels N goes to infinity. There are at least two groups of questions here.

The *global* fluctuations of Dyson Brownian motions as $N \to \infty$ are known to be Gaussian (see [3, Section 4.3]); moreover, the limiting covariance structure can be described by the Gaussian Free Field (see [5,12]). In addition, the asymptotic fluctuations of the Hermite β corners processes of Definition 1.1 are also Gaussian and can be described via the Gaussian Free Field (cf. [12]). This raises the question of whether the 3-dimensional global fluctuations of the solution to (1.4) are also asymptotically (as $N \to \infty$) Gaussian and how the limiting covariance structure might look like. A partial result in this direction was obtained for $\beta = 2$ in [9].

The *edge* fluctuations (that is, the fluctuations of the rightmost particle as $N \to \infty$) in the Hermite β ensemble given by (1.1) can be described via the β-Tracy–Widom distribution (see [41]). Moreover, in the present article we link the Hermite β ensemble to a certain discrete interacting particle system. This suggests that one might find the β-Tracy–Widom distribution in the limit of the edge fluctuations of that interacting particle system or its simplified versions.

2 Discrete space dynamics via Jack polynomials

2.1 Preliminaries on Jack polynomials

In this section we collect certain facts about Jack symmetric polynomials. A reader familiar with these polynomials can proceed to Sect. 2.2. Our notations generally follow the ones in [31].

In what follows Λ^N is the algebra of symmetric polynomials in N variables. In addition, we let Λ be the algebra of symmetric polynomials in countably many variables, that is, of *symmetric functions*. An element of Λ is a formal symmetric power series of bounded degree in the variables $x_1, x_2, \dots$. One way to view Λ is as an algebra of polynomials in the Newton power sums $p_k = \sum_i (x_i)^k$. There exists a unique canonical projection $\pi_N : \Lambda \to \Lambda_N$, which sets all variables except for $x_1, x_2, \dots, x_N$ to zero (see [31, Chapter 1, Section 2] for more details).

A partition of size n, or a *Young diagram* with n boxes, is a sequence of non-negative integers $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ such that $\sum_i \lambda_i = n$. $|\lambda|$ stands for the number of boxes in λ and $\ell(\lambda)$ is the number of non-empty rows in λ (that is, the number of non-zero sequence elements λ_i in λ). Let $\mathbb{Y}$ denote the set of all Young diagrams, and $\mathbb{Y}^N$ the set of all Young diagrams λ with at most N rows (that is, such that $\lambda_{N+1} = 0$). Typically, we will use the symbols λ, μ for Young diagrams. We adopt the convention that the empty Young diagram $\emptyset$ with $|\emptyset| = 0$ also belongs to $\mathbb{Y}$ and $\mathbb{Y}^N$. For a box $\square = (i, j)$ of a Young diagram λ (that is, a pair (i, j) such that $\lambda_i \geq j$), $a(i, j; \lambda)$ and $l(i, j; \lambda)$ are its arm and leg lengths:

$$a(i, j; \lambda) = \lambda_i - j, \quad l(i, j; \lambda) = \lambda'_j - i,$$

where λ'_j is the row length in the *transposed* diagram λ' defined by

$$\lambda'_j = |\{i : \lambda_i \geq j\}|.$$

Further, $a'(i,j)$, $l'(i,j)$ stand for the co-arm and the co-leg lengths, which do not depend on λ

$$a'(i,j) = j-1, \quad l'(i,j) = i-1.$$

When it is clear from the context which Young diagram is used, we omit it from the notation and write simply $a(i,j)$ (or $a(\square)$) and $l(i,j)$ (or $l(\square)$).

We write $J_\lambda(\cdot;\theta)$ for *Jack polynomials*, which are indexed by Young diagrams λ and positive reals θ. Many facts about these polynomials can be found in [31, Chapter VI, Section 10]. Note however that in that book Macdonald uses the parameter α given by our θ^{-1}. We use θ, following [27]. J_λ can be viewed either as an element of the algebra Λ of symmetric functions in countably many variables $x_1, x_2, \ldots$, or (specializing all but finitely many variables to zeros) as a symmetric polynomial in $x_1, x_2, \ldots, x_N$ from the algebra Λ^N. In both interpretations the leading term of J_λ is given by $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_{\ell(\lambda)}^{\lambda_{\ell(\lambda)}}$. When N is finite, the polynomials $J_\lambda(x_1, \ldots, x_N; \theta)$ are known to be the eigenfunctions of the Sekiguchi differential operator:

$$\frac{1}{\prod_{i<j}(x_i - x_j)} \det\left[x_i^{N-j}\left(x_i \frac{\partial}{\partial x_i} + (N-j)\theta + u\right)\right]_{i,j=1,2,\ldots,N} J_\lambda(x_1, \ldots, x_N; \theta)$$
$$= \left(\prod_{i=1}^{N} \left(\lambda_i + (N-i)\theta + u\right)\right) J_\lambda(x_1, \ldots, x_N; \theta). \tag{2.1}$$

The eigenrelation (2.1) can be taken as a definition for the Jack polynomials. We also need dual polynomials $\widetilde{J}_\lambda$ which differ from J_λ by an *explicit* multiplicative constant:

$$\widetilde{J}_\lambda = J_\lambda \cdot \prod_{\square \in \lambda} \frac{a(\square) + \theta\, l(\square) + \theta}{a(\square) + \theta\, l(\square) + 1}. \tag{2.2}$$

Next, we recall the definition of *skew Jack polynomials* $J_{\lambda/\mu}$. Take two infinite sets of variables x and y, and consider a Jack polynomial $J_\lambda(x, y; \theta)$. The latter is, in particular, a symmetric polynomial in the x variables. The coefficients $J_{\lambda/\mu}(y;\theta)$ in its decomposition in the linear basis of Jack polynomials in the x variables are symmetric polynomials in the y variables and are referred to as skew Jack polynomials:

$$J_\lambda(x, y; \theta) = \sum_\mu J_\mu(x;\theta)\, J_{\lambda/\mu}(y;\theta). \tag{2.3}$$

Similarly, one writes

$$\widetilde{J}_\lambda(x, y; \theta) = \sum_\mu \widetilde{J}_\mu(x;\theta)\, \widetilde{J}_{\lambda/\mu}(y;\theta).$$

It is known (see e.g. [31, Chapter VI, Section 10]) that $J_{\lambda/\mu}(y;\theta) = \widetilde{J}_{\lambda/\mu}(y;\theta) = 0$ unless $\mu \subset \lambda$, which means that $\lambda_i \geq \mu_i$ for $i = 1, 2, \ldots$. Also $J_{\lambda/\emptyset}(y;\theta) = J_\lambda(y;\theta)$ and $\widetilde{J}_{\lambda/\emptyset}(y;\theta) = \widetilde{J}_\lambda(y;\theta)$.

Throughout the article the parameter θ remains fixed and, thus, we usually omit it, writing simply $J_\lambda(x)$, $\widetilde{J}_\lambda(x)$, $J_{\lambda/\mu}(x)$, $\widetilde{J}_{\lambda/\mu}(x)$.

A *specialization* ρ is an algebra homomorphism from Λ to the set of complex numbers. A specialization is called Jack-positive if its values on all (skew) Jack polynomials with a fixed parameter $\theta > 0$ are real and non-negative. The following statement gives a classification of all Jack-positive specializations.

Proposition 2.1 ([27]) *For any fixed $\theta > 0$, Jack-positive specializations can be parameterized by triplets (α, β, γ), where α, β are sequences of real numbers with*

$$\alpha_1 \geq \alpha_2 \geq \cdots \geq 0, \quad \beta_1 \geq \beta_2 \geq \cdots \geq 0, \quad \sum_{i=1}^{\infty}(\alpha_i + \beta_i) < \infty$$

and γ is a non-negative real number. The specialization corresponding to a triplet (α, β, γ) is given by its values on the Newton power sums p_k, $k \geq 1$:

$$p_1 \mapsto p_1(\alpha, \beta, \gamma) = \gamma + \sum_{i=1}^{\infty}(\alpha_i + \beta_i),$$
$$p_k \mapsto p_k(\alpha, \beta, \gamma) = \sum_{i=1}^{\infty}\alpha_i^k + (-\theta)^{k-1}\sum_{i=1}^{\infty}\beta_i^k, \quad k \geq 2.$$

The specialization with all parameters taken to be zero is called the *empty* specialization. This specialization maps a polynomial to its constant term (that is, the degree zero summand).

We prepare the following explicit formulas for Jack-positive specializations for future use.

Proposition 2.2 ([31, Chapter VI, (10.20)]) *Consider the Jack-positive specialization $\mathfrak{a}^N$ with $\alpha_1 = \alpha_2 = \cdots = \alpha_N = \mathfrak{a}$ and all other parameters set to zero. We have*

$$J_\lambda(\mathfrak{a}^N) = \begin{cases} \mathfrak{a}^{|\lambda|}\prod_{\square\in\lambda}\dfrac{N\theta + a'(\square) - \theta\, l'(\square)}{a(\square) + \theta\, l(\square) + \theta}, & \text{if } \ell(\lambda) \leq N, \\ 0, & \text{otherwise.} \end{cases}$$

Taking the limit $N \to \infty$ of specializations $\left(\frac{s}{N}\right)^N$ of Proposition 2.2 we obtain the following.

Proposition 2.3 *Consider the Jack-positive specialization $\mathfrak{r}_s$ with $\gamma = s$ and all other parameters set to zero. We have*

$$J_\lambda(\mathfrak{r}_s) = s^{|\lambda|}\,\theta^{|\lambda|}\prod_{\square\in\lambda}\frac{1}{a(\square) + \theta\, l(\square) + \theta}.$$

Certain specializations of skew Jack polynomials also admit explicit formulas. We say that two Young diagrams λ and μ interlace and write $\mu \prec \lambda$ if

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots .$$

Proposition 2.4 *For any complex number $\mathfrak{a} \neq 0$, the specialization value $J_{\lambda/\mu}(\mathfrak{a}^1)$ vanishes unless $\mu \prec \lambda$. In the latter case,*

$$J_{\lambda/\mu}(\mathfrak{a}^1) = \mathfrak{a}^{|\lambda|-|\mu|} \prod_{1\leq i\leq j\leq k-1} \frac{(\mu_i - \mu_j + \theta\,(j-i) + \theta)_{\mu_j-\lambda_{j+1}}}{(\mu_i - \mu_j + \theta\,(j-i) + 1)_{\mu_j-\lambda_{j+1}}} \cdot \frac{(\lambda_i - \mu_j + \theta\,(j-i) + 1)_{\mu_j-\lambda_{j+1}}}{(\lambda_i - \mu_j + \theta\,(j-i) + \theta)_{\mu_j-\lambda_{j+1}}}, \tag{2.4}$$

where k is any integer satisfying $\ell(\lambda) \leq k$ and we have used the Pochhammer symbol notation

$$(b)_n = b\,(b+1)\cdots(b+n-1).$$

When μ differs from λ by one box $\lambda = \mu \sqcup (i, j)$ the formula can be simplified to read in terms of $\widetilde{J}_{\lambda/\mu}$

$$\widetilde{J}_{\mu/\lambda}(\mathfrak{a}^1) = \mathfrak{a}\,\theta \prod_{l=1}^{i-1} \frac{a(l, j; \mu) + \theta\,(i-l+1)}{a(l, j; \mu) + \theta\,(i-l)} \cdot \frac{a(l, j; \mu) + 1 + \theta\,(i-l-1)}{a(l, j; \mu) + 1 + \theta\,(i-l)}. \tag{2.5}$$

Note that the arm lengths in the latter formula are computed with respect to the (smaller) diagram μ.

Proof The evaluation of (2.4) is known as the *branching rule for Jack polynomials* and is also a limit of a similar rule for Macdonald polynomials, see e.g. [31, (7.14'), Section VII, Chapter VI] or [36, (2.3)]. The formula (2.5) is obtained from (2.4) using (2.2). However, this computation is quite involved and we also provide an alternative way: formulas [31, (7.13), (7.14), Chapter VI] relate the skew Macdonald polynomials to certain functions $\varphi_{\lambda/\mu}$. Further, formulas [31, (6.20),(6.24), Chapter VI] give explicit expressions for $\varphi_{\lambda/\mu}$, and [31, Section 10, Chapter VI] explains that (skew) Jack polynomials are obtained from (skew) Macdonald polynomials parametrized by pairs (q, t) by the limit transition $q \to 1$, $t = q^\theta$. This limit in the expression for $\varphi_{\lambda/\mu}$ of [31, (6.24), Chapter VI] gives (2.5). □

We also need the following two summation formulas for Jack polynomials.

Proposition 2.5 *Take two specializations ρ_1, ρ_2 such that the series $\sum_{k=1}^{\infty} \frac{p_k(\rho_1)\,p_k(\rho_2)}{k}$ is absolutely convergent, and define*

$$H_\theta(\rho_1; \rho_2) = \exp\left(\sum_{k=1}^{\infty} \frac{\theta}{k}\, p_k(\rho_1)\, p_k(\rho_2)\right).$$

Then

$$\sum_{\lambda\in\mathbb{Y}} J_\lambda(\rho_1)\,\widetilde{J}_\lambda(\rho_2) = H_\theta(\rho_1;\,\rho_2), \tag{2.6}$$

and more generally for any $\nu, \kappa \in \mathbb{Y}$

$$\sum_{\lambda\in\mathbb{Y}} J_{\lambda/\nu}(\rho_1)\,\widetilde{J}_{\lambda/\kappa}(\rho_2) = H_\theta(\rho_1;\,\rho_2)\sum_{\mu\in\mathbb{Y}} J_{\kappa/\mu}(\rho_1)\,\widetilde{J}_{\nu/\mu}(\rho_2). \tag{2.7}$$

Proof (2.6) is the specialized version of a Cauchy-type identity for Jack polynomials, see e.g. [31, (10.4), Section 10, Chapter VI]. The latter is also a $(q, t) \to (1, 1)$ limit of a similar identity for Macdonald polynomials [31, (4.13), Section 4, Chapter VI], as is explained in [31, Section 10, Chapter VI]. Similarly, (2.7) is the specialized version of the limit of a skew-Cauchy identity for Macdonald polynomials, see e.g. [31, Exercise 6, Section 7, Chapter VI]. □

2.2 Probability measures related to Jack polynomials

We start with the definition of Jack probability measures which is based on (2.6).

Definition 2.6 Given two Jack-positive specializations ρ_1 and ρ_2 such that the series $\sum_{k=1}^{\infty}\frac{p_k(\rho_1)\,p_k(\rho_2)}{k}$ is absolutely convergent, the Jack probability measure $\mathcal{J}_{\rho_1;\rho_2}$ on $\mathbb{Y}$ is defined through

$$\mathcal{J}_{\rho_1;\rho_2}(\lambda) = \frac{J_\lambda(\rho_1)\,\widetilde{J}_\lambda(\rho_2)}{H_\theta(\rho_1;\,\rho_2)}, \tag{2.8}$$

with the normalization constant being given by

$$H_\theta(\rho_1;\,\rho_2) = \exp\Big(\sum_{k=1}^{\infty}\frac{\theta}{k}\,p_k(\rho_1)\,p_k(\rho_2)\Big).$$

Remark 2.7 The construction of probability measures via specializations of symmetric polynomials was originally suggested by Okounkov in the context of *Schur measures* [37]. Recently, similar constructions for more general polynomials have led to many interesting results starting from the paper [7] by Borodin and Corwin. We refer to [8, Introduction] for the chart of probabilistic objects which are linked to various degenerations of *Macdonald polynomials*.

The following statement is a corollary of Propositions 2.2, 2.3 and formula (2.2).

Proposition 2.8 *Take specializations* 1^N *and* $\mathfrak{r}_s$ *of Propositions 2.2 and 2.3, respectively. Then* $\mathcal{J}_{1^N;\mathfrak{r}_s}(\lambda)$ *vanishes unless* $\lambda \in \mathbb{Y}^N$*, and in the latter case we have*

$$\mathcal{J}_{1^N;\mathfrak{r}_s}(\lambda) = \exp\big(-\theta s N\big)\,s^{|\lambda|}\,\theta^{|\lambda|}\prod_{\square\in\lambda}\frac{N\theta + a'(\square) - \theta\,l'(\square)}{(a(\square)+\theta\,l(\square)+\theta)(a(\square)+\theta\,l(\square)+1)}. \tag{2.9}$$

Next, we consider limits of the measures $\mathcal{J}_{1^N;\mathfrak{r}_s}$ under a diffusive rescaling of s and λ. Define the open Weyl chamber $\mathcal{W}^N = \{y \in \mathbb{R}^N : y_1 < y_2 < \cdots < y_N\}$ and let $\overline{\mathcal{W}^N}$ be its closure.

Proposition 2.9 *Fix some $N \in \mathbb{N}$. Then under the rescaling*

$$s = \varepsilon^{-1}\frac{t}{\theta}, \quad \lambda_i = \varepsilon^{-1}t + \varepsilon^{-1/2}y_{N+1-i}, \quad i = 1, 2, \ldots, N,$$

the measures $\mathcal{J}_{1^N;\mathfrak{r}_s}$ converge weakly in the limit $\varepsilon \to 0$ to the probability measure with density

$$\frac{1}{Z}\prod_{i<j}(y_j - y_i)^{2\theta}\prod_{i=1}^{N}\exp\left(-\frac{y_i^2}{2t}\right) \tag{2.10}$$

on the closed Weyl chamber $\overline{\mathcal{W}^N}$ where

$$Z = t^{\theta\frac{N(N-1)}{2}+\frac{N}{2}}(2\pi)^{N/2}\prod_{j=1}^{N}\frac{\Gamma(j\theta)}{\Gamma(\theta)}. \tag{2.11}$$

Note that we have chosen the notation is such a way that the row lengths λ_i are non-increasing, while the continuous coordinates y_i are non-decreasing in i.

Proof of Proposition 2.9 We start by observing that (2.10), (2.11) define a probability density, namely that the total mass of the corresponding measure is equal to one. Indeed, the computation of the normalization constant is a particular case of the *Selberg integral* (see [20,32,44]). Since $\mathcal{J}_{1^N;\mathfrak{r}_s}$ is also a probability measure, it suffices to prove that as $\varepsilon \to 0$

$$\mathcal{J}_{1^N;\mathfrak{r}_s}(\lambda) = \varepsilon^{N/2}\frac{1}{Z}\prod_{i<j}(y_j - y_i)^{2\theta}\prod_{i=1}^{N}\exp\left(-\frac{y_i^2}{2t}\right)\big(1 + o(1)\big)$$

with the error term $o(1)$ being uniformly small on compact subsets of $\mathcal{W}^N$. The product over boxes in the first row of λ in (2.9) is (with the convention $\lambda_{N+1} = 0$)

$$\prod_{i=1}^{\lambda_1}(N\theta + i - 1)\prod_{i=1}^{N}\prod_{j=\lambda_{i+1}+1}^{\lambda_i}\frac{1}{(\lambda_1 - j + \theta(i-1) + \theta)(\lambda_1 - j + \theta(i-1) + 1)}$$
$$= \frac{\Gamma(N\theta + \lambda_1)}{\Gamma(N\theta)}\prod_{i=1}^{N}\frac{\Gamma(\lambda_1 - \lambda_i + i\theta)}{\Gamma(\lambda_1 - \lambda_{i+1} + i\theta)}\frac{\Gamma(\lambda_1 - \lambda_i + i\theta + 1 - \theta)}{\Gamma(\lambda_1 - \lambda_{i+1} + i\theta + 1 - \theta)}$$

$$
\begin{aligned}
&= \frac{\Gamma(\theta)}{\Gamma(N\theta)\,\Gamma((N-1)\theta+\lambda_1+1)} \prod_{i=2}^{N} \frac{\Gamma(\lambda_1-\lambda_i+i\theta)}{\Gamma(\lambda_1-\lambda_i+(i-1)\,\theta)} \\
&\quad \times \frac{\Gamma(\lambda_1-\lambda_i+(i-1)\,\theta+1)}{\Gamma(\lambda_1-\lambda_i+(i-2)\,\theta+1)} \\
&\sim \frac{\Gamma(\theta)}{\Gamma(N\theta)\Gamma((N-1)\,\theta+\lambda_1+1)} \prod_{i=1}^{N-1} (\varepsilon^{-1/2}(y_N-y_i))^{2\theta}
\end{aligned}
$$

where we have written $A(\varepsilon) \sim B(\varepsilon)$ for $\lim_{\varepsilon\to 0} \frac{A(\varepsilon)}{B(\varepsilon)} = 1$. Further,

$$
\begin{aligned}
&\frac{e^{-t\varepsilon^{-1}}\, t^{\lambda_1} \varepsilon^{-\lambda_1}}{\Gamma((N-1)\theta+\lambda_1+1)} \sim e^{-t\varepsilon^{-1}} \left(t\varepsilon^{-1}\right)^{t\varepsilon^{-1}+y_N\varepsilon^{-1/2}+1/2} \\
&\quad \times \frac{1}{\sqrt{2\pi}} \left(\frac{(N-1)\theta+t\varepsilon^{-1}+y_N\varepsilon^{-1/2}+1}{e}\right)^{-(N-1)\theta-t\varepsilon^{-1}-y_N\varepsilon^{-1/2}-1} \\
&\quad \sim \frac{\left(t\varepsilon^{-1}\right)^{1/2-(N-1)\theta-1}}{\sqrt{2\pi}}\, e^{(N-1)\theta+1+y_N\varepsilon^{-1/2}} \\
&\quad \times \left(1+\varepsilon^{1/2}\,\frac{y_N}{t}+\varepsilon\,\frac{(N-1)\theta+1}{t}\right)^{-(N-1)\theta-t\varepsilon^{-1}-y_N\varepsilon^{-1/2}-1} \\
&\quad \sim \frac{\left(t\varepsilon^{-1}\right)^{1/2-(N-1)\theta+\theta}}{\sqrt{2\pi}} e^{(N-1)\theta+1+y_N\varepsilon^{-1/2}} \\
&\quad \times \exp\left(\left(\varepsilon^{1/2}\,\frac{y_N}{t}+\varepsilon\,\frac{(N-1)\theta+1}{t}-\varepsilon\frac{y_N^2}{2\,t^2}\right)\left(-t\varepsilon^{-1}-y_N\varepsilon^{-1/2}\right)\right) \\
&\quad = \varepsilon^{1/2+(N-1)\theta}\,\frac{t^{-1/2-(N-1)\theta}}{\sqrt{2\pi}}\,\exp\left(-\frac{y_N^2}{2\,t}\right).
\end{aligned}
$$

Therefore, the factors coming from the first row of λ in (2.9) are asymptotically given by

$$
\frac{\Gamma(\theta)}{\Gamma(N\theta)}\,\varepsilon^{1/2}\,\frac{t^{-1/2-(N-1)\theta}}{\sqrt{2\pi}}\,\exp\left(-\frac{y_N^2}{2\,t}\right)\prod_{i=1}^{N-1}(y_N-y_i)^{2\theta}.
$$

Performing similar computations for the other rows we get

$$
\mathcal{J}_{1^N;\mathfrak{r}_s}(\lambda) = \varepsilon^{N/2} \prod_{j=1}^{N} \frac{\Gamma(\theta)\,t^{-1/2-(j-1)\theta}}{\Gamma(j\theta)\sqrt{2\pi}} \prod_{i<j}(y_j-y_i)^{2\theta} \prod_{i=1}^{N} \exp\left(-\frac{y_i^2}{2\,t}\right)\left(1+o(1)\right)
$$

which finishes the proof. □

We now proceed to the definition of probability measures on multilevel structures associated with Jack polynomials. Let $\mathbb{GT}^{(N)}$ denote the set of sequences of Young

diagrams $\lambda^1 \prec \lambda^2 \prec \cdots \prec \lambda^N$ such that $\ell(\lambda^i) \le i$ for every i and the Young diagrams interlace, that is,

$$\lambda_1^{i+1} \ge \lambda_1^i \ge \lambda_2^{i+1} \ge \cdots \ge \lambda_i^i \ge \lambda_{i+1}^{i+1}, \quad i = 1, 2, \ldots, N-1.$$

The following definition is motivated by the property (2.3) of skew Jack polynomials.

Definition 2.10 A probability distribution P on arrays $(\lambda^1 \prec \cdots \prec \lambda^N) \in \mathbb{GT}^{(N)}$ is called a *Jack–Gibbs* distribution, if for any $\mu \in \mathbb{Y}^N$ such that $P(\lambda^N = \mu) > 0$ the conditional distribution of $\lambda^1 \ldots, \lambda^{N-1}$ given $\lambda^N = \mu$ is

$$P\big(\lambda^1, \ldots, \lambda^{N-1} \mid \lambda^N = \mu\big) = \frac{J_{\mu/\lambda^{N-1}}(1^1)\, J_{\lambda^{N-1}/\lambda^{N-2}}(1^1) \cdots J_{\lambda^2/\lambda^1}(1^1)\, J_{\lambda^1}(1^1)}{J_\mu(1^N)}. \tag{2.12}$$

Remark 2.11 When $\theta = 1$, (2.12) implies that the conditional distribution of $\lambda^1, \ldots, \lambda^{N-1}$ is *uniform* on the polytope defined by the interlacing conditions.

One important example of a Jack–Gibbs measure is given by the following definition, see [7,8,11,13] for a review of related constructions in the context of Schur and, more generally, Macdonald polynomials.

Definition 2.12 Given a Jack-positive specialization ρ such that $\sum_{k=1}^\infty \frac{p_k(\rho)}{k} < \infty$ we define the ascending Jack process $\mathcal{J}^{asc}_{\rho;N}$ as the probability measure on $\mathbb{GT}^{(N)}$ given by

$$\mathcal{J}^{asc}_{\rho;N}(\lambda^1, \lambda^2, \ldots, \lambda^N) = \frac{\widetilde{J}_{\lambda^N}(\rho)\, J_{\lambda^N/\lambda^{N-1}}(1^1) \cdots J_{\lambda^2/\lambda^1}(1^1)\, J_{\lambda^1}(1^1)}{H_\theta(\rho; 1^N)}. \tag{2.13}$$

Remark 2.13 If ρ is the empty specialization, then $\mathcal{J}^{asc}_{\rho;N}$ assigns mass 1 to the single element of $\mathbb{GT}^{(N)}$ such that $\lambda^i_j = 0$, $1 \le i \le j \le N$.

Lemma 2.14 *The formula* (2.13) *defines a Jack–Gibbs probability distribution. Furthermore, for any* $1 \le k \le N$ *the projection of* $\mathcal{J}^{asc}_{\rho;N}$ *to* $(\lambda^1, \ldots, \lambda^k)$ *is* $\mathcal{J}^{asc}_{\rho;k}$*, and the projection of* $\mathcal{J}^{asc}_{\rho;N}$ *to* λ^k *is* $\mathcal{J}_{\rho;1^k}$.

Proof The formula (2.4) yields that $J_{\lambda/\mu}(1^1)$ vanishes unless $\mu \prec \lambda$, thus, the support of $\mathcal{J}^{asc}_{\rho;N}$ is indeed a subset of $\mathbb{GT}^{(N)}$. Now we can sum (2.13) sequentially over λ^1, …, λ^{N-1} using (2.3). This proves that the projection of $\mathcal{J}^{asc}_{\rho;N}$ to λ^N is $\mathcal{J}_{\rho;1^N}$. Thus, since $\mathcal{J}_{\rho;1^N}$ is a probability measure, so is $\mathcal{J}^{asc}_{\rho;N}$. Further, dividing $\mathcal{J}^{asc}_{\rho;N}(\lambda^1, \ldots, \lambda^N)$ by $\mathcal{J}_{\rho;1^N}(\lambda^N)$ we get the conditional distribution (2.12), which proves that (2.13) is Jack–Gibbs.

To compute the projection onto $(\lambda^1, \ldots, \lambda^k)$ we sum (2.13) sequentially over λ^N,…, λ^{k+1} using (2.7) and arrive at $\mathcal{J}^{asc}_{\rho;k}$. In order to further compute the projection to λ^k we also sum over $\lambda^1, \ldots, \lambda^{k-1}$ using (2.3) and get $\mathcal{J}_{\rho;1^k}$. □

Define the (open) Gelfand–Tsetlin cone via

$$\mathcal{G}^N = \left\{ y \in \mathbb{R}^{N(N+1)/2} : y_i^{j+1} < y_i^j < y_{i+1}^{j+1},\ 1 \le i \le j \le N-1 \right\},$$

and let $\overline{\mathcal{G}^N}$ be its closure. A natural continuous analogue of Definition 2.10 is:

Definition 2.15 An absolutely continuous (with respect to the Lebesgue measure) probability distribution P on arrays $y \in \mathcal{G}^N$ is called θ–Gibbs, if the conditional distribution of the first $N-1$ levels y_i^k, $1 \le i \le k \le N-1$ given the N–th level $y_1^N, \dots, y_N^N$ has density

$$\begin{aligned} &P\left(y_i^k, 1 \le i \le k \le N-1 \mid y_1^N, \dots, y_N^N\right) \\ &= \prod_{k=2}^{N} \left(\frac{\Gamma(k\theta)}{\Gamma(\theta)^k} \prod_{1 \le i < m \le k-1} (y_m^{k-1} - y_i^{k-1}) \prod_{i=1}^{k-1} \prod_{j=1}^{k} |y_j^k \right. \\ &\left. - y_i^{k-1}|^{\theta-1} \prod_{1 \le j < n \le k} (y_n^k - y_j^k)^{1-2\theta} \right). \end{aligned} \tag{2.14}$$

To see that (2.14) indeed defines a probability measure, one can use a version of the Dixon-Anderson identity (see [2,17], [20, Chapter 4]), which reads

$$\begin{aligned} &\int \int \dots \int \prod_{1 \le i < j \le m} |u_i - u_j| \prod_{i=1}^{m} \prod_{j=1}^{m+1} |u_i - v_j|^{\theta-1} \, \mathrm{d}u_1 \, \mathrm{d}u_2 \, \dots \, \mathrm{d}u_m \\ &= \frac{\Gamma(\theta)^{m+1}}{\Gamma((m+1)\theta)} \prod_{1 \le i < j \le m+1} |v_i - v_j|^{2\theta-1}, \end{aligned} \tag{2.15}$$

where the integration is performed over the domain

$$v_1 < u_1 < v_2 < u_2 < v_3 < \dots < u_m < v_{m+1}.$$

Applying (2.15) sequentially to integrate the density in (2.14) with respect to the variable y_1^1, then the variables y_1^2, y_2^2 and so on, we eventually arrive at 1.

Proposition 2.16 *Let $P(q)$, $q = 1, 2, \dots$ be a sequence of Jack–Gibbs measures on $\mathbb{GT}^{(N)}$, and for each q let $\{\lambda_i^k(q)\}$ be a $P(q)$–distributed random element of $\mathbb{GT}^{(N)}$. Suppose that there exist two sequences $m(q)$ and $b(q)$ such that $\lim_{q\to\infty} b(q) = \infty$ and as $q \to \infty$ the N–dimensional vector*

$$\left(\frac{\lambda_N^N - m(q)}{b(q)}, \frac{\lambda_{N-1}^N - m(q)}{b(q)}, \dots, \frac{\lambda_1^N - m(q)}{b(q)} \right)$$

converges weakly to a random vector whose distribution is absolutely continuous with respect to the Lebesgue measure. Then the whole $N(N+1)/2$–dimensional vector

$$\left(\frac{\lambda_i^k - m(q)}{b(q)}\right), \quad 1 \le i \le k \le N$$

also converges weakly and its limiting distribution is θ–Gibbs.

Proof Since we deal with probability distributions converging to another probability distribution, it suffices to check that the quantity in (2.12) (written in the rescaled and reordered coordinates $y_i^j = \frac{\lambda_{j+1-i}^j - m(q)}{b(q)}$) converges to (2.14), uniformly on compact subsets of $\mathcal{G}^N$.

$J_{\lambda/\mu}(1^1)$ can be evaluated according to the identity (2.4). Thus, with the notation $f(\alpha) = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha+\theta)}$ we have

$$J_{\lambda^k/\lambda^{k-1}}(1^1) = \prod_{1\le i\le j\le k-1} \frac{f(\lambda_i^{k-1} - \lambda_j^{k-1} + \theta\,(j-i))\, f(\lambda_i^k - \lambda_{j+1}^k + \theta\,(j-i))}{f(\lambda_i^{k-1} - \lambda_{j+1}^k + \theta\,(j-i))\, f(\lambda_i^k - \lambda_j^{k-1} + \theta\,(j-i))}.$$

The asymptotics $f(\alpha) \sim \alpha^{1-\theta}$ as $\alpha \to \infty$ shows

$$J_{\lambda^k/\lambda^{k-1}}(1^1) \sim \frac{b(q)^{(1-k)(1-\theta)}}{\Gamma(\theta)^{k-1}} \times \prod_{1\le i<j\le k-1} (y_j^{k-1} - y_i^{k-1})^{1-\theta} \prod_{1\le i<j\le k} (y_j^k - y_i^k)^{1-\theta} \prod_{i=1}^{k-1}\prod_{j=1}^{k} |y_i^{k-1} - y_j^k|^{\theta-1}.$$

It remains to analyze the asymptotics of $J_\mu(1^N)$ in the denominator of (2.12). To this end, we use the expression for $J_\mu(1^N)$ in Proposition 2.2 and recall that μ can be identified with λ^N to find that the product over the boxes in the first row of μ asymptotically (in the limit $q \to \infty$) behaves as

$$\begin{aligned}
&\prod_{i=1}^{\mu_1}(N\theta + i - 1) \prod_{i=1}^{N} \prod_{j=\mu_{i+1}+1}^{\mu_i} \frac{1}{\mu_1 - j + \theta\,(i-1) + \theta} \\
&= \frac{\Gamma(N\theta + \lambda_1)}{\Gamma(N\theta)} \prod_{i=1}^{N} \frac{\Gamma(\lambda_1 - \lambda_i + i\theta)}{\Gamma(\lambda_1 - \lambda_{i+1} + i\theta)} \\
&= \prod_{i=2}^{N} \left(\frac{\Gamma(\lambda_1 - \lambda_i + i\theta)}{\Gamma(\lambda_1 - \lambda_i + (i-1)\,\theta)}\right) \sim \prod_{i=1}^{N-1} \left(b(q)(y_N^N - y_i^N)\right)^{\theta}.
\end{aligned}$$

Performing the same computations for the other rows we find that, as $q \to \infty$,

$$\frac{1}{J_\mu(1^N)} \sim b(q)^{-\theta N(N-1)/2} \prod_{1\le i<j\le N} (y_j^N - y_i^N)^{-\theta}.$$

One obtains the desired convergence to (2.14) by putting together the asymptotics of the factors in (2.12) and multiplying the result by the term $b(q)^{N(N-1)/2}$ coming from the space rescaling. □

As a combination of Propositions 2.9 and 2.16 we obtain the following statement.

Corollary 2.17 *Fix some $N \in \mathbb{N}$. Then, under the rescaling*

$$s = \varepsilon^{-1}\frac{t}{\theta}, \quad \lambda_i^j = \varepsilon^{-1} t + \varepsilon^{-1/2} y_{j+1-i}^j, \quad 1 \le i \le j \le N,$$

the measures $\mathcal{J}^{asc}_{\mathfrak{r}_s;N}$ converge weakly in the limit $\varepsilon \to 0$ to the probability measure on the Gelfand–Tsetlin cone $\mathcal{G}^N$ with density

$$\frac{1}{Z}\prod_{i<j}\left(y_j^N - y_i^N\right)\prod_{i=1}^{N}\exp\left(-\frac{(y_i^N)^2}{2t}\right)\prod_{n=1}^{N-1}\prod_{1\le i<j\le n}\left(y_j^n - y_i^n\right)^{2-2\theta} \times \prod_{a=1}^{n}\prod_{b=1}^{n+1}\left|y_a^n - y_b^{n+1}\right|^{\theta-1} \tag{2.16}$$

where

$$Z = t^{\theta N(N-1)+N/2}\,(2\pi)^{N/2}\prod_{j=1}^{N}\frac{(\Gamma(j\theta))^2}{\Gamma(\theta)^{j+1}}. \tag{2.17}$$

Note that the probability measure of (2.16) is precisely the Hermite $\beta = 2\theta$ corners process with variance t of Definition 1.1.

Remark 2.18 When $\theta = 1$, the factors $(y_j^n - y_i^n)^{2-2\theta}$ and $|y_a^n - y_b^{n+1}|^{\theta-1}$ in (2.16) disappear, and the conditional distribution of $y^1, y^2, \ldots, y^{N-1}$ given y^N becomes *uniform* on the polytope defined by the interlacing conditions. This distribution is known to be that of eigenvalues of *corners* of a random Gaussian $N \times N$ Hermitian matrix sampled from the Gaussian Unitary Ensemble (see e.g. [4]). Similarly, for $\theta = 1/2$ and $\theta = 2$ one gets the joint distribution of the eigenvalues of corners of the Gaussian Orthogonal Ensemble and the Gaussian Symplectic Ensemble, respectively (see e.g. [33], [35, Section 4]).

2.3 Dynamics related to Jack polynomials

We are now ready to construct the stochastic dynamics related to Jack polynomials. Similar constructions for Schur, q-Whittacker and Macdonald polynomials can be found in [6,7,9,11].

Definition 2.19 Given two specializations ρ, ρ' define their union (ρ, ρ') through the formulas

$$p_k(\rho, \rho') = p_k(\rho) + p_k(\rho'), \quad k \geq 1$$

where $p_k, k \geq 1$ are the Newton power sums as before.

Let ρ and ρ' be two Jack-positive specializations such that $H_\theta(\rho; \rho') < \infty$. Define matrices $p^{\uparrow}_{\lambda\to\mu}$ and $p^{\downarrow}_{\lambda\to\mu}$ with rows and columns indexed by Young diagrams as follows:

$$p^{\uparrow}_{\lambda\to\mu}(\rho; \rho') = \frac{1}{H_\theta(\rho; \rho')} \frac{J_\mu(\rho)}{J_\lambda(\rho)} \widetilde{J}_{\mu/\lambda}(\rho'), \quad \lambda, \mu \in \mathbb{Y}, \quad J_\lambda(\rho) \neq 0, \quad (2.18)$$

$$p^{\downarrow}_{\lambda\to\mu}(\rho; \rho') = \frac{J_\mu(\rho)}{J_\lambda(\rho, \rho')} J_{\lambda/\mu}(\rho'), \qquad \lambda, \mu \in \mathbb{Y}, \quad J_\lambda(\rho, \rho') \neq 0. \quad (2.19)$$

The next three propositions follow from (2.3), (2.6), (2.7) (see also [6,7,11] for analogous results in the cases of Schur, q-Whittacker and Macdonald polynomials).

Proposition 2.20 *The matrices $p^{\uparrow}_{\lambda\to\mu}$ and $p^{\downarrow}_{\lambda\to\mu}$ are stochastic, that is, all matrix elements are non-negative, and for every $\lambda \in \mathbb{Y}$ we have*

$$\sum_{\mu\in\mathbb{Y}} p^{\uparrow}_{\lambda\to\mu}(\rho, \rho') = 1, \quad \text{if } J_\lambda(\rho) \neq 0,$$

$$\sum_{\mu\in\mathbb{Y}} p^{\downarrow}_{\lambda\to\mu}(\rho, \rho') = 1, \quad \text{if } J_\lambda(\rho, \rho') \neq 0.$$

Proposition 2.21 *For any $\mu \in \mathbb{Y}$ and any Jack-positive specializations ρ_1, ρ_2, ρ_3 we have*

$$\sum_{\lambda\in\mathbb{Y}:\, \mathcal{J}_{\rho_1;\rho_2}(\lambda)\neq 0} \mathcal{J}_{\rho_1;\rho_2}(\lambda)\, p^{\uparrow}_{\lambda\to\mu}(\rho_2; \rho_3) = \mathcal{J}_{\rho_1,\rho_3;\rho_2}(\mu),$$

$$\sum_{\lambda\in\mathbb{Y}:\, \mathcal{J}_{\rho_1;\rho_2,\rho_3}(\lambda)\neq 0} \mathcal{J}_{\rho_1;\rho_2,\rho_3}(\lambda)\, p^{\downarrow}_{\lambda\to\mu}(\rho_2; \rho_3) = \mathcal{J}_{\rho_1;\rho_2}(\mu).$$

Proposition 2.22 *The following commutation relation on matrices $p^{\uparrow}_{\lambda\to\mu}$ and $p^{\downarrow}_{\lambda\to\mu}$ holds:*

$$p^{\uparrow}(\rho_1, \rho_2; \rho_3)\, p^{\downarrow}(\rho_1; \rho_2) = p^{\downarrow}(\rho_1; \rho_2)\, p^{\uparrow}(\rho_1; \rho_3).$$

Let $X^N_{disc}(s), s \geq 0$ denote the continuous time Markov chain on $\mathbb{Y}^N$ with transition probabilities given by $p^{\uparrow}(1^N; \mathfrak{r}_s)$, $s \geq 0$ (and arbitrary initial condition $X^N_{disc}(0) \in \mathbb{Y}^N$). We record the jump rates of X^N_{disc} for later use.

Proposition 2.23 *The jump rates of the Markov chain X^N_{disc} on $\mathbb{Y}^N$ are given by*

$$q_{\lambda\to\mu} = \begin{cases} \dfrac{J_\mu(1^N)}{J_\lambda(1^N)}\,\widetilde{J}_{\mu/\lambda}(\mathfrak{r}_1), & \mu = \lambda \sqcup \square, \\ -\sum\limits_{\nu=\lambda\sqcup\square} q_{\lambda\to\nu}, & \mu = \lambda, \\ 0, & \text{otherwise.} \end{cases} \tag{2.20}$$

Explicitly, for $\mu = \lambda \sqcup (i, j)$ we have

$$q_{\lambda\to\mu} = \frac{\prod_{\square\in\mu} \dfrac{N\theta + a'(\square) - \theta\, l'(\square)}{a(\square;\mu) + \theta\, l(\square;\mu) + \theta}}{\prod_{\square\in\lambda} \dfrac{N\theta + a'(\square) - \theta\, l'(\square)}{a(\square;\lambda) + \theta\, l(\square;\lambda) + \theta}}\,\theta \prod_{k=1}^{i-1} \frac{a(k,j;\lambda) + \theta\,(i-k+1)}{a(k,j;\lambda) + \theta\,(i-k)} \times \frac{a(k,j;\lambda) + 1 + \theta\,(i-l-1)}{a(k,j;\lambda) + 1 + \theta\,(i-l)}. \tag{2.21}$$

Remark 2.24 While the jump rates $q_{\lambda\to\mu}$ are explicit, we are not aware of any fairly simple formulas for the transition probabilities $p^\uparrow(1^N;\mathfrak{r}_s)$, $s \geq 0$ of X^N_{disc}.

Proof of Proposition 2.23 The formula (2.20) is readily obtained from the definition of the transition probabilities in (2.18). In order to get (2.21) we note that $J_\lambda(1^N)$ and $J_\mu(1^N)$ have been computed in Proposition 2.2. Further, observe that $\widetilde{J}_{\mu/\lambda}$ is a symmetric polynomial of degree 1, thus, it is proportional to the sum of indeterminates p_1. Therefore, $\widetilde{J}_{\mu/\lambda}(1^1) = \widetilde{J}_{\mu/\lambda}(\mathfrak{r}_1)$ and we can use the formula (2.5) to evaluate it.

The following proposition will prove useful below.

Proposition 2.25 *The process $|X^N_{disc}| := \sum_{i=1}^N (X^N_{disc})_i$ is a Poisson process with intensity $N\theta$.*

Proof According to Proposition 2.23 the process $|X^N_{disc}|$ increases by 1 with rate

$$\sum_{\mu=\lambda\sqcup\square} \frac{J_\mu(1^N)}{J_\lambda(1^N)}\,\widetilde{J}_{\mu/\lambda}(\mathfrak{r}_1).$$

In order to evaluate the latter sum we use *Pieri's rule* for Jack polynomials, which is a formula for the product of a Jack polynomial with the sum of indeterminates and in our case reads

$$\sum_{\mu=\lambda\sqcup\square} J_\mu(1^N)\,\widetilde{J}_{\mu/\lambda}(\mathfrak{r}_1) = \theta\, p_1(1^N)\, J_\lambda(1^N) = N\,\theta\, J_\lambda(1^N).$$

The proof of Pieri's fule (for Macdonald polynomials, with the case of Jack polynomials being given by the limit transition $q \to 1$, $t = q^\theta$) can be found in [31, (6.24) and Section 10 in Chapter VI]. Note that in the formulas of [31] the notation $\varphi_{\mu/\lambda}$ is used for $\widetilde{J}_{\mu/\lambda}(\mathfrak{r}_1) = \widetilde{J}_{\mu/\lambda}(1^1)$ and the g_1 there is proportional to our p_1. □

Proposition 2.21 implies the following statement.

Proposition 2.26 *Suppose that the initial condition $X^N_{disc}(0)$ is the empty Young diagram, that is, $\lambda_1 = \lambda_2 = \cdots = \lambda_N = 0$. Then, for any fixed $s > 0$, the law of $X^N_{disc}(s)$ is given by $\mathcal{J}_{1^N;\mathfrak{r}_s}$ which was computed explicitly in Proposition 2.8.*

Our next goal is to define a stochastic dynamics on $\mathbb{GT}^{(N)}$. The construction we use is parallel to those of [6,7,9–11]; it is based on an idea going back to [16], which allows to couple the dynamics of Young diagrams of different sizes. We start from the degenerate discrete time dynamics $\lambda^0(n) = \emptyset$, $n \in \mathbb{N}_0$ and construct the discrete time dynamics of $\lambda^1, \lambda^2, \ldots, \lambda^N$ inductively. Given $\lambda^{k-1}(n)$, $n \in \mathbb{N}_0$ and a Jack-positive specialization ρ we define the process $\lambda^k(n)$, $n \in \mathbb{N}_0$ with a given initial condition $\lambda^k(0)$ satisfying $\lambda^{k-1}(0) \prec \lambda^k(0)$ as follows. We let the distribution of $\lambda^k(n+1)$ depend only on $\lambda^k(n)$ and $\lambda^{k-1}(n+1)$ and be given by

$$\mathbb{P}(\lambda^k(n+1) = \nu \mid \lambda^k(n) = \lambda,\ \lambda^{k-1}(n+1) = \mu) = \frac{\widetilde{J}_{\nu/\lambda}(\rho)\, J_{\nu/\mu}(1^1)}{\sum_{\kappa \in \mathbb{Y}} \widetilde{J}_{\kappa/\lambda}(\rho)\, J_{\kappa/\mu}(1^1)}. \tag{2.22}$$

Carrying out this procedure for $k = 1, 2, \ldots, N$ we end up with a discrete time Markov chain $\hat{X}^{multi}_{disc}(n;\rho)$, $n \in \mathbb{N}_0$ on $\mathbb{GT}^{(N)}$.

Definition 2.27 Define the continuous time dynamics $X^{multi}_{disc}(s)$, $s \geq 0$ on $\mathbb{GT}^{(N)}$ with an initial condition $X^{multi}_{disc}(0) \in \mathbb{GT}^{(N)}$ as the distributional limit

$$\lim_{\varepsilon \to 0} \hat{X}^{multi}_{disc}(\lfloor \varepsilon^{-1} s \rfloor; \mathfrak{r}_\varepsilon)$$

where all dynamics $\hat{X}^{multi}_{disc}(\cdot;\mathfrak{r}_\varepsilon)$ are started from the initial condition $X^{multi}_{disc}(0)$ and the specialization $\mathfrak{r}_\varepsilon$ is defined as in Proposition 2.3.

Remark 2.28 Alternatively, we could have started from the specialization ρ with a single α parameter $\alpha_1 = \varepsilon$ and we would have arrived at the same continuous time dynamics. Analogous constructions of the continuous time dynamics in the context of Schur and Macdonald polynomials can be found in [9] and [7].

Note that when $\lambda \prec \kappa$ the term $\widetilde{J}_{\kappa/\lambda}(\mathfrak{r}_\varepsilon)$ is of order $\varepsilon^{|\kappa|-|\lambda|}$ as $\varepsilon \to 0$. Therefore, the leading order term in the sum on the right-hand side of (2.22) comes from the choice $\kappa = \lambda$ unless that κ violates $\mu \prec \kappa$, in which case the leading term corresponds to taking $\kappa = \mu$. Moreover, the first-order terms come from the choices $\kappa = \lambda \sqcup \square$ and the resulting terms turn into the jump rates of the continuous time dynamics. Summing up, the continuous time dynamics $X^{multi}_{disc}(s)$, $s \geq 0$ looks as follows: given the trajectory of λ^{k-1}, a box $\square$ is added to the Young diagram λ^k at time t at the rate

$$q(\square, \lambda^k(s-), \lambda^{k-1}(s)) = \widetilde{J}_{(\lambda^k(s-)\sqcup\square)/\lambda^k(s-)}(\mathfrak{r}_1)\, \frac{J_{(\lambda^k(s-)\sqcup\square)/\lambda^{k-1}(s)}(1^1)}{J_{\lambda^k(s-)/\lambda^{k-1}(s)}(1^1)}. \tag{2.23}$$

In particular, the latter jump rates incorporate the following push interaction: if the coordinates of λ^{k-1} evolve in a way which violates the interlacing condition $\lambda^{k-1} \prec$

λ^k, then the appropriate coordinate of λ^k is pushed in the sense that a box is added immediately to the Young diagram λ^k to restore the interlacing. The factors on the right-hand side of (2.23) are explicit and given by (2.4) and (2.5). Simulations of the continuous time dynamics for $\theta = 0.5$ and $\theta = 2$ can be found at [25].

The following statement is based on the results of Propositions 2.20, 2.21, 2.22 and can be proved by the argument of [9, Sections 2.2, 2.3], see also [6,7,11].

Proposition 2.29 *Suppose that $X^{multi}_{disc}(s)$, $s \geq 0$ is started from a random initial condition with a Jack–Gibbs distribution. Then:*

- *the restriction of $X^{multi}_{disc}(s)$ to level N coincides with $X^{N}_{disc}(s)$, $s \geq 0$ started from the restriction to level N of the initial condition $X^{multi}_{disc}(0)$;*
- *the law of $X^{multi}_{disc}(s)$ at a fixed time $s > 0$ is a Jack–Gibbs distribution. Moreover, if $X^{multi}_{disc}(0)$ has law $\mathcal{J}^{asc}_{\rho;N}$, then $X^{multi}_{disc}(s)$ has law $\mathcal{J}^{asc}_{\rho,\mathfrak{r}_s;N}$.*

Remark 2.30 In fact, there is a way to generalize Proposition 2.29 to a statement describing the restriction of our multilevel dynamics started from Jack–Gibbs initial conditions to any monotone space-time path (meaning that we look at level N for some time, then at level $N-1$ and so on). We refer the reader to [9, Proposition 2.5] for a precise statement in the setting of multilevel dynamics based on Schur polynomials.

3 Convergence to Dyson Brownian motion

The goal of this section is to prove that the Markov chain $X^{N}_{disc}(s)$ converges in the diffusive scaling limit to a Dyson Brownian Motion.

To start with, we recall the existence and uniqueness result for Dyson Brownian motions with $\beta > 0$ (see e.g. [3, Proposition 4.3.5] for the case $\beta \geq 1$ and [14, Theorem 3.1] for the case $0 < \beta < 1$).

Proposition 3.1 *For any $N \in \mathbb{N}$ and $\beta > 0$, the system of SDEs*

$$\mathrm{d}X_i(t) = \frac{\beta}{2} \sum_{j \neq i} \frac{1}{X_i(t) - X_j(t)} \,\mathrm{d}t + \mathrm{d}W_i(t), \tag{3.1}$$

$i = 1, 2, \ldots, N$, with $W_1, W_2, \ldots, W_N$ being independent standard Brownian motions, has a unique strong solution taking values in the Weyl chamber $\overline{\mathcal{W}^N}$ for any initial condition $X(0) \in \overline{\mathcal{W}^N}$. Moreover, for all initial conditions, the stopping time

$$\tau := \inf\{t > 0 : X_i(t) = X_{i+1}(t) \ \textit{for some } i\} \tag{3.2}$$

is infinite with probability one if $\beta \geq 1$ and finite with positive probability if $0 < \beta < 1$.

We write $D^N = D([0, \infty), \mathbb{R}^N)$ for the space of right-continuous paths with left limits taking values in $\mathbb{R}^N$ and endow it with the usual Skorokhod topology (see e.g. [19]).

Theorem 3.2 *Fix $\theta \geq 1/2$ and let $\varepsilon > 0$ be a small parameter. Let the N–dimensional stochastic process $Y_\varepsilon^N(t) = (Y_\varepsilon^N(t)_1, \ldots, Y_\varepsilon^N(t)_N)$ be defined through*

$$(Y_\varepsilon^N(t))_i = \varepsilon^{1/2}\big((X_{disc}^N)_{N+1-i}\big(\varepsilon^{-1}\theta^{-1}t\big) - \varepsilon^{-1}t\big), \quad i = 1, \ldots, N$$

where $(X_{disc}^N)_i$ is i-th coordinate of the process X_{disc}^N. Suppose that, as $\varepsilon \to 0$, the initial conditions $Y_\varepsilon^N(0)$ converge to a point $Y(0)$ in the interior of $\overline{\mathcal{W}^N}$. Then the process $Y_\varepsilon^N(t)$ converges in the limit $\varepsilon \downarrow 0$ in law on D^N to the $\beta = 2\theta$–Dyson Brownian motion, that is, to the unique strong solution of (3.1) with $\beta = 2\theta$.

Remark 3.3 We believe that Theorem 3.2 should hold for any $\theta > 0$. However, the case $0 < \theta < 1/2$ presents additional technical challenges, since there the stopping time τ of (3.2) may be finite.

Let us first present the plan of the proof of Theorem 3.2. In Step 1 we study the asymptotics of the jump rates of X_{disc}^N in the scaling limit of Theorem 3.2. In Step 2 we prove the tightness of the processes Y_ε^N as $\varepsilon \to 0$. In Step 3 we show that subsequential limits of that family solve the SDE (3.1). This fact and the uniqueness of the solution to (3.1) yield together Theorem 3.2.

3.1 Step 1: Rates

$Y_\varepsilon^N(t)$ is a continuous time Markov process with state space $\overline{\mathcal{W}^N}$, a (constant) drift of $-\varepsilon^{-1/2}$ in each coordinate and jump rates

$$p_\varepsilon^N(y, y', t) = \theta^{-1}\varepsilon^{-1} q_{\frac{t}{\theta}\varepsilon^{-1}+\hat{y}\varepsilon^{-1/2} \to \frac{t}{\theta}\varepsilon^{-1}+\hat{y'}\varepsilon^{-1/2}}$$

where $\hat{y}, \hat{y'}$ are the vectors (viewed as Young diagrams) obtained from y, y' by reordering the components in decreasing order, and the intensities $q_{\lambda\to\mu}$ are given in Proposition 2.23. If we write $y' \approx_\varepsilon y$ for vectors y', y which differ in exactly one coordinate with the difference being $\varepsilon^{1/2}$, then $p_\varepsilon^N(y, y', t) = 0$ unless $y' \approx_\varepsilon y$. As we will see, in fact, $p_\varepsilon^N(y, y', t)$ does not depend on t.

Now, take two sequences $y' \approx_\varepsilon y$ with $y'_{N+1-i} - y_{N+1-i} = \varepsilon^{1/2}$ for some fixed $i \in \{1, 2, \ldots, N\}$. Define Young diagrams λ and μ via $\lambda_l = \frac{t}{\theta}\varepsilon^{-1} + y_{N+1-l}\varepsilon^{-1/2}$, $\mu_l = \frac{t}{\theta}\varepsilon^{-1} + y'_{N+1-j}\varepsilon^{-1/2}$. Then $\mu_i = \lambda_i + 1$ and $\mu_l = \lambda_l$ for $j \neq i$. Also, set $j = \lambda_i + 1$, so that $\mu = \lambda \sqcup (i, j)$.

Lemma 3.4 *For sequences $y' \approx_\varepsilon y$ differing in the $(N+1-i)$-th coordinate as above, we have in the limit $\varepsilon \to 0$:*

$$p_\varepsilon^N(y, y', t) = \varepsilon^{-1} + \varepsilon^{-1/2}\left(\sum_{j\neq i} \frac{\theta}{y_{N+1-i} - y_{N+1-j}}\right) + O(1)$$

where the error $O(1)$ is uniform on compact subsets of the open Weyl chamber $\mathcal{W}^N$.

Proof Using Proposition 2.23 we have

$$p_\varepsilon^N(y,y',t)=\frac{\varepsilon^{-1}}{\theta}\,\frac{J_\mu(1^N)}{J_\lambda(1^N)}\,\tilde{J}_{\mu/\lambda}(\mathfrak{r}_1)=\frac{\varepsilon^{-1}}{\theta}\,\frac{\prod_{\square\in\mu}\dfrac{N\theta+a'(\square)-\theta\,l'(\square)}{a(\square;\mu)+\theta\,l(\square;\mu)+\theta}}{\prod_{\square\in\lambda}\dfrac{N\theta+a'(\square)-\theta\,l'(\square)}{a(\square;\lambda)+\theta\,l(\square;\lambda)+\theta}}$$

$$\times\theta\prod_{m=1}^{i-1}\frac{a(m,j;\lambda)+\theta\,(i-m+1)}{a(m,j;\lambda)+\theta\,(i-m)}\,\frac{a(m,j;\lambda)+1+\theta\,(i-m-1)}{a(m,j;\lambda)+1+\theta\,(i-m)}$$

$$=\frac{\varepsilon^{-1}}{\theta}\prod_{k=1}^{j-1}\frac{j-k-1+\theta\,(\lambda'_k-i+1)}{j-k+\theta\,(\lambda'_k-i+1)}\prod_{m=1}^{i-1}\frac{\lambda_m-j+\theta\,(i-m)}{\lambda_m-j+\theta\,(i-m+1)}$$

$$\times\big((N-i+1)\,\theta+j-1\big)\prod_{m=1}^{i-1}\frac{\lambda_m-j+\theta\,(i-m+1)}{\lambda_m-j+\theta\,(i-m)}\cdot\frac{\lambda_m-j+1+\theta\,(i-m-1)}{\lambda_m-j+1+\theta\,(i-m)}$$

$$=\frac{\varepsilon^{-1}}{\theta}\prod_{k=1}^{j-1}\frac{j-k-1+\theta\,(\lambda'_k-i+1)}{j-k+\theta\,(\lambda'_k-i+1)}\big((N-i+1)\,\theta+j-1\big)$$

$$\times\prod_{m=1}^{i-1}\frac{(y_{N+1-m}-y_{N+1-i})\,\varepsilon^{-1/2}+\theta\,(i-m-1)}{(y_{N+1-m}-y_{N+1-i})\,\varepsilon^{-1/2}+\theta\,(i-m)}.\tag{3.3}$$

Now, for any y the corresponding Young diagram λ has λ_N columns of length N, $(\lambda_{N-1}-\lambda_N)$ columns of length $(N-1)$, $(\lambda_{N-2}-\lambda_{N-1})$ columns of length $(N-2)$ etc. Therefore, the latter expression for $p_\varepsilon^N(y,y',t)$ can be simplified to

$$\varepsilon^{-1}\prod_{r=0}^{N-i-1}\frac{j-1-\lambda_{N-r}+\theta\,(N-i-r+1)}{j-1-\lambda_{N-r}+\theta\,(N-i-r)}\prod_{m=1}^{i-1}$$

$$\times\frac{(y_{N+1-m}-y_{N+1-i})\,\varepsilon^{-1/2}+\theta\,(i-m-1)}{(y_{N+1-m}-y_{N+1-i})\,\varepsilon^{-1/2}+\theta\,(i-m)}$$

$$=\varepsilon^{-1}\prod_{k=i+1}^{N}\frac{(y_{N+1-i}-y_{N+1-k})\,\varepsilon^{-1/2}+\theta\,(k-i+1)}{(y_{N+1-i}-y_{N+1-k})\,\varepsilon^{-1/2}+\theta\,(k-i)}\prod_{m=1}^{i-1}$$

$$\times\frac{(y_{N+1-m}-y_{N+1-i})\,\varepsilon^{-1/2}+\theta\,(i-m-1)}{(y_{N+1-m}-y_{N+1-i})\,\varepsilon^{-1/2}+\theta\,(i-m)}$$

$$=\varepsilon^{-1}+\varepsilon^{-1/2}\left(\sum_{j\neq i}\frac{\theta}{y_{N+1-i}-y_{N+1-j}}\right)+O(1),\tag{3.4}$$

with the remainder $O(1)$ being uniform over y such that $|y_{N+1-i}-y_{N+1-j}|>\delta$ for $j\neq i$ and a fixed $\delta>0$. □

3.2 Step 2: Tightness

Let us show that the family Y_ε^N, $\varepsilon\in(0,1)$ is tight on D^N. To this end, we aim to apply the necessary and sufficient condition for tightness of [19, Corollary 3.7.4] and

need to show that, for any fixed $t \geq 0$, the random variables $Y_\varepsilon^N(t)$ are tight on $\mathbb{R}^N$ as $\varepsilon \downarrow 0$ and that for every $\Delta > 0$ and $T > 0$ there exists a $\delta > 0$ such that

$$\limsup_{\varepsilon \downarrow 0} \mathbb{P}\left(\sup_{0 \leq s < t \leq T, t-s < \delta} \left| (Y_\varepsilon^N)_i(t) - (Y_\varepsilon^N)_i(s) \right| > \Delta \right) < \Delta, \quad i = 1, 2, \ldots, N. \tag{3.5}$$

We first explain how to obtain the desired controls on $(Y_\varepsilon^N(t))_+$ (the vector of positive parts of the components of $Y_\varepsilon^N(t)$) and

$$\sup_{0 \leq s < t \leq T, t-s < \delta} \Big((Y_\varepsilon^N)_i(t) - (Y_\varepsilon^N)_i(s) \Big), \quad i = 1, 2, \ldots, N. \tag{3.6}$$

To control $(Y_\varepsilon^N(t))_+$ and the expressions in (3.6) we proceed by induction over the index of the coordinates in Y_ε^N. For the first coordinate $(Y_\varepsilon^N)_1$ the explicit formula (3.4) in Step 1 shows that the jump rates of the process $(Y_\varepsilon^N)_1$ are bounded above by ε^{-1}. Hence, a comparison with a Poisson process with jump rate ε^{-1}, jump size $\varepsilon^{1/2}$ and drift $-\varepsilon^{-1/2}$ shows that $((Y_\varepsilon^N)_1(t))_+$ and the expression in (3.6) for $i = 1$ behave in accordance with the conditions of Corollary 3.7.4 in [19] as stated above. Next, we consider $(Y_\varepsilon^N)_i$ for some $i \in \{2, 3, \ldots, N\}$. In this case, the formula (3.4) in Step 1 shows that, whenever the spacing $(Y_\varepsilon^N)_i - (Y_\varepsilon^N)_{i-1}$ exceeds $\Delta/3$, the jump rate of $(Y_\varepsilon^N)_i$ is bounded above by

$$\varepsilon^{-1} + \sum_{j=1}^{i-1} \frac{\theta\, \varepsilon^{-1/2}}{(Y_\varepsilon^N)_i(t) - (Y_\varepsilon^N)_j(t)} + C(\Delta) \leq \varepsilon^{-1} + \frac{3(i-1)\theta}{\Delta}\, \varepsilon^{-1/2} + C(\Delta). \tag{3.7}$$

Let us show that $(Y_\varepsilon^N)_i$ can be coupled with a Poisson jump process R_ε with jump size $\varepsilon^{1/2}$, jump rate given by the right-hand side of the last inequality and drift $-\varepsilon^{-1/2}$, so that, whenever $(Y_\varepsilon^N)_i - (Y_\varepsilon^N)_{i-1}$ exceeds $\Delta/3$ and $(Y_\varepsilon^N)_i$ has a jump to the right, the process R_ε has a jump to the right as well.

To do this, recall that (by definition) the law of the jump times of Y_ε^N can be described as follows. We take N independent exponential random variables $a_1, \ldots, a_N$ with means $r_j(Y_\varepsilon^N)$, $j = 1, \ldots, N$ defined by (3.3) with y and y' differing in the j-th coordinate. If we let k be the index for which $a_k = \min(a_1, \ldots, a_N)$, then at time a_k the k-th particle (that is, $(Y_\varepsilon^N)_k$) jumps. After this jump we repeat the procedure again to determine the next jump.

Let M denote the right-hand side of (3.7) and consider in each time interval between the jumps of Y_ε^N an additional independent exponential random variable b with mean $M - r_i(Y_\varepsilon^N)$ if $(Y_\varepsilon^N)_i - (Y_\varepsilon^N)_{i-1}$ exceeds $\Delta/3$ and with mean M otherwise. Now, instead of considering $\min(a_1, \ldots, a_N)$, we consider $\min(a_1, \ldots, a_N, b)$. If the minimum is given by b, then no jump happens and the whole procedure is repeated. Now, we *define* the jump times of process R_ε to be all times when the clock of the i-th particle rings provided that $(Y_\varepsilon^N)_i - (Y_\varepsilon^N)_{i-1}$ exceeds $\Delta/3$, and also all times when

the auxilliary random variable b constitutes the minimum. One readily checks that R_ε is given by a Poisson jump process of constant intensity M and drift $-\varepsilon^{-1/2}$.

We further use the convergence of R_ε to Brownian motion with drift $3(i-1)\theta/\Delta$, which implies the tightness and conditions (3.5) are satisfied for R_ε. Now we get the desired control for $((Y_\varepsilon^N)_i(t))_+$ and the quantities in (3.6) by invoking the induction hypothesis when spacing $(Y_\varepsilon^N)_i - (Y_\varepsilon^N)_{i-1}$ is less than $\Delta/3$ and by comparison with R_ε when the spacing is larger.

It remains to observe that $(Y_\varepsilon^N(t))_-$ (the vector of negative parts of the components of Y_ε^N) and

$$\sup_{0\le s<t\le T, t-s<\delta} -\big((Y_\varepsilon^N)_i(t) - (Y_\varepsilon^N)_i(s)\big), \quad i = 1, 2, \dots, N$$

can be dealt with in a similar manner (but considering the rightmost particle first and moving from right to left). Together these controls yield the conditions of [19, Corollary 3.7.4].

We also note that, since the maximal size of the jumps tends to zero as $\varepsilon \downarrow 0$, any limit point of the family Y_ε^N, $\varepsilon \in (0, 1)$ as $\varepsilon \downarrow 0$ must have continuous paths (see e.g. [19, Theorem 3.10.2]).

Remark 3.5 Note that in the proof of the tightness result the condition $\theta \ge 1/2$ is not used.

3.3 Step 3: SDE for subsequential limits

Throughout this section we let Y^N be an arbitrary limit point of the family Y_ε^N as $\varepsilon \downarrow 0$. Our goal is to identify Y^N with the solution of (3.1). We pick a sequence of Y_ε^N which converges to Y^N in law, and by virtue of the Skorokhod Embedding Theorem (see e.g. Theorem 3.5.1 in [18]) may assume that all processes involved are defined on the same probability space and that the convergence holds in the almost sure sense. In the rest of this section all limits $\varepsilon \to 0$ are taken along this sequence.

Let $\mathcal{F}$ denote the set of all infinitely differentiable functions on $\overline{\mathcal{W}^N}$ whose support is a compact subset of $\mathcal{W}^N$. Define

$$\mathcal{F}_\delta := \big\{ f \in \mathcal{F} : f(\mathbf{x}) = 0 \ \text{ whenever } \ \mathrm{dist}(\mathbf{x}, \partial\overline{\mathcal{W}^N}) \le \delta \big\}$$

where $\partial\overline{\mathcal{W}^N}$ denotes the boundary of $\overline{\mathcal{W}^N}$ and dist stands for the L^∞ distance:

$$\mathrm{dist}(\mathbf{x}, \partial\overline{\mathcal{W}^N}) = \min_{i=1,\dots,N-1} |x_{i+1} - x_i|.$$

Clearly, $\mathcal{F} = \bigcup_{\delta>0} \mathcal{F}_\delta$.

For functions $f \in \mathcal{F}$ we consider the processes

$$M^f(t) := f(Y^N(t)) - f(Y^N(0)) - \int_0^t \sum_{1\le i\ne j\le N} \frac{\theta}{Y_i^N(s) - Y_j^N(r)} f_{y_i}(Y^N(r))\,\mathrm{d}r - \frac{1}{2}\int_0^t \sum_{i=1}^N f_{y_i y_i}(Y^N(r))\,\mathrm{d}r. \tag{3.8}$$

Here, f_{y_i} ($f_{y_i y_i}$ resp.) stands for the first (second resp.) partial derivative of f with respect to y_i.

In Step 3a we show that the processes in (3.8) are martingales and identify their quadratic covariations. In step 3b we use the latter results to derive the SDEs for the processes $Y_1^N,\ldots, Y_N^N$.

Step 3a. We now fix an $f \in \mathcal{F}_\delta$ for some $\delta > 0$ and consider the family of martingales

$$M_\varepsilon^f(t) := f(Y_\varepsilon^N(t)) - f(Y_\varepsilon^N(0)) - \int_0^t \Bigg(\sum_{i=1}^N -\varepsilon^{-1/2} f_{y_i}(Y_\varepsilon^N(r)) + \sum_{y'\approx_\varepsilon Y^N(r)} p_\varepsilon^N(Y_\varepsilon^N(r), y', s)(f(y') - f(Y^N(s)))\Bigg)\mathrm{d}r, \quad \varepsilon > 0. \tag{3.9}$$

Lemma 3.4 implies that the integrand in (3.9) behaves asymptotically as

$$\frac{1}{2}\sum_{i=1}^N f_{y_i y_i}(Y_\varepsilon^N(r)) + \sum_{i=1}^N b_i(Y_\varepsilon^N(r)) f_{y_i}(Y_\varepsilon^N(r)) + O(\varepsilon^{1/2}),$$

where $b_i(y) = \sum_{j\ne i} \frac{\theta}{y_i - y_j}$. Note also that, for any fixed function $f \in \mathcal{F}$, the error terms can be bounded uniformly for all sequences $y' \approx_\varepsilon y$ as above, since f and all its partial derivatives are bounded and vanish in a neighborhood of the boundary $\partial\overline{\mathcal{W}^N}$ of $\overline{\mathcal{W}^N}$.

By taking the limit of the corresponding martingales M_ε^f for a fixed $f \in \mathcal{F}$ and noting that their limit M^f can be bounded uniformly on every compact time interval, we conclude that M^f must be a martingale as well.

In order to proceed further we recall the following definitions. For a real-valued function f defined on an interval $[0, T]$ (T can be $+\infty$ here), its quadratic variation $\langle f\rangle(t)$ is defined for $0 \le t \le T$ via

$$\langle f\rangle(t) = \lim_{\substack{||\mathfrak{P}||\to 0\\ \mathfrak{P}=(t_0<t_1<\cdots<t_k)}} \sum_{i=1}^k \big(f(t_i) - f(t_{i-1})\big)^2$$

where $\mathfrak{P}$ ranges over all ordered collections of points $0 = t_0 < t_1 < \cdots < t_k = t$ with k being arbitrary, and $||\mathfrak{P}|| = \min_{1\le i\le k}(t_i - t_{i-1})$. Similarly, for two function f and g their quadratic covariation $\langle f, g\rangle(t)$ is defined as

$$\langle f, g\rangle(t) = \lim_{\substack{||\mathfrak{P}||\to 0 \\ P=(t_0<t_1<\cdots<t_k)}} \sum_{i=1}^{k} \big(f(t_i) - f(t_{i-1})\big)\big(g(t_i) - g(t_{i-1})\big).$$

Lemma 3.6 *For any two functions $g, h \in \mathcal{F}$, the quadratic covariation of M^g and M^h is given by*

$$\big\langle M^g, M^h\big\rangle(t) = \sum_{j=1}^{N} \int_0^t g_{y_j}(Y^N(r)) h_{y_j}(Y^N(r))\,\mathrm{d}r. \tag{3.10}$$

Proof Due to the polarization identity

$$2\,\langle M^g, M^h\rangle(t) = \langle M^g + M^h\rangle(t) - \langle M^g, M^g\rangle(t) - \langle M^h, M^h\rangle(t)$$

it is enough to consider the case $g = h$, that is, to determine the quadratic variation $\langle M^g\rangle(t)$.

We proceed by finding the limit of the quadratic variation processes $\langle M^g_\varepsilon\rangle$ of M^g_ε as $\varepsilon \to 0$. For each $\varepsilon > 0$ and $j = 1, \ldots, N$ define $\mathcal{S}^j_\varepsilon$ as the (random) set of all times when the j-th coordinate of Y^N_ε jumps. Note that the sets $\mathcal{S}^j_\varepsilon$ are pairwise disjoint and their union $\bigcup_{j=1}^N S^j_\varepsilon$ is a Poisson point process of intensity $\varepsilon^{-1}N$ (see Proposition 2.25).

Recall that the quadratic variation process $\langle M^g_\varepsilon\rangle(t)$ of M^g_ε is given by the sum of squares of the jumps of the process M^g_ε (see e.g. [15, Proposition 8.9]) and conclude

$$\big\langle M^g_\varepsilon\big\rangle(t) = \sum_{j=1}^{N} \sum_{r\in\mathcal{S}^j_\varepsilon\cap[0,t]} \big(\varepsilon^{1/2} g_{y_j}(Y^N_\varepsilon(r) + O(\varepsilon)\big)^2, \tag{3.11}$$

with a uniform error term $O(\varepsilon)$. Suppose that $g \in \mathcal{F}_{2\delta}$ and consider new N pairwise disjoint sets $\widehat{\mathcal{S}}^j_\varepsilon$, $j = 1, \ldots, N$ satisfying $\bigcup_{j=1}^N \widehat{\mathcal{S}}^j_\varepsilon = \bigcup_{j=1}^N \mathcal{S}^j_\varepsilon$ and defined through the following procedure. Take any $r \in \bigcup_{j=1}^N S^j_\varepsilon$ and suppose that $r \in \mathcal{S}^k_\varepsilon$. If $\mathrm{dist}(Y^N_\varepsilon(r), \partial\overline{\mathcal{W}^N}) \geq \delta$, then put $r \in \widehat{\mathcal{S}}^k_\varepsilon$. Otherwise, take an independent random variable κ sampled from the uniform distribution on the set $\{1, 2, \ldots, N\}$ and put $r \in \widehat{\mathcal{S}}^\kappa_\varepsilon$. The definition implies that, for small ε,

$$\big\langle M^g_\varepsilon\big\rangle(t) = \sum_{j=1}^{N} \sum_{r\in\widehat{\mathcal{S}}^j_\varepsilon\cap[0,t]} \big(\varepsilon^{1/2} g_{y_j}(Y^N_\varepsilon(r) + O(\varepsilon)\big)^2, \tag{3.12}$$

with a uniform error term $O(\varepsilon)$. Now, take any two reals $a < b$. We claim that the sets $\widehat{\mathcal{S}}^k_\varepsilon$ satisfy the following property almost surely:

$$\lim_{\varepsilon\to 0} \varepsilon\,|\widehat{\mathcal{S}}^k_\varepsilon \cap [a, b]| = b - a. \tag{3.13}$$

Indeed, the Law of Large Numbers for Poisson Point Processes implies

$$\lim_{\varepsilon\to 0}\varepsilon\left|\left(\bigcup_{k=1}^{N}\widehat{\mathcal{S}}_\varepsilon^k\right)\cap[a,b]\right| = N(b-a). \tag{3.14}$$

On the other hand, Lemma 3.4 implies the following *uniform* asymptotics as $\varepsilon\to 0$:

$$\mathbb{P}\left(t\in\widehat{\mathcal{S}}_\varepsilon^k \mid \Theta_{<t}, t\in\bigcup_{j=1}^{N}\widehat{\mathcal{S}}_\varepsilon^j\right) = \frac{1}{N}+o(1),$$

where $\Theta_{<t}$ is the σ-algebra generated by the point process $\widehat{\mathcal{S}}_\varepsilon^k$, $j=1,\dots,N$ up to time t. Therefore, the conditional distribution of $|\widehat{\mathcal{S}}_\varepsilon^k\cap[a,b]|$ given $|(\bigcup_{k=1}^{N}\widehat{\mathcal{S}}_\varepsilon^k)\cap[a,b]|$ can be sandwiched between two binomial distributions with parameters $\frac{1}{N}\pm C(\varepsilon)$, where $\lim_{\varepsilon\to 0}C(\varepsilon)=0$ (see e.g. [30, Lemma 1.1]). Now, (3.14) and the Law of Large Numbers for the Binomial Distribution imply (3.13).

It follows that the sums in (3.12) approximate the corresponding integrals and we obtain

$$\lim_{\varepsilon\downarrow 0}\langle M_\varepsilon^g\rangle(t) = \sum_{j=1}^{N}\int_0^t g_{y_j}(Y^N(r))^2\,\mathrm{d}r. \tag{3.15}$$

Note that for each $g\in\mathcal{F}$, both $M_\varepsilon^g(t)^2$ and $\langle M_\varepsilon^g\rangle(t)$ are uniformly integrable on compact time intervals (this can be shown for example by another comparison with a Poisson jump process). Further, one of the properties of the quadratic variation (see e.g. [19, Chapter 2, Proposition 6.1]) is that $M_\varepsilon^g(t)^2-\langle M_\varepsilon^g\rangle(t)$ is a martingale. Sending $\varepsilon\to 0$ (see e.g. [19, Chapter 7, Problem 7] for a justification) it follows that the process

$$M^g(t)^2-\sum_{j=1}^{N}\int_0^t g_{y_j}(Y^N(r))^2\,\mathrm{d}r = \lim_{\varepsilon\to 0}\left(M_\varepsilon^g(t)^2-[M_\varepsilon^g](t)\right) \tag{3.16}$$

is a martingale. On the other hand, since $M^g(t)$ is continuous in t (see the end of Step 2), its quadratic variation $\langle M^g\rangle(t)$ is a unique increasing predictable process such that $M^g(t)-\langle M^g\rangle(t)$ is a martingale (see [19, Chapter 2, Section 6]). We conclude that

$$\langle M^g\rangle(t) = \sum_{j=1}^{N}\int_0^t g_{y_j}(Y^N(r))^2\,\mathrm{d}r.$$

□

Step 3b. We are now ready to derive the SDEs for the processes $Y_1^N,\dots,Y_N^N$. Define the stopping times τ_δ, $\delta>0$ by

$$\tau_\delta = \inf\{t \geq 0 : Y_i^N(t) - Y_{i-1}^N(t) \leq \delta \text{ for some } i\} \wedge \inf\{t \geq 0 : |Y_i^N(t)| \geq 1/\delta \text{ for some } i\}. \tag{3.17}$$

Our next aim is to derive the stochastic integral equations for the processes

$$\left(Y_j^N(t \wedge \tau_\delta)\right)_{j=1,\ldots,N}.$$

Let f_j, $j = 1, \ldots, N$, be an arbitrary function from $\mathcal{F}$ such that $f_j(\mathbf{y}) = y_j$ for $\mathbf{y}$ inside the box $|y_j| \leq 1/\delta$ and such that $\text{dist}(\mathbf{y}, \partial\overline{\mathcal{W}^N}) \geq \delta$. The results of Step 3a imply that the processes M^{f_j} are martingales. Note that the definition of stopping times τ_δ imply that on the time interval $[0, \tau_\delta]$ the processes M^{f_j} and M^{y_j} almost surely coincide. At this point we can use Lemma 3.6 to conclude that

$$Y_j^N(t \wedge \tau_\delta) - Y_j^N(0) - \int_0^{t\wedge\tau_\delta} \sum_{n\neq j} \frac{\theta}{Y_j^N(r) - Y_n^N(r)}\,\mathrm{d}r, \qquad j = 1, \ldots, N$$

are martingales with quadratic variations given by $t \wedge \tau_\delta$ and with the quadratic covariation between any two of them being zero. We may now apply the Martingale Representation Theorem in the form of [29, Theorem 3.4.2] to deduce the existence of independent standard Brownian motions $W_1, \ldots, W_N$ (possibly on an extension of the underlying probability space) such that

$$Y_j^N(t\wedge\tau_\delta) - Y_j^N(0) - \int_0^{t\wedge\tau_\delta} \sum_{n\neq j} \frac{\theta}{Y_j^N(r) - Y_n^N(r)}\,\mathrm{d}r = \int_0^{t\wedge\tau_\delta} \mathrm{d}W_j(s), \quad j = 1, \ldots, N. \tag{3.18}$$

To finish the proof of Theorem 3.2 it remains to observe that Proposition 3.1 implies

$$\lim_{\delta\to 0} \tau_\delta = \infty$$

with probability one.

Remark 3.7 An alternative way to derive the system of SDEs for the components of Y^N is to use [29, Chapter 5, Proposition 4.6]. We will employ this strategy in Sect. 5.3 due to the lack of a straightforward generalization of Lemma 3.6 to the multilevel setting.

3.4 Zero initial condition

A refinement of the proof of Theorem 3.2 involving Proposition 2.9 allows us to deal with the limiting process which is started from $0 \in \overline{\mathcal{W}^N}$.

Corollary 3.8 *Fix $\theta \geq 1$. In the notations of Theorem 3.2 and assuming the convergence of the initial conditions to $0 \in \overline{\mathcal{W}^N}$, the process X_{disc}^N converges in the limit*

$\varepsilon \downarrow 0$ in law on D^N to the $\beta = 2\theta$–Dyson Brownian motion started from $0 \in \overline{\mathcal{W}^N}$, that is, to the unique strong solution of (3.1) with $\beta = 2\theta$ and $Y(0) = 0 \in \overline{\mathcal{W}^N}$.

Proof Using Proposition 2.9 and arguing as in the proof of Theorem 3.2 one obtains the convergence of the rescaled versions of the process X^N_{disc} on every time interval $[t, \infty)$ with $t > 0$ to the solution of (3.1) starting according to the initial distribution of (2.10). Since (2.10) converges to the delta–function at the origin as $t \to 0$, we identify the limit points of the rescaled versions of X^N_{disc} with the solution of (3.1) started from $0 \in \overline{\mathcal{W}^N}$. □

4 Existence and uniqueness for multilevel DBM

The aim of this section is to prove an analogue of Proposition 3.1 for the multilevel Dyson Brownian motion.

Theorem 4.1 *For any $N \in \mathbb{N}$ and $\theta > 1$ (that is, $\beta = 2\theta > 2$), and for any initial condition $X(0)$ in the interior of $\overline{\mathcal{G}^N}$, the system of SDEs*

$$\mathrm{d}X_i^k(t) = \left(\sum_{m \neq i} \frac{1-\theta}{X_i^k(t) - X_m^k(t)} - \sum_{m=1}^{k-1} \frac{1-\theta}{X_i^k(t) - X_m^{k-1}(t)} \right) \mathrm{d}t + \mathrm{d}W_i^k, \quad 1 \leq i \leq k \leq N \tag{4.1}$$

with W_i^k, $1 \leq i \leq k \leq N$ being independent standard Brownian motions, possesses a unique weak solution taking values in the Gelfand–Tsetlin cone $\overline{\mathcal{G}^N}$.

Proof Given a stochastic process $X(t)$ taking values in $\overline{\mathcal{G}^N}$, for any fixed $\delta > 0$, let $\widehat{\tau}_\delta(X)$ denote

$$\widehat{\tau}_\delta[X] = \inf\{t \geq 0 \,:\, |X_i^k(t) - X_{i'}^{k'}(t)| \leq \delta,\ |k - k'| \leq 1\},$$

that is, the first time when two particles on adjacent levels are at the distance of at most δ. Further, we define the stopping time $\tau_\delta[X]$ as the first time when *three* particles on adjacent levels are at the distance of at most δ:

$$\tau_\delta[X] = \inf\Big\{ t \geq 0 \,:\, |X_i^k(t) - X_{i'}^{k'}(t)| \leq \delta,\ |X_i^k(t) - X_{i''}^{k''}(t)| \leq \delta, \\ \text{for } (k', i') \neq (k'', i'') \text{ such that } |k - k'| = |k - k''| = 1 \Big\}.$$

Figure 2 shows schematically the six possible triplets of nearby particles at time τ_δ.

The following proposition will be proved in Sect. 4.1.

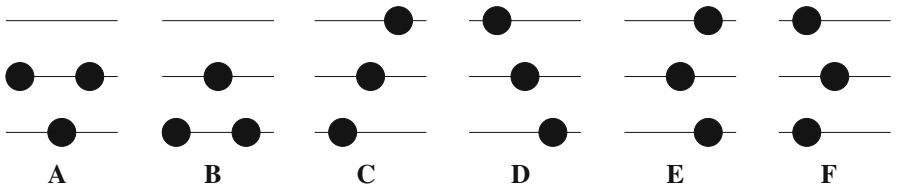

Fig. 2 Six possible triplets of nearby particles: one of these situations occurs at time τ_δ

Proposition 4.2 *For any $N \in \mathbb{N}$, $\delta > 0$ and $\theta > 1$ (that is, $\beta = 2\theta > 2$), and for any initial condition $X(0)$ in the interior of $\overline{\mathcal{G}^N}$ the system of stochastic integral equations*

$$X_i^k(t) - X_i^k(0) = \int_0^{t \wedge \tau_\delta[X]} \left(\sum_{m \neq i} \frac{1-\theta}{X_i^k(s) - X_m^k(s)} - \sum_{m=1}^{k-1} \frac{1-\theta}{X_i^k(t) - X_m^{k-1}(t)} \right) \mathrm{d}s + W_i^k(t \wedge \tau_\delta[X]), \ 1 \leq i \leq k \leq N, \tag{4.2}$$

with W_i^k, $1 \leq i \leq k \leq N$ being independent standard Brownian motions, possesses a unique weak solution.

In view of Proposition 4.2 we can consider a product probability space which supports independent weak solutions of (4.2) for all $\delta > 0$ and all initial conditions in the interior of $\overline{\mathcal{G}^N}$. Choosing a sequence δ_l, $l \in \mathbb{N}$ decreasing to zero, we can define on this space a process X such that the law of $X(t \wedge \tau_{\delta_1}[X])$, $t \geq 0$ coincides with the law of the solution of (4.2) with $\delta = \delta_1$ and initial condition $X(0)$, the law of $X((\tau_{\delta_1}[X] + t) \wedge \tau_{\delta_2}[X])$, $t \geq 0$ is given by the law of the solution of (4.2) with $\delta = \delta_2$ and initial condition $X(\tau_{\delta_1[X]})$ etc. The uniqueness part of Proposition 4.2 now shows that, for each $l \in \mathbb{N}$, the law of $X(t \wedge \tau_{\delta_l[X]})$, $t \geq 0$ is that of the weak solution of (4.2) with $\delta = \delta_l$. Since the paths of X are continuous by construction and hence $\lim_{l \to \infty} \tau_{\delta_l}[X] = \tau_0[X]$, we have constructed a weak solution of the system

$$X_i^k(t) - X_i^k(0) = \int_0^{t \wedge \tau_0[X]} \left(\sum_{m \neq i} \frac{1-\theta}{X_i^k(s) - X_m^k(s)} - \sum_{m=1}^{k-1} \frac{1-\theta}{X_i^k(t) - X_m^{k-1}(t)} \right) \mathrm{d}s + W_i^k(t \wedge \tau_0[X]), \ 1 \leq i \leq k \leq N \tag{4.3}$$

with W_i^k, $1 \leq i \leq k \leq N$ being independent standard Brownian motions as before. In addition, we note that the law of the solution to (4.3) is uniquely determined. Indeed, for any $\delta > 0$, the process X stopped at time τ_δ would give a solution to (4.2). Uniqueness of the latter for any $\delta > 0$ now readily implies the uniqueness of the weak solution to (4.3). At this point, Theorem 4.1 is a consequence of the following statement which will be proved in Sect. 4.2.

Proposition 4.3 *Suppose that $X(0)$ lies in the interior of the cone $\overline{\mathcal{G}^N}$ and let X be a solution to (4.3).*

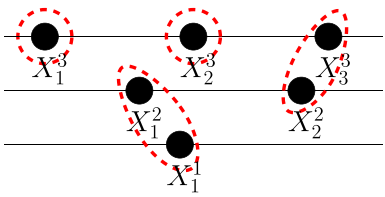

Fig. 3 A pair-partition with $N = 3$, two pairs and two singletons

(a) If $\theta > 1$, then almost surely $\tau_0[X] = \infty$.
(b) If $\theta \geq 2$, then almost surely $\widehat{\tau}_0[X] = \infty$.

4.1 Proof of Proposition 4.2

Our proof of Proposition 4.2 is based on a Girsanov change of measure that will dramatically simplify the SDE in consideration. We refer the reader to [29, Section 3.5] and [29, Section 5.3] for general information about Girsanov's theorem and weak solutions of SDEs.

We start with the uniqueness part. Fix $N \in \{1, 2, \ldots\}$, $\theta \geq 1$, $\delta > 0$, and let X be a solution of (4.2). Let $\mathcal{I}$ denote the set of $N(N+1)/2$ pairs (k, i), $k = 1, \ldots, N$, $i = 1, \ldots, k$ which represent different coordinates (particles) in the process X. We will subdivide $\mathcal{I}$ into disjoint singletons and pairs of neighboring particles, that is, pairs of the form $((k, i), (k-1, i))$ or $((k, i), (k-1, i-1))$. We call any such subdivision a *pair-partition* of $\mathcal{I}$. An example is shown in Fig. 3.

Lemma 4.4 *There exists a sequence of stopping times $0 = \sigma_0 \leq \sigma_1 \leq \sigma_2 \leq \cdots \leq \tau_\delta[X]$ and (random) pair–partitions $A_1, A_2, \ldots$ such that*

- *for any $n = 1, 2, \ldots$, any $\sigma_{n-1} \leq t < \sigma_n$, any two pairs (k, i), (k', i'), $1 \leq i \leq k \leq N$, $1 \leq i' \leq k' \leq N$, $|k - k'| \leq 1$, we have $|X_i^k(t) - X_{i'}^{k'}(t)| \geq \delta/2$ unless the pair $((k, i), (k', i'))$ is one of the pairs of the pair-partition A_n, and*
- *for any $n = 1, 2, \ldots$, either $\sigma_n = \tau_\delta$ or $|X_i^k(\sigma_{n+1}) - X_i^k(\sigma_n)| \geq \delta/2$ for some (k, i).*

Proof Define the (random) sets $\mathcal{B}_i^k$, $\mathcal{D}_i^k$ by setting

$$\mathcal{B}_i^k = \left\{0 \leq t \leq \tau_\delta[X] \mid |X_i^k(t) - X_i^{k-1}(t)| \leq \delta\right\},$$
$$\mathcal{D}_i^k = \left\{0 \leq t \leq \tau_\delta[X] \mid |X_i^k(t) - X_{i-1}^{k-1}(t)| \leq \delta\right\}.$$

Note that these sets are closed due to the continuity of the trajectories of X, which in turn is a consequence of (4.2). Define $A(t; \delta)$ as a pair-partition such that pair $((k, i), (k-1, i))$ belongs to $A(t; \delta)$ iff $t \in \mathcal{B}_i^k$ and pair $((k, i), (k-1, i-1))$ belongs to $A(t; \delta)$ iff $t \in \mathcal{D}_i^k$. Define $A(t, \delta/2)$ similarly. The definition of $\tau_\delta[X]$ implies that such pair-partitions $A(t; \delta/2) \subset A(t; \delta)$ are well-defined for any $0 \leq t \leq \tau_\delta[X]$.

Now, we define σ_n and A_n inductively. First, set $\sigma_0 = 0$. Further, for $n = 1, 2, \ldots$ let $A_n = A(\sigma_{n-1}; \delta)$ and set σ_n to be the minimal t satisfying $\tau_\delta[X] \geq t \geq \sigma_{n-1}$ and

such that $A(t;\delta/2)$ has a pair which $A(\sigma_{n-1};\delta)$ does not have. Since the sets $\mathcal{B}_i^k$, $\mathcal{D}_i^k$ are closed, either such t exists or no new pairs are added after time σ_{n-1} and up to time $\tau_\delta[X]$. In the latter case we set $\sigma_n = \tau_\delta$. □

Next, we fix a $T > 0$, set $I_n = [\sigma_{n-1}, \sigma_n)$, $n = 1, 2, \dots$ and apply a Girsanov change of measure (see e.g. [29, Theorem 5.1, Chapter 3] and note that Novikov's condition as in [29, Corollary 5.13, Chapter 3] is satisfied due to the boundedness of the integrand in the stochastic exponential) with a density of the form

$$\exp\left(\sum_{n=0}^{\infty}\sum_{1\le i\le k\le N}\left(\int_0^T b_{i,n}^k(t)\,\mathbf{1}_{I_n}(t)\,\mathrm{d}W_i^k(t) - \frac{1}{2}\int_0^T (b_{i,n}^k(t))^2\,\mathbf{1}_{I_n}(t)\,\mathrm{d}t\right)\right), \quad (4.4)$$

so that under the new measure $\widetilde{P}$ for every fixed k, i, n and $0 \le t \le T$:

$$\begin{aligned} &X_i^k(t\wedge\sigma_n) - X_i^k(t\wedge\sigma_{n-1}) \\ &= \begin{cases} \int\limits_{t\wedge\sigma_{n-1}}^{t\wedge\sigma_n}\left(\mathrm{d}\tilde{W}_i^k - \frac{(1-\theta)\,\mathrm{d}s}{X_i^k(s)-X_{i-1}^{k-1}(s)}\right), & \text{if } ((k,i),(k-1,i-1))\in A_n, \\ \int\limits_{t\wedge\sigma_{n-1}}^{t\wedge\sigma_n}\left(\mathrm{d}\tilde{W}_i^k - \frac{(1-\theta)\mathrm{d}s}{X_i^k(s)-X_i^{k-1}(s)}\right), & \text{if } ((k,i),(k-1,i))\in A_n, \\ \int\limits_{t\wedge\sigma_{n-1}}^{t\wedge\sigma_n}\mathrm{d}\tilde{W}_i^k, & \text{otherwise} \end{cases} \end{aligned} \quad (4.5)$$

where $\tilde{W}_i^k$, $1 \le i \le k \le N$ are independent standard Brownian motions under the measure $\widetilde{P}$. We claim that the solution of the resulting system of SDEs (4.5) is pathwise unique on $[0, \lim_{n\to\infty}\sigma_n)$ (that is, for any two strong solutions of (4.5) adapted to the same Brownian filtration, the quantities $\lim_{n\to\infty}\sigma_n$ for the two solutions will be the same with probability one and the trajectories of the two solutions on $[0, \lim_{n\to\infty}\sigma_n)$ will be identical with probability one). Indeed, on each time interval $\sigma_{n-1} \le t \le \sigma_n$ the system (4.5) splits into $|A_n|$ non-interacting systems of SDEs each of which consists of one equation

$$X_i^k(t\wedge\sigma_n) - X_i^k(t\wedge\sigma_{n-1}) = \int\limits_{t\wedge\sigma_{n-1}}^{t\wedge\sigma_n}\mathrm{d}\tilde{W}_i^k, \quad (4.6)$$

if (k,i) is a singleton in A_n, or of a system of two equations

$$\begin{aligned} &\left(X_i^k(t\wedge\sigma_n) - X_{i'}^{k-1}(t\wedge\sigma_n)\right) - \left(X_i^k(t\wedge\sigma_{n-1}) - X_{i'}^k(t\wedge\sigma_{n-1})\right) \\ &= \int\limits_{t\wedge\sigma_{n-1}}^{t\wedge\sigma_n}\left(\mathrm{d}\tilde{W}_i^k - \mathrm{d}\tilde{W}_{i'}^{k-1}\right) \\ &\quad - \int\limits_{t\wedge\sigma_{n-1}}^{t\wedge\sigma_n}\frac{(1-\theta)\,\mathrm{d}s}{X_i^k(s) - X_{i'}^{k-1}(s)}, \end{aligned} \quad (4.7)$$

$$X_{i'}^{k-1}(t\wedge\sigma_n)-X_i^k(t\wedge\sigma_{n-1})=\int\limits_{t\wedge\sigma_{n-1}}^{t\wedge\sigma_n}\mathrm{d}\tilde{W}_i^k, \tag{4.8}$$

if $((k,i),(k-1,i'))$ is a pair in A_n. Therefore, one can argue by induction over n and, once pathwise uniqueness of the triplet $((X(t\wedge\sigma_{n-1}):t\geq 0),\sigma_{n-1},A_{n-1})$ is established, appeal to the pathwise uniqueness for (4.6), (4.8) and (4.7) (the latter being the equation for the Bessel process of dimension $\theta>1$, see [42, Section 1, Chapter XI]) to deduce the pathwise uniqueness of the triplet $((X(t\wedge\sigma_n):t\geq 0),\sigma_n,A_n)$.

The SDEs in (4.5) also allow us to prove the following statement.

Lemma 4.5 *The identity* $\lim_{n\to\infty}\sigma_n=\tau_\delta[X]$ *holds with probability one.*

Proof It suffices to show that $\lim_{n\to\infty}\sigma_n\wedge T=\tau_\delta[X]\wedge T$ for any given $T>0$. Indeed, then

$$\lim_{n\to\infty}\sigma_n\geq\lim_{T\to\infty}\lim_{n\to\infty}\sigma_n\wedge T=\lim_{T\to\infty}\tau_\delta[X]\wedge T=\tau_\delta[X]$$

and $\lim_{n\to\infty}\sigma_n\leq\tau_\delta[X]$ holds by the definitions of the stopping times involved. If for some n we have $\tau_\delta[X]\wedge T=\sigma_n\wedge T$, then we are done. Otherwise, $\sigma_n<T$ for all n and the definition in Lemma 4.4 shows that $|X_i^k(\sigma_{n+1})-X_i^k(\sigma_n)|\geq\delta/2$ for some (k,i). In addition, (4.5) yields that, under the measure $\tilde{P}$, $|X_i^k(\sigma_{n+1})-X_i^k(\sigma_n)|$ is bounded above by the sum of absolute values of the increments of at most two Brownian motions and one Bessel process in time $(\sigma_{n+1}-\sigma_n)$. Since the trajectories of such processes are uniformly continuous on the compact interval $[0,T]$ with probability one, there exist two constants $c>0$ and $p>0$ such that $\tilde{P}(\sigma_{n+1}-\sigma_n>c)>p$. Consequently, σ_n/c stochastically dominates a binomial random variable $Bin(n,p)$. In view of the law of large numbers for the latter, this is a contradiction to $\sigma_n<T$ for all n. □

Now, we make a Girsanov change of measure back to the original probability measure and conclude that the joint law of $X(t\wedge\tau_\delta\wedge T),t\geq 0,\sigma_n\wedge T$ and $\tau_\delta\wedge T$ under the original probability measure is determined by such law under the measure $\widetilde{P}$ (the justification for this conclusion can be found for example in the proof of [29, Proposition 5.3.10]). Since the latter is uniquely defined (by the law of the solution to (4.5)), so is the former. Finally, since $T>0$ was arbitrary, we conclude that the joint law of $X(t\wedge\tau_\delta[X]),t\geq 0$ and $\tau_\delta[X]$ is uniquely determined.

To construct a weak solution to (4.2) we start with a probability space $(\Omega,\mathcal{F},\mathbb{P})$ that supports a family of independent standard Brownian motions $\tilde{W}_i^k$, $1\leq i\leq k\leq N$. In addition, we note (see [42, Section XI] for a proof) that to each pair of Brownian motions of the form $(\tilde{W}_i^k,\tilde{W}_{i-1}^{k-1})$ or $(\tilde{W}_i^k,\tilde{W}_i^{k-1})$ and all initial conditions we can associate the unique strong solutions of the SDEs

$$\mathrm{d}R_i^{k,-}(t)=\frac{\theta-1}{R_i^{k,-}(t)}\,\mathrm{d}t+\mathrm{d}\tilde{W}_i^k(t)-\mathrm{d}\tilde{W}_{i-1}^{k-1}(t), \tag{4.9}$$

$$\mathrm{d}R_i^{k,+}(t)=\frac{\theta-1}{R_i^{k,+}(t)}\,\mathrm{d}t+\mathrm{d}\tilde{W}_i^k(t)-\mathrm{d}\tilde{W}_i^{k-1}(t), \tag{4.10}$$

defined on the same probability space.

We will now construct an $N(N+1)/2$-dimensional process $X(t)$, $t \geq 0$, stopping times $\tau_\delta, \sigma_n, n = 0, 1, 2, \ldots$ and pair-partitions $A_n, n = 0, 1, 2, \ldots$ which satisfy the conditions of Lemma 4.4 and the system of equations (4.5).

The construction proceeds for each $\omega \in \Omega$ independently, and is inductive. If the initial condition $X(0)$ is such that $\tau_\delta[X] = 0$, then there is nothing to prove. Otherwise, we set $\sigma_0 = 0$ and $A_1 = A(0; \delta)$ (see the proof of Lemma 4.4 for the definition of $A(t; \delta)$). Next, we define $\hat{X}$ as the unique strong solution of

$$\hat{X}_i^k(t) - \hat{X}_i^k(0) = \begin{cases} \int_0^t \left(\mathrm{d}\tilde{W}_i^k - \frac{(1-\theta)\,\mathrm{d}s}{\hat{X}_i^k(s) - \hat{X}_{i-1}^{k-1}(s)} \right), & \text{if } ((k,i),(k-1,i-1)) \in A_1, \\ \int_0^t \left(\mathrm{d}\tilde{W}_i^k - \frac{(1-\theta)\mathrm{d}s}{\hat{X}_i^k(s) - \hat{X}_i^{k-1}(s)} \right), & \text{if } ((k,i),(k-1,i)) \in A_1, \\ \int_0^t \mathrm{d}\tilde{W}_i^k, & \text{otherwise,} \end{cases} \tag{4.11}$$

with initial condition $\hat{X}(0) = X(0)$.

Now, we can define σ_1 as in Lemma 4.4, but with $\hat{X}(t)$ instead of $X(t)$. After this, we set $X(t)$ to be equal to $\hat{X}(t)$ on the time interval $[0, \sigma_1]$. We further define $A_2 = A(\sigma_1; \delta)$ and repeat the above procedure to define $X(t)$ on the time interval $[\sigma_1, \sigma_2]$. Iterating this process we can define $X(t)$ up to time $\tau_\delta[X]$ thanks to Lemma 4.5. We extend it to all $t \geq 0$ by setting $X_i^k(t) = X_i^k(\tau_\delta[X])$ for $t > \tau_\delta[X]$.

Next, we apply the Girsanov Theorem as in the uniqueness part to conclude that, for each $T > 0$, there exists a probability measure $\mathbb{Q}_T$ which is absolutely continuous with respect to $\mathbb{P}$ and such that the representation

$$X_i^k(t \wedge T) - X_i^k(0) = \int_0^{t \wedge T \wedge \tau_\delta[X]} \left(\sum_{m \neq i} \frac{1-\theta}{X_i^k(s) - X_m^k(s)} - \sum_{m=1}^{k-1} \frac{1-\theta}{X_i^k(s) - X_m^{k-1}(s)} \right) \mathrm{d}s + W_i^k(t \wedge T \wedge \tau_\delta[X]), \quad 1 \leq i \leq k \leq N$$

holds with W_i^k, $1 \leq i \leq k \leq N$ being independent standard Brownian motions under $\mathbb{Q}_T$.

Finally, replacing T by a sequence $T_n \uparrow \infty$ and using the Kolmogorov Extension Theorem (see e.g. [28, Theorem 6.16] and note that the consistency condition is satisfied due to the uniqueness of the solution to (4.2)), we deduce the existence of processes X_i^k, $1 \leq i \leq k \leq N$ defined on a suitable probability space which solve (4.2).

4.2 Proof of Proposition 4.3

We start with a version of Feller's test for explosions that will be used below (see e.g. [29, Section 5.5.C] and the references therein for related results).

Lemma 4.6 *Let Z be a one-dimensional continuous semimartingale satisfying $Z(0) > 0$ and*

$$\forall\, 0 \le t_1 < t_2 : \quad Z(t_2) - Z(t_1) = b(t_2 - t_1) + M(t_2) - M(t_1) \tag{4.12}$$

with a constant $b > 0$ and a local martingale M. If the quadratic variation of M satisfies

$$\forall\, 0 \le t_1 < t_2 : \langle M \rangle (t_2) - \langle M \rangle (t_1) \le 2b \int_{t_1}^{t_2} Z(t)\, \mathrm{d}t, \tag{4.13}$$

then the process Z does not reach zero in finite time with probability one.

Proof We fix two constants $0 < r_1 < Z(0) < R_1 < \infty$ and let τ_{r_1,R_1} be the first time that Z reaches r_1 or R_1. Next, we apply Itô's formula (see e.g. [29, Section 3.3.A]) to obtain

$$\begin{aligned} \ln Z(t \wedge \tau_{r_1,R_1} \wedge \zeta) - \ln Z(0) = &\int_0^{t \wedge \tau_{r_1,R_1} \wedge \zeta} \left(\frac{b\, \mathrm{d}s}{Z(s)} - \frac{\mathrm{d}\langle M \rangle (s)}{2\, Z(s)^2} \right) \\ &+ \int_0^{t \wedge \tau_{r_1,R_1} \wedge \zeta} \frac{\mathrm{d}M(s)}{Z(s)} \end{aligned} \tag{4.14}$$

for any stopping time ζ. By (4.13), the first integral in (4.14) takes non-negative values. Hence, picking a localizing sequence of stopping times $\zeta = \zeta_m$ for the local martingale given by the second integral in (4.14), taking the expectation in (4.14) and passing to the limit $m \to \infty$, we obtain

$$\mathbb{E}\left[\ln Z(t \wedge \tau_{r_1,R_1}) \right] \ge \ln Z(0).$$

Now, Fatou's Lemma and $Z(t \wedge \tau_{r_1,R_1}) \le R_1$ yield the chain of estimates

$$\ln Z(0) \le \mathbb{E}\left[\limsup_{t \to \infty} \ln Z(t \wedge \tau_{r_1,R_1}) \right] \le p_{r_1} \ln r_1 + (1 - p_{r_1}) \ln R_1$$

where $p_{r_1} = P(\limsup_{t\to\infty} \ln Z(t \wedge \tau_{r_1,R_1}) = r_1)$. Consequently,

$$p_{r_1} \le \frac{\ln R_1 - \ln Z(0)}{\ln R_1 - \ln r_1}.$$

The lemma now follows by taking the limit $r_1 \downarrow 0$. □

We will now show that $\tau_0[X] = \infty$ for the solution of (4.2) with initial condition $X(0)$ such that $\tau_0[X] > 0$ (in particular, this includes the case that $X(0)$ belongs to the interior of $\overline{\mathcal{G}^N}$). Recall that τ_0 was defined as the first time at which one of the events in Fig. 2 with $\delta = 0$ occurs. We will show that neither of the cases $A - F$ in Fig. 2 can occur in finite time. We will argue by induction over k to prove that none of these events can happen on the first k levels for $k = 1, 2, \ldots, N$.

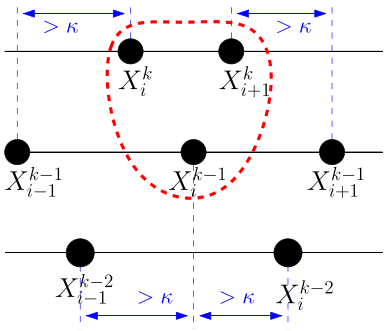

Fig. 4 Decoupling of the particles X_i^k, X_{i+1}^k, X_i^{k-1}

First, we focus on the cases A and B.

Lemma 4.7 *An event of the form $X_i^k(t) = X_{i+1}^k(t)$ cannot occur in finite time without one of the events*

$$X_i^{k-1}(t) - X_{i-1}^{k-2}(t) = 0, \quad X_i^{k-2}(t) - X_i^{k-1}(t) = 0 \tag{4.15}$$

occurring at the same time.

Proof If the statement of the lemma was not true, then the continuity of the paths of the particles would allow us to find stopping times σ, σ' similar to the ones introduced in Sect. 4.1 and a real number $\kappa > 0$ such that $\sigma < \sigma'$ with probability one, the spacings in (4.15) are at least κ during the time interval $[\sigma, \sigma']$ and the event $X_i^k(t) = X_{i+1}^k(t)$ occurs for the first time at time σ'. Moreover, the interlacing condition and the induction hypothesis imply together that $[\sigma, \sigma']$ and κ can be chosen such that the spacings

$$X_i^k - X_{i-1}^{k-1}, \quad X_{i+1}^{k-1} - X_{i+1}^k$$

do not fall below κ on $[\sigma, \sigma']$ (otherwise at least one of the events $X_{i-1}^{k-1}(t) = X_i^{k-1}(t)$ or $X_i^{k-1}(t) = X_{i+1}^{k-1}(t)$ would have occurred at time σ' in contradiction to the induction hypothesis). The described inequalities are shown in Fig. 4.

Now, making a Girsanov change of measure similar to the one in Sect. 4.1, we can decouple the particles X_i^k, X_{i+1}^k, X_i^{k-1} from the rest of the particle system, thus reducing their dynamics on the time interval $[\sigma, \sigma']$ to the two-level dynamics:

$$\begin{aligned}
X_i^k(t \wedge \sigma') - X_i^k(t \wedge \sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \left(\mathrm{d}\tilde{W}_i^k + \frac{(1-\theta)\,\mathrm{d}s}{X_i^k(s) - X_{i+1}^k(s)} - \frac{(1-\theta)\,\mathrm{d}s}{X_i^k(s) - X_i^{k-1}(s)} \right), \\
X_{i+1}^k(t \wedge \sigma') - X_{i+1}^k(t \wedge \sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \left(\mathrm{d}\tilde{W}_{i+1}^k + \frac{(1-\theta)\,\mathrm{d}s}{X_{i+1}^k(s) - X_i^k(s)} - \frac{(1-\theta)\,\mathrm{d}s}{X_{i+1}^k(s) - X_i^{k-1}(s)} \right), \\
X_i^{k-1}(t \wedge \sigma') - X_i^{k-1}(t \wedge \sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \mathrm{d}\tilde{W}_i^{k-1},
\end{aligned}$$

with $\tilde{W}_i^k$, $\tilde{W}_{i+1}^k$, $\tilde{W}_i^{k-1}$ being standard Brownian motions under the new probability measure. Next, we note that the process $X_{i+1}^k - X_i^k$ hits zero if and only if the process

$$Z = \frac{1}{2}\Big((X_i^{k-1} - X_i^k)^2 + (X_{i+1}^k - X_i^{k-1})^2\Big) \tag{4.16}$$

hits zero and in this case both events occur at the same time. Moreover, applying Itô's formula (see e.g. [29, Section 3.3.A]) and simplifying the result we obtain

$$\begin{aligned} Z(t\wedge\sigma') - Z(t\wedge\sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \Big((1+\theta)\,\mathrm{d}s + (X_i^{k-1} - X_i^k)\,\mathrm{d}(\tilde{W}_i^{k-1} - \tilde{W}_i^k) \\ &\quad + (X_{i+1}^k - X_i^{k-1})\,\mathrm{d}(\tilde{W}_{i+1}^k - \tilde{W}_i^{k-1})\Big) =: \int_{t\wedge\sigma}^{t\wedge\sigma'} \big((1+\theta)\,\mathrm{d}s + \mathrm{d}M\big) \end{aligned}$$

where M is a local martingale whose quadratic variation process satisfies

$$\begin{aligned} \langle M\rangle(t\wedge\sigma') - \langle M\rangle(t\wedge\sigma) = \int_{t\wedge\sigma}^{t\wedge\sigma'} \Big(&2\,(X_i^{k-1} - X_i^k)^2 + 2\,(X_{i+1}^k - X_i^{k-1})^2 \\ &-(X_i^{k-1} - X_i^k)(X_{i+1}^k - X_i^{k-1})\Big)\,\mathrm{d}s, \quad t \ge 0. \end{aligned}$$

We can now define the (random) time change

$$s(t) = \inf\left\{ s \ge 0 : \int_{s\wedge\sigma}^{s\wedge\sigma'} (1+\theta)\,\mathrm{d}u = t \right\}, \quad 0 \le t \le \int_{\sigma}^{\sigma'} (1+\theta)\,\mathrm{d}u$$

and rewrite the stochastic integral equation for Z as

$$Z(s(t_2)) - Z(s(t_1)) = (t_2 - t_1) + M(s(t_2)) - M(s(t_1)), \quad 0 \le t_1 \le t_2 \le \int_{\sigma}^{\sigma'} (1+\theta)\,\mathrm{d}u.$$

A standard application of the Optional Sampling Theorem (see e.g. the proof of [29, Theorem 4.6, Chapter 3] for a similar argument) shows that the process $M(s(t)), t \ge 0$ is a local martingale in its natural filtration. In addition,

$$\begin{aligned} \langle M\rangle(s(t_2)) - \langle M\rangle(s(t_1)) &\le \int_{t_1}^{t_2} \frac{2\,(X_i^{k-1}(s(t)) - X_i^k(s(t)))^2 + 2\,(X_{i+1}^k(s(t)) - X_i^{k-1}(s(t)))^2}{1+\theta}\,\mathrm{d}t \\ &\le 2\int_{t_1}^{t_2} Z(s(t))\,\mathrm{d}t, \qquad 0 \le t_1 \le t_2 \le \int_{\sigma}^{\sigma'} (1+\theta)\,\mathrm{d}u. \end{aligned}$$

It follows that the process

$$\tilde{Z}(t) = \begin{cases} Z(s(t)) & \text{if } t \in \big[0, \int_{\sigma}^{\sigma'} (1+\theta)\,\mathrm{d}u\big] \\ Z\big(s\big(\int_{\sigma}^{\sigma'} (1+\theta)\,\mathrm{d}u\big)\big) + \big(t - \int_{\sigma}^{\sigma'} (1+\theta)\,\mathrm{d}u\big) & \text{if } t \in \big(\int_{\sigma}^{\sigma'} (1+\theta)\,\mathrm{d}u, \infty\big) \end{cases}$$

falls into the framework of Lemma 4.6. Consequently, the original process $Z(t), t \geq 0$ does not reach zero on the time interval $[\sigma, \sigma']$ with probability one. Using Girsanov's Theorem again (now to go back to the original probability measure) we conclude that $X_{i+1}^k - X_i^k$ does not hit zero on $[\sigma, \sigma']$ under the original probability measure, which is the desired contradiction. □

Next, we study the events $C - F$ in Fig. 2. All of them can be dealt in exactly the same manner (in particular, using a Lyapunov function of the same form) and we will only show the following:

Lemma 4.8 *The event*

$$X_i^k(t) = X_i^{k-1}(t) = X_i^{k-2}(t) \tag{4.17}$$

cannot occur in finite time.

Proof To show the non-occurrence of the event in (4.17), we again argue by induction over k and by contradiction. Assuming that the event in (4.17) occurs in finite time, we may invoke the induction hypothesis and Lemma 4.7 to find a random time interval $[\sigma, \sigma']$ with σ, σ' being stopping times and a real number $\kappa > 0$ such that the event in (4.17) occurs for the first time at σ', and either $X_i^k(\sigma') = X_{i+1}^k(\sigma') = X_i^{k-1}(\sigma') = X_i^{k-2}(\sigma')$ and the spacings $X_i^k - X_{i-1}^{k-1}$, $X_i^{k-1} - X_{i-1}^{k-2}$, $X_i^{k-2} - X_{i-1}^{k-3}$, $X_i^{k-3} - X_i^{k-2}$, $X_{i+1}^{k-1} - X_i^{k-2}$, $X_{i+1}^{k-1} - X_{i+1}^k$ are bounded below by κ on $[\sigma, \sigma']$; or $X_i^k(\sigma') = X_i^{k-1}(\sigma') = X_i^{k-2}(\sigma')$ and the spacings $X_i^k - X_{i-1}^{k-1}$, $X_i^{k-1} - X_{i-1}^{k-2}$, $X_i^{k-2} - X_{i-1}^{k-3}$, $X_i^{k-3} - X_i^{k-2}$, $X_{i+1}^{k-1} - X_i^{k-2}$, $X_{i+1}^k - X_i^{k-1}$ are bounded below by κ on $[\sigma, \sigma']$. Figure 5 shows schematic illustrations of these two cases. In the first case, we can make a Girsanov change of measure such that under the new measure the evolution of the particles X_i^k, X_{i+1}^k, X_i^{k-1}, X_i^{k-2} decouples from the rest of the particle system

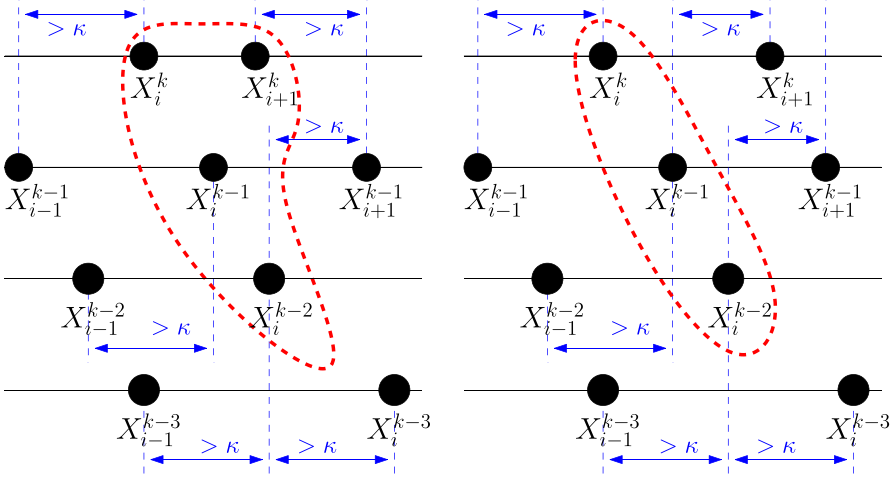

Fig. 5 Decoupling of the particles X_i^k, X_{i+1}^k, X_i^{k-1}, X_i^{k-2} (*left panel*) and the particles X_i^k, X_i^{k-1}, X_i^{k-2} (*right panel*)

on the time interval $[\sigma, \sigma']$. Similarly, in the second case, we can apply a Girsanov change of measure such that under the new measure the dynamics of the particles X_i^k, X_i^{k-1}, X_i^{k-2} decouples from the dynamics of the rest of the particle configuration.

We only treat the first of the two cases in detail (the second case can be dealt with by proceeding as below with the same definitions for R, S, and Z redefined to $\frac{1}{2}(R^2 + S^2 + 2\,R\,S)$). In the first case, the decoupled particles satisfy under the new measure:

$$\begin{aligned}
X_i^k(t\wedge\sigma') - X_i^k(t\wedge\sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \Big(\mathrm{d}\tilde{W}_i^k + \frac{(1-\theta)\,\mathrm{d}s}{X_i^k(s) - X_{i+1}^k(s)} - \frac{(1-\theta)\,\mathrm{d}s}{X_i^k(s) - X_i^{k-1}(s)}\Big),\\
X_{i+1}^k(t\wedge\sigma') - X_{i+1}^k(t\wedge\sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \Big(\mathrm{d}\tilde{W}_{i+1}^k + \frac{(1-\theta)\,\mathrm{d}s}{X_{i+1}^k(s) - X_i^k(s)} - \frac{(1-\theta)\,\mathrm{d}s}{X_{i+1}^k(s) - X_i^{k-1}(s)}\Big),\\
X_i^{k-1}(t\wedge\sigma') - X_i^{k-1}(t\wedge\sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \Big(\mathrm{d}\tilde{W}_i^{k-1} - \frac{(1-\theta)\,\mathrm{d}s}{X_i^{k-1}(s) - X_i^{k-2}(s)}\Big),\\
X_i^{k-2}(t\wedge\sigma') - X_i^{k-2}(t\wedge\sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \mathrm{d}\tilde{W}_i^{k-2}
\end{aligned}$$

where $\tilde{W}_i^k$, $\tilde{W}_{i+1}^k$, $\tilde{W}_i^{k-1}$, $\tilde{W}_i^{k-2}$ are independent standard Brownian motions under the new measure. Next, we set $R := X_i^{k-1} - X_i^k$, $S := X_i^{k-2} - X_i^{k-1}$, $U := X_{i+1}^k - X_i^{k-1}$, $B_1 := \tilde{W}_i^{k-1} - \tilde{W}_i^k$, $B_2 := \tilde{W}_i^{k-2} - \tilde{W}_i^{k-1}$, $B_3 := \tilde{W}_{i+1}^k - \tilde{W}_i^{k-1}$ and define $Z := \frac{1}{2}(R^2 + S^2 + U^2 + 2\,R\,U)$. Applying Itô's formula (see e.g. [29, Section 3.3.A]) and simplifying, we obtain

$$\begin{aligned}
Z(t\wedge\sigma') - Z(t\wedge\sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} \left(4 + \theta + (\theta-1)\left(\frac{U(s)}{R(s)} + \frac{R(s)}{U(s)}\right)\right)\mathrm{d}s\\
&\quad + (R+U)\mathrm{d}(B_1+B_3) + S\,\mathrm{d}B_2 =: \int_{t\wedge\sigma}^{t\wedge\sigma'} \Big(D(s)\,\mathrm{d}s + \mathrm{d}M\Big)
\end{aligned}$$

where M is a local martingale whose quadratic variation process satisfies

$$\begin{aligned}
\langle M\rangle(t\wedge\sigma') - \langle M\rangle(t\wedge\sigma) &= \int_{t\wedge\sigma}^{t\wedge\sigma'} 2\,(R^2 + S^2 + U^2 + 2\,R\,U)\,\mathrm{d}s\\
&= \int_{t\wedge\sigma}^{t\wedge\sigma'} 4\,Z\,\mathrm{d}s, \quad t \geq 0.
\end{aligned}$$

Next, we introduce the time change

$$s(t) = \inf\left\{ s \geq 0 : \int_{t\wedge\sigma}^{t\wedge\sigma'} D(u)\,\mathrm{d}u = t \right\}, \quad 0 \leq t \leq \int_{\sigma}^{\sigma'} D(u)\,\mathrm{d}u$$

and rewrite the latter stochastic integral equation for Z as

$$Z(s(t_2)) - Z(s(t_1)) = (t_2 - t_1) + M(s(t_2)) - M(s(t_1)), \quad 0 \le t_1 \le t_2 \le \int_\sigma^{\sigma'} D(u)\,\mathrm{d}u.$$

At this point, a routine application of the Optional Sampling Theorem (see e.g. the proof of [29, Theorem 4.6, Chapter 3] for an argument of this type) shows that $M(s(t))$, $t \ge 0$ is a local martingale in its natural filtration. In addition,

$$\langle M\rangle(s(t_2)) - \langle M\rangle(s(t_1)) = \int_{t_1}^{t_2} \frac{4\,Z(s(t))}{D(s(t))}\,\mathrm{d}t \le 2\int_{t_1}^{t_2} Z(s(t))\,\mathrm{d}t, \quad 0 \le t_1 \le t_2 \le \int_\sigma^{\sigma'} D(u)\,\mathrm{d}u.$$

Hence, the process

$$\tilde{Z}(t) = \begin{cases} Z(s(t)) & \text{if} \quad t \in \left[0, \int_\sigma^{\sigma'} D(u)\,\mathrm{d}u\right] \\ Z\Big(s\Big(\int_\sigma^{\sigma'} D(u)\,\mathrm{d}u\Big)\Big) + \Big(t - \int_\sigma^{\sigma'} D(u)\,\mathrm{d}u\Big) & \text{if} \quad t \in \Big(\int_\sigma^{\sigma'} D(u)\,\mathrm{d}u, \infty\Big) \end{cases}$$

falls into the setting of Lemma 4.6. The result of that lemma implies that the original process $Z(t)$, $t \ge 0$ does not hit zero on the time interval $[\sigma, \sigma']$ with probability one. Changing the measure back to the original probability measure by a suitable application of Girsanov's Theorem we conclude that the same is true under the original probability measure. This is the desired contradiction. □

Putting together Lemmas 4.7 and 4.8 we deduce that $\tau_0[X] = \infty$ for all $\theta > 1$.

Finally, for $\theta \ge 2$ and any $T > 0$, we have shown that the law of our process up to time $\tau_\delta[X] \wedge T$ is absolutely continuous with respect to the law of a process comprised of a number of Brownian motions and Bessel processes of dimension θ. The definition of the latter implies that two of its components can collide only if the corresponding Bessel process hits zero. However, it is well-known that the Bessel process of dimension $\theta \ge 2$ does not reach zero with probability one (see e.g. [42, Chapter XI, Section 1]). It follows that $\widehat{\tau}_\delta[X] \ge \lim_{T\to\infty} \tau_\delta[X] \wedge T = \tau_\delta[X]$. Passing to the limit $\delta \downarrow 0$, we conclude that $\widehat{\tau}_0[X] = \infty$.

5 Convergence to multilevel Dyson Brownian Motion

In this section we study the diffusive scaling limit of the multilevel process X^{multi}_{disc} of Definition 2.27. We start by formulating our main results.

We fix $\theta > 0$, let $\varepsilon > 0$ be a small parameter and define the $\frac{N(N+1)}{2}$-dimensional stochastic process $Y^{mu}_\varepsilon = ((Y^{mu}_\varepsilon)^k_i : 1 \le i \le k \le N)$ by

$$(Y^{mu}_\varepsilon)^k_i(t) = \varepsilon^{1/2}\Big((X^{multi}_{disc})^k_{N+1-i}\Big(\frac{t}{\theta\varepsilon}\Big) - \frac{t}{\varepsilon}\Big), \quad t \ge 0, \qquad 1 \le i \le k \le N$$

where $(X^{multi}_{disc})^k_i$, $1 \le i \le k \le N$ are the coordinate processes of X^{multi}_{disc}. Here, in contrast to Definition 2.27, we allow X^{multi}_{disc} (or, equivalently, Y^{mu}_ε) to start from an arbitrary initial condition, in particular, one that depends on ε. In addition, we use the

notation $D^{N(N+1)/2} = D([0,\infty), \mathbb{R}^{N(N+1)/2})$ for the space of right-continuous paths with left limits taking values in $\mathbb{R}^{N(N+1)/2}$, endowed with the Skorokhod topology.

Theorem 5.1 *Let $\theta > 0$ and suppose that the family of initial conditions $Y_\varepsilon^{mu}(0)$, $\varepsilon \in (0,1)$ is tight on $\mathbb{R}^{N(N+1)/2}$. Then the family Y_ε^{mu}, $\varepsilon \in (0,1)$ is tight on $D^{N(N+1)/2}$.*

We defer the proof of Theorem 5.1 to Sect. 5.2.

Next, we let Y^{mu} be an arbitrary limit point as $\varepsilon \downarrow 0$ of the tight family Y_ε^{mu}, $\varepsilon \in (0,1)$. For $\theta \geq 2$, we can uniquely identify the limit point with the solution of (4.1), thus, obtaining the diffusive scaling limit. For $\theta \in \left[\frac{1}{2}, 2\right)$, we give a partial result towards such an identification.

Theorem 5.2 *Let $\theta \geq 2$ (that is, $\beta = 2\theta \geq 4$) and suppose that the initial conditions $Y_\varepsilon^{mu}(0)$, $\varepsilon \in (0,1)$ converge as $\varepsilon \downarrow 0$ in distribution to a limit $Y^{mu}(0)$ which takes values in the interior of the Gelfand–Tsetlin cone $\overline{\mathcal{G}^N}$ with probability one. Then the family Y_ε^{mu}, $\varepsilon \in (0,1)$ converges as $\varepsilon \downarrow 0$ in distribution in $D^{N(N+1)/2}$ to the unique solution of the system of SDEs*

$$\mathrm{d}(Y^{mu})_i^k = \left(\sum_{m \neq i} \frac{1-\theta}{(Y^{mu})_i^k - (Y^{mu})_m^k} - \sum_{m=1}^{k-1} \frac{1-\theta}{(Y^{mu})_i^k - (Y^{mu})_m^{k-1}} \right) \mathrm{d}t + \mathrm{d}W_i^k, \quad 1 \leq i \leq k \leq N$$

started from $Y^{mu}(0)$ and where W_i^k, $1 \leq i \leq k \leq N$ are independent standard Brownian motions.

We give the proof of Theorem 5.2 in Sect. 5.3. We expect Theorem 5.2 to be valid for all $\theta > 1$, but we are not able to prove this generalization.

Theorem 5.3 *Suppose that the initial conditions $Y_\varepsilon^{mu}(0)$, $\varepsilon \in (0,1)$ converge as $\varepsilon \downarrow 0$ in distribution to a limit $Y^{mu}(0)$ which takes values in the interior of the Gelfand–Tsetlin cone $\overline{\mathcal{G}^N}$ with probability one. In addition, suppose that the distribution of $Y^{mu}(0)$ is θ-Gibbs in the sense of Definition 2.15.*

(a) If $\theta \geq \frac{1}{2}$ (that is, $\beta = 2\theta \geq 1$), then the restriction of Y^{mu} to level N, that is the process $((Y^{mu})_1^N, \ldots, (Y^{mu})_N^N)$, is a (2θ)-Dyson Brownian motion:

$$\mathrm{d}(Y^{mu})_i^N(t) = \sum_{m \neq i} \frac{\theta}{(Y^{mu})_i^N(t) - (Y^{mu})_m^N(t)} \mathrm{d}t + \mathrm{d}W_i, \quad 1 \leq i \leq N$$

with W_i^k, $1 \leq i \leq k \leq N$ being independent standard Brownian motions.

(b) For any $\theta > 0$ and any fixed $t > 0$, the distribution of $Y^{mu}(t)$ is θ-Gibbs.

We expect the first part of Theorem 5.3 to be valid for all $\theta > 0$, but we are currently not able to prove this.

Proof of Theorem 5.3 The theorem follows from a combination of Propositions 2.29, 2.16 and Theorem 3.2. □

Corollary 5.4 *Take any $\theta > 0$ and suppose that $Y_\varepsilon^{mu}(0) = 0 \in \mathbb{R}^{N(N+1)/2}$, $\varepsilon \in (0,1)$. Then, for any $t \geq 0$, the distribution of $Y^{mu}(t)$ is given by the Hermite $\beta = 2\theta$ corners process of variance t (see Definition 1.1).*

Remark 5.5 Since for any $t > 0$, the Hermite $\beta = 2\theta$ corners process of variance t is supported by the interior of the Gelfand–Tsetlin cone $\overline{\mathcal{G}^N}$, it follows that Theorem 5.2 (for $\theta \geq 2$) can be applied in this case as well. Consequently, for $\theta \geq 2$, the process Y^{mu} started from the zero initial condition is a diffusion that combines Dyson Brownian motions and corners processes into a single picture as desired.

Proof of Corollary 5.4 The corollary is a consequence of Proposition 2.29 and Corollary 2.17. □

The rest of this section is devoted to the proofs of Theorems 5.1 and 5.2. Our proof strategy is similar to that in the proof of Theorem 3.2: In Sect. 5.1 we analyze the asymptotic behavior of the jump rates of the processes Y_ε^{mu}, $\varepsilon \in (0,1)$, in Sect. 5.2 we use this asymptotics to prove that the family Y_ε^{mu}, $\varepsilon \in (0,1)$ is tight, and in Sect. 5.3 we deduce the SDE (4.1) for subsequential limits as $\varepsilon \downarrow 0$ of this family when $\theta \geq 2$. We omit the details in the parts that are parallel to the arguments of Sect. 3.

5.1 Step 1: Rates

We start by noting that, for each $\varepsilon \in (0,1)$, Y_ε^{mu} is a continuous time Markov process with state space $\overline{\mathcal{G}^N}$, a (constant) drift of $-\varepsilon^{-1/2}$ in each coordinate and jump rates

$$q_\varepsilon^{mu}(y, y', t) = \frac{1}{\theta\varepsilon}\, q_{\frac{t}{\theta}\varepsilon^{-1} + \hat{y}\,\varepsilon^{-1/2} \to \frac{t}{\theta}\varepsilon^{-1} + \hat{y'}\,\varepsilon^{-1/2}}$$

where $\hat{y}$, $\hat{y'}$ are the vectors obtained from y, y' by reordering the coordinates on each level in decreasing order, and intensities q are given by (2.23). Write $y' \approx_\varepsilon y$ for vectors $y, y' \in \overline{\mathcal{G}^N}$ such that y' can be obtained from y by increasing one coordinate (say, y_i^k) by $\varepsilon^{1/2}$, and, if necessary, by increasing other coordinates as well to preserve the interlacing condition (in the sense of the push interaction as explained after (2.23)). Clearly, $q_\varepsilon^{mu}(y, y', t) = 0$ unless $y' \approx_\varepsilon y$. As we will see, in fact, $q_\varepsilon^{mu}(y, y', t)$ does not depend on t.

Lemma 5.6 *For any sequence of vectors $y' \approx_\varepsilon y$ and any fixed $k \in \{1, \dots, N\}$, $i \in \{1, 2, \dots, k\}$ as above one has the following $\varepsilon \downarrow 0$ asymptotics:*

$$q_\varepsilon^{mu}(y, y', t) = \varepsilon^{-1} + \varepsilon^{-1/2}\Big(\sum_{m \neq i} \frac{1-\theta}{y_i^k - y_m^k} - \sum_{m=1}^{k-1} \frac{1-\theta}{y_i^k - y_m^{k-1}}\Big) + O(1) \qquad (5.1)$$

with a uniform $O(1)$ remainder on compact subsets of the open Gelfand–Tsetlin cone $\mathcal{G}^N$.

Proof We write $\hat{y}_i^k$ for y_{k+1-i}^k and $(\hat{y}')_i^k$ for $(y')_{k+1-i}^k$. Using (2.23), Proposition 2.4 and arguing as in Lemma 3.4 we rewrite $q_\varepsilon^{mu}(y, y', t)$ as

$$
\begin{aligned}
&\varepsilon^{-1}\prod_{l=1}^{i-1}\frac{(\hat{y}_l^k-\hat{y}_i^k)\,\varepsilon^{-1/2}-1+\theta\,(i-l+1)}{(\hat{y}_l^k-\hat{y}_i^k)\,\varepsilon^{-1/2}-1+\theta\,(i-l)}\cdot\frac{(\hat{y}_l^k-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-l-1)}{(\hat{y}_l^k-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-l)}\\
&\quad\times\prod_{1\le m\le n\le k-1}\frac{\big(((\hat{y}')_m^{k-1}-(\hat{y}')_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+\theta\big)_{((\hat{y}')_n^{k-1}-(\hat{y}')_{n+1}^k)\,\varepsilon^{-1/2}}}{\big(((\hat{y}')_m^{k-1}-(\hat{y}')_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+1\big)_{((\hat{y}')_n^{k-1}-(\hat{y}')_{n+1}^k)\,\varepsilon^{-1/2}}}\\
&\quad\times\frac{\big(((\hat{y}')_m^{k}-(\hat{y}')_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+1\big)_{((\hat{y}')_n^{k-1}-(\hat{y}')_{n+1}^k)\,\varepsilon^{-1/2}}}{\big(((\hat{y}')_m^{k}-(\hat{y}')_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+\theta\big)_{((\hat{y}')_n^{k-1}-(\hat{y}')_{n+1}^k)\,\varepsilon^{-1/2}}}\\
&\quad\times\prod_{1\le m\le n\le k-1}\frac{\big((y_m^{k-1}-y_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+1\big)_{(y_n^{k-1}-y_{n+1}^k)\,\varepsilon^{-1/2}}}{\big((y_m^{k-1}-y_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+\theta\big)_{(y_n^{k-1}-y_{n+1}^k)\,\varepsilon^{-1/2}}}\\
&\quad\times\frac{\big((y_m^{k}-y_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+\theta\big)_{(y_n^{k-1}-y_{n+1}^k)\,\varepsilon^{-1/2}}}{\big((y_m^{k}-y_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-m)+1\big)_{(y_n^{k-1}-y_{n+1}^k)\,\varepsilon^{-1/2}}}.
\end{aligned}
$$

Using the fact that y and y' differ only in one coordinate, we can simplify the latter expression to

$$
\begin{aligned}
&\varepsilon^{-1}\prod_{m=1}^{i-1}\frac{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}-1+\theta\,(i-m+1)}{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}-1+\theta\,(i-m)}\,\frac{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-m-1)}{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-m)}\\
&\quad\times\prod_{m=1}^{i-1}\frac{(\hat{y}_m^{k-1}-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-1-m)}{(\hat{y}_m^{k-1}-\hat{y}_i^k)\,\varepsilon^{-1/2}-1+\theta\,(i-m)}\prod_{n=i}^{k-1}\frac{(\hat{y}_i^k-\hat{y}_{n+1}^k)\,\varepsilon^{-1/2}+\theta\,(n-i)+1}{(\hat{y}_i^k-\hat{y}_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-i)+1}\\
&\quad\times\prod_{m=1}^{i-1}\frac{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-m)-1}{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-1-m)}\prod_{n=i}^{k-1}\frac{(\hat{y}_i^k-\hat{y}_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-i+1)}{(\hat{y}_i^k-\hat{y}_{n+1}^k)\,\varepsilon^{-1/2}+\theta\,(n-i+1)}\\
&=\varepsilon^{-1}\prod_{m=1}^{i-1}\frac{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}-1+\theta\,(i-m+1)}{(\hat{y}_m^k-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-m)}\\
&\quad\times\frac{(\hat{y}_m^{k-1}-\hat{y}_i^k)\,\varepsilon^{-1/2}+\theta\,(i-1-m)}{(\hat{y}_m^{k-1}-\hat{y}_i^k)\,\varepsilon^{-1/2}-1+\theta\,(i-m)}\\
&\quad\times\prod_{n=i}^{k-1}\frac{(\hat{y}_i^k-\hat{y}_{n+1}^k)\,\varepsilon^{-1/2}+\theta\,(n-i)+1}{(\hat{y}_i^k-\hat{y}_{n+1}^k)\,\varepsilon^{-1/2}+\theta\,(n-i+1)}\,\frac{(\hat{y}_i^k-\hat{y}_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-i+1)}{(\hat{y}_i^k-\hat{y}_n^{k-1})\,\varepsilon^{-1/2}+\theta\,(n-i)+1}.
\end{aligned}
$$

Expanding the last expression into a Taylor series in terms of $\varepsilon^{1/2}$ we get

$$\varepsilon^{-1} + \varepsilon^{-1/2}\left(\sum_{m=1}^{i-1}\left(\frac{\theta-1}{\hat{y}_m^k - \hat{y}_i^k} + \frac{1-\theta}{\hat{y}_m^{k-1} - \hat{y}_i^k}\right)\right) + \left(\sum_{n=i}^{k-1}\left(\frac{1-\theta}{\hat{y}_i^k - \hat{y}_{n+1}^k} + \frac{\theta-1}{\hat{y}_i^k - \hat{y}_n^{k-1}}\right)\right) + O(1).$$

The lemma now readily follows. □

5.2 Step 2: Tightness

We show next that the family Y_ε^{mu}, $\varepsilon \in (0,1)$ is tight on $D^{N(N+1)/2}$. To this end, we aim to apply the necessary and sufficient conditions for tightness of [19, Corollary 3.7.4] which amount to showing that, for any fixed $t \geq 0$, the random variables $Y_\varepsilon^{mu}(t)$, $\varepsilon \in (0,1)$ are tight on $\mathbb{R}^{N(N+1)/2}$ and that, for every $\Delta > 0$ and $T > 0$, there exists a $\delta > 0$ such that

$$\limsup_{\varepsilon \downarrow 0} \mathbb{P}\left(\sup_{0\leq s<t\leq T, t-s<\delta} \left|(Y_\varepsilon^{mu})_i^k(t) - (Y_\varepsilon^{mu})_i^k(s)\right| > \Delta\right) < \Delta, \quad 1 \leq i \leq k \leq N.$$

We start by explaining how to deal with $(Y_\varepsilon^{mu}(t))_+$ (the vector of positive parts of components of $Y_\varepsilon^{mu}(t)$) and

$$\sup_{0\leq s<t\leq T, t-s<\delta} \left((Y_\varepsilon^{mu})_i^k(t) - (Y_\varepsilon^{mu})_i^k(s)\right), \quad 1 \leq i \leq k \leq N. \tag{5.2}$$

We argue by induction over k. For $k = 1$, it is sufficient to observe that $(Y_\varepsilon^{mu})_1^1$ is a Poisson process with jump size $\varepsilon^{1/2}$, jump rate ε^{-1} and drift $-\varepsilon^{-1/2}$ and, hence, converges to a standard Brownian motion in the limit $\varepsilon \downarrow 0$. Therefore, the necessary and sufficient conditions of [19, Corollary 3.7.4] hold for $(Y_\varepsilon^{mu})_1^1$.

We now fix some $k \geq 2$ and distinguish the cases $0 < \theta \leq 1$ and $\theta > 1$. In the first case, we consider first $i = k$. It is easy to see from the formulas for the jump rates in the proof of Lemma 5.6 that for $0 < \theta < 1$, whenever the spacing $(Y_\varepsilon^{mu})_k^k - (Y_\varepsilon^{mu})_{k-1}^{k-1}$ exceeds $\Delta/3$, the jump rate to the right of $(Y_\varepsilon^{mu})_k^k$ is bounded above by ε^{-1}. For $\theta = 1$ the jump rate is *equal* to ε^{-1}. Hence, arguing as in Sect. 3.2, we conclude that for $0 < \theta \leq 1$, $(Y_\varepsilon^{mu})_k^k$ can be coupled with a Poisson process with jump size $\varepsilon^{1/2}$, jump rate ε^{-1} and drift $-\varepsilon^{-1/2}$ in such a way that, whenever $(Y_\varepsilon^{mu})_k^k - (Y_\varepsilon^{mu})_{k-1}^{k-1}$ exceeds $\Delta/3$ and $(Y_\varepsilon^{mu})_k^k$ has a jump to the right, the Poisson process jumps to the right as well. Therefore, the convergence of such Poisson processes to a standard Brownian motion and the necessary and sufficient conditions of [19, Corollary 3.7.4] for them imply the corresponding conditions for $((Y_\varepsilon^{mu})_k^k(t))_+$ and the quantity in (5.2) with $i = k$.

For $i \in \{1, 2, \ldots, k-1\}$, the quantity $((Y_\varepsilon^{mu})_i^k(t))_+$ can be bounded above by the quantity $((Y_\varepsilon^{mu})_i^{k-1}(t))_+$ and the latter satisfies the required condition by the induction hypothesis. Moreover, the formula for the jump rates (5.1) reveals that, whenever $(Y_\varepsilon^{mu})_i^k - (Y_\varepsilon^{mu})_{i-1}^{k-1}$ and $(Y_\varepsilon^{mu})_i^{k-1} - (Y_\varepsilon^{mu})_i^k$ both exceed $\Delta/4$, the jump rate to the right of $(Y_\varepsilon^{mu})_i^k$ is bounded above by

$$\varepsilon^{-1} + \frac{\varepsilon^{-1/2}(1-\theta)}{(Y_\varepsilon^{mu})_i^{k-1}(t) - (Y_\varepsilon^{mu})_i^k(t)} + O(1) \le \varepsilon^{-1} + 4\,\varepsilon^{-1/2}(1-\theta)/\Delta + O(1). \tag{5.3}$$

Hence, $(Y_\varepsilon^{mu})_i^k$ can be coupled with a Poisson process with jump size $\varepsilon^{1/2}$, jump rate given by the right-hand side of the latter inequality and drift $-\varepsilon^{-1/2}$ in such a way that, whenever $(Y_\varepsilon^{mu})_i^k - (Y_\varepsilon^{mu})_{i-1}^{k-1}$ and $(Y_\varepsilon^{mu})_i^{k-1} - (Y_\varepsilon^{mu})_i^k$ both exceed $\Delta/4$ and $(Y_\varepsilon^{mu})_i^k$ has a jump to the right, the Poisson process jumps to the right as well. Thus, the convergence of such Poisson processes to Brownian motion with drift $4\,(1-\theta)/\Delta$ and the necessary and sufficient conditions of [19, Corollary 3.7.4] for them imply the corresponding control on the quantities in (5.2).

In the case $\theta > 1$, we first consider $i = 1$. From the formulas for the jump rates in the proof of Lemma 5.6 it is not hard to see that the jump rate to the right of the process $(Y_\varepsilon^{mu})_1^k$ is bounded above by ε^{-1}. Therefore, it can be coupled with a Poisson process with jump size $\varepsilon^{1/2}$, jump rate ε^{-1} and drift $-\varepsilon^{-1/2}$ in such a way that, whenever $(Y_\varepsilon^{mu})_1^k$ has a jump to the right, the Poisson process jumps to the right as well. Thus, as in Sect. 3.2, the convergence of such Poisson processes to a Brownian motion and the necessary and sufficient conditions of [19, Corollary 3.7.4] for them give the desired control on $((Y_\varepsilon^{mu})_1^k(t))_+$ and the quantity in (5.2) with $i = 1$.

For $i \in \{2, 3, \ldots, k\}$, the formulas for the jump rates in the proof of Lemma 5.6 reveal that, whenever $(Y_\varepsilon^{mu})_i^k - (Y_\varepsilon^{mu})_{i-1}^{k-1}$ exceeds $\Delta/3$, the jump rate to the right of $(Y_\varepsilon^{mu})_i^k$ is bounded above by

$$\varepsilon^{-1} - \frac{\varepsilon^{-1/2}(1-\theta)}{(Y_\varepsilon^{mu})_i^k(t) - (Y_\varepsilon^{mu})_{i-1}^{k-1}(t)} + O(1) \le \varepsilon^{-1} - 3\,\varepsilon^{-1/2}(1-\theta)/\Delta + O(1). \tag{5.4}$$

Hence, $(Y_\varepsilon^{mu})_i^k$ can be coupled with a Poisson process with jump size $\varepsilon^{1/2}$, jump rate given by the right-hand side of the last inequality and drift $-\varepsilon^{-1/2}$, so that, whenever $(Y_\varepsilon^{mu})_i^k - (Y_\varepsilon^{mu})_{i-1}^{k-1}$ exceeds $\Delta/3$ and $(Y_\varepsilon^{mu})_i^k$ has a jump to the right, the Poisson process jumps to the right as well. Therefore, as in Sect. 3.2, the convergence of such Poisson processes to Brownian motion with drift $-3\,(1-\theta)/\Delta$ and the necessary and sufficient conditions of [19, Corollary 3.7.4] for them yield the corresponding conditions for $((Y_\varepsilon^{mu})_i^k(t))_+$ and the quantities in (5.2).

Finally, we note that the quantities $(Y_\varepsilon^{mu}(t))_-$ (the vector of negative parts of components of $Y_\varepsilon^{mu}(t)$) and

$$\sup_{0 \le s < t \le T,\, t-s < \delta} -\big((Y_\varepsilon^{mu})_i^k(t) - (Y_\varepsilon^{mu})_i^k(s)\big), \quad 1 \le i \le k \le N$$

can be analyzed in a similar manner (however, now by moving from the leftmost to the rightmost particle on every level for $0 < \theta \leq 1$ and vice versa for $\theta > 1$). By combining everything together and using [19, Corollary 3.7.4] we conclude that the family Y_ε^{mu}, $\varepsilon \in (0, 1)$ is tight.

We also note that, since the maximal size of the jumps tends to zero as $\varepsilon \downarrow 0$, any limit point of the family Y_ε^{mu}, $\varepsilon \in (0, 1)$ as $\varepsilon \downarrow 0$ must have continuous paths (see e.g. [19, Theorem 3.10.2]).

5.3 Step 3: SDE for subsequential limits

Writing Y^{mu} for an arbitrary limit point as $\varepsilon \downarrow 0$ of the tight family Y_ε^{mu}, $\varepsilon \in (0, 1)$ as before, our goal now is to prove that Y^{mu} solves the SDE (4.1). We pick a sequence of Y_ε^{mu} which converges to Y^{mu} in law, and by virtue of the Skorokhod Embedding Theorem (see e.g. [18, Theorem 3.5.1]) may assume that all processes involved are defined on the same probability space and that the convergence holds in the almost sure sense. In the rest of this section all the limits $\varepsilon \downarrow 0$ are taken along such a sequence.

Define $\mathcal{F}^{mu}$ as the set of all infinitely differentiable functions on $\overline{\mathcal{G}^N}$ with support in a compact subset of $\mathcal{G}^N$. Further, set

$$\mathcal{F}_\delta^{mu} := \left\{ f \in \mathcal{F}^{mu} \,|\, f(x) = 0 \text{ whenever } \operatorname{dist}(x, \partial\overline{\mathcal{G}^N}) \leq \delta \right\},$$

where $\partial\overline{\mathcal{G}^N}$ stands for the boundary of $\overline{\mathcal{G}^N}$ and $\operatorname{dist}(x, \partial\overline{\mathcal{G}^N})$ is the L^∞ distance to the boundary, i.e.

$$\operatorname{dist}(x, \partial\overline{\mathcal{G}^N}) = \inf_{\substack{1 \leq i \leq j \leq N, \\ 1 \leq i' \leq j' \leq N, \\ (i,j) \neq (i',j')}} |x_i^j - x_{i'}^{j'}|.$$

Clearly $\mathcal{F}^{mu} = \bigcup_{\delta > 0} \mathcal{F}_\delta^{mu}$.

In addition, for every test function $f \in \mathcal{F}^{mu}$, define the process

$$\begin{aligned} M^f(t) := f(Y^{mu}(t)) - f(Y^{mu}(0)) - \sum_{1 \leq i \leq k \leq N} \int_0^t \frac{1}{2} f_{y_i^k y_i^k}(Y^{mu}(s))\, \mathrm{d}s \\ - \sum_{1 \leq i \leq k \leq N} \int_0^t \left(\sum_{m \neq i} \frac{1-\theta}{(Y^{mu})_i^k(s) - (Y^{mu})_m^k(s)} \right. \\ \left. - \sum_{m=1}^{k-1} \frac{1-\theta}{(Y^{mu})_i^k(s) - (Y^{mu})_m^{k-1}(s)} \right) f_{y_i^k}(Y^{mu}(s))\, \mathrm{d}s. \end{aligned}$$

Our first aim is to show that each M^f is a martingale. To this end, we fix an $f \in \mathcal{F}^{mu}$ and consider the family of martingales

$$M_\varepsilon^f(t) := f(Y_\varepsilon^{mu}(t)) - f(Y_\varepsilon^{mu}(0)) + \int_0^t \sum_{1\le i\le k\le N} \varepsilon^{-1/2} f_{y_i^k}(Y_\varepsilon^{mu}(s))\,\mathrm{d}s$$
$$- \int_0^t \sum_{y'\approx_\varepsilon Y_\varepsilon^{mu}(s)} q_\varepsilon^{mu}(Y_\varepsilon^{mu}(s), \hat{y}', s)\big(f(y') - f(Y_\varepsilon^{mu}(s))\big)\,\mathrm{d}s, \quad \varepsilon > 0$$

where the notations are the same as in Sect. 5.1. Now, one can argue as in Sect. 3.3, Step 3a to conclude that M^f is a martingale, as the $\varepsilon \downarrow 0$ limit of the martingales M_ε^f which can be bounded uniformly on every compact time interval.

Next, for each $\delta > 0$ and $K > 0$, we define the stopping time

$$\widehat{\tau}_{\delta,K} := \inf\big\{t \ge 0 : \mathrm{dist}(Y^{mu}(t), \partial\overline{\mathcal{G}^N}) \le \delta \quad \text{or} \quad |(Y^{mu})_i^k(t)| \ge K \text{ for some } (k,i)\big\}.$$

Now, note that for every function $\overline{\mathcal{G}^N} \to \mathbb{R}$ of the form $x \mapsto x_i^k$ or $x \mapsto x_i^k\, x_{i'}^{k'}$, there is a function $f \in \mathcal{F}^{mu}$ which coincides with that function on

$$\{x \in \overline{\mathcal{G}^N} : \mathrm{dist}(x, \partial\overline{\mathcal{G}^N}) \ge \delta,\ |x_i^k| \le K \text{ for all } (k,i)\}.$$

Combining this observation, the Optional Sampling Theorem and the conclusion of the preceeding paragraph we deduce that all processes of the two forms

$$M^{(k,i)}(t) := (Y^{mu})_i^k(t \wedge \widehat{\tau}_{\delta,K}) - (Y^{mu})_i^k(0)$$
$$- \int_0^{t\wedge\widehat{\tau}_{\delta,K}} \left(\sum_{m\ne i} \frac{1-\theta}{(Y^{mu})_i^k(s) - (Y^{mu})_m^k(s)} - \sum_{m=1}^{k-1} \frac{1-\theta}{(Y^{mu})_i^k(s) - (Y^{mu})_m^{k-1}(s)}\right)\mathrm{d}s,$$
$$M^{(k,i),(k',i')}(t) := (Y^{mu})_i^k(t \wedge \widehat{\tau}_{\delta,K})(Y^{mu})_{i'}^{k'}(t \wedge \widehat{\tau}_{\delta,K})$$
$$-(Y^{mu})_i^k(0)(Y^{mu})_{i'}^{k'}(0) - \mathbf{1}_{\{k=k',i=i'\}} \cdot t \wedge \widehat{\tau}_{\delta,K}$$
$$- \int_0^{t\wedge\widehat{\tau}_{\delta,K}} \left(\sum_{m\ne i} \frac{(1-\theta)(Y^{mu})_{i'}^{k'}(s)}{(Y^{mu})_i^k(s) - (Y^{mu})_m^k(s)} - \sum_{m=1}^{k-1} \frac{(1-\theta)(Y^{mu})_{i'}^{k'}(s)}{(Y^{mu})_i^k(s) - (Y^{mu})_m^{k-1}(s)}\right)\mathrm{d}s$$
$$- \int_0^{t\wedge\widehat{\tau}_{\delta,K}} \left(\sum_{m\ne i'} \frac{(1-\theta)(Y^{mu})_i^k(s)}{(Y^{mu})_{i'}^{k'}(s) - (Y^{mu})_m^{k'}(s)} - \sum_{m=1}^{k'-1} \frac{(1-\theta)(Y^{mu})_i^k(s)}{(Y^{mu})_{i'}^{k'}(s) - (Y^{mu})_m^{k'-1}(s)}\right)\mathrm{d}s$$

are martingales. Now, the argument of [29, Chapter 5, Proposition 4.6] (only replacing every occurrence of t by $t \wedge \widehat{\tau}_{\delta,K}$ there) yields that Y^{mu} satisfies the system of stochastic integral equations

$$(Y^{mu})_i^k(t \wedge \widehat{\tau}_{\delta,K}) - (Y^{mu})_i^k(0)$$
$$= \int_0^{t\wedge\widehat{\tau}_{\delta,K}} \left(\sum_{m\ne i} \frac{1-\theta}{(Y^{mu})_i^k(s) - (Y^{mu})_m^k(s)} - \sum_{m=1}^{k-1} \frac{1-\theta}{(Y^{mu})_i^k(s) - (Y^{mu})_m^{k-1}(s)}\right)\mathrm{d}s$$
$$+ \mathrm{W}_i^k(t \wedge \widehat{\tau}_{\delta,K}),\ 1 \le i \le k \le N$$

where W_i^k, $1 \le i \le k \le N$ are independent standard Brownian motions, possibly defined on an extension of the underlying probability space. At this stage, one can repeat the argument at the end of Sect. 4.2 to show that $\lim_{K\uparrow\infty} \widehat{\tau}_{\delta,K} \ge \tau_\delta[Y^{mu}]$ and then combine Propositions 4.2 and 4.3 to end up with the SDE of Theorem 5.2 as desired.

Remark 5.7 In Sect. 3.3 we have used Lemma 3.6 instead of [29, Chapter 5, Proposition 4.6], but we could have used the latter as well. On the other hand, it is not straightforward to generalize Lemma 3.6 to the setting of the current section, because there is no obvious multilevel analogue for Proposition 2.25.

Acknowledgments We would like to thank Amir Dembo for many fruitful discussions and, in particular, for suggesting the use of a suitable Girsanov change of measure in the proof of Theorem 4.1. We also want to thank Alexei Borodin, Grigori Olshanski and Mark Adler for helpful comments. V. G. was partially supported by RFBR-CNRS Grant 11-01-93105 and by NSF Grant DMS-1407562.

References

1. Adler, M., Nordenstam, E., Van Moerbeke, P.: Consecutive minors for Dyson Brownian motions. Stoch. Process. Applic. **124**(6), 2023–2051 (2014). arXiv:1007.0220
2. Anderson, G.W.: A short proof of Selberg's generalized beta formula. Forum Math. **3**, 415–417 (1991)
3. Anderson, G.W., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics, vol. 118. Cambridge University Press, Cambridge (2010)
4. Baryshnikov, Y.: GUEs and queues. Probab. Theory Relat. Fields **119**(2), 256–274 (2001)
5. Borodin, A.: CLT for spectra of submatrices of Wigner random matrices II. Stochastic evolution. arXiv:1011.3544
6. Borodin, A.: Schur dynamics of the Schur processes. Adv. Math. **228**, 2268–2291 (2011). arXiv:1001.3442
7. Borodin, A., Corwin, I.: Macdonald processes. Probab. Theory Relat. Fields **158**(1–2), 225–400 (2014). arXiv:1111.4408
8. Borodin, A., Corwin, I., Gorin, V., Shakirov, S. : Observables of Macdonald processes. Trans. AMS (to appear). arXiv:1306.0659
9. Borodin, A., Ferrari, P.: Anisotropic growth of random surfaces in 2 + 1 dimensions. Commun. Math. Phys. **325**(2), 603–684 (2014). arXiv:0804.3035
10. Borodin, A., Gorin, V.: Shuffling algorithm for boxed plane partitions. Adv. Math. **220**(6), 1739–1770 (2009). arXiv:0804.3071
11. Borodin, A., Gorin, V.: Lectures on Integrable Probability. arXiv:1212.3351
12. Borodin, A., Gorin, V.: General β Jacobi corners process and the Gaussian free field. arXiv:1305.3627
13. Borodin, A., Petrov, L.: Integrable probability: from representation theory to Macdonald processes. Probab. Surv. **11**, 1–58 (2014). arXiv:1310.8007
14. Cépa, E., Lépingle, D.: Diffusing particles with electrostatic repulsion. Probab. Theory Relat. Fields **107**, 429–449 (2012)
15. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Financial Mathematics Series. Chapman & Hall/CRC, London (2004)
16. Diaconis, P., Fill, J.A.: Strong stationary times via a new form of duality. Ann. Probab. **18**, 1483–1522 (1990)
17. Dixon, A.L.: Generalizations of Legendre formula $KE' - (K - E)K' = \frac{1}{2}\pi$. Proc. Lond. Math. Soc. **3**, 206–224 (1905)
18. Dudley, R.M.: Uniform Central Limit Theorems. Cambridge University Press, Cambridge (1999)
19. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New York (1986)
20. Forrester, P.J.: Log-gases and Random Matrices. Princeton University Press, Princeton (2010)
21. Gel'fand, I.M., Naimark, M.A.: Unitary Representations of the Classical Groups. Trudy Mat. Inst. Steklov, Leningrad, Moscow (1950) (in Russian). (German transl.: Academie-Verlag, Berlin, 1957)

22. Gorin, V.: From alternating sign matrices to the gaussian unitary ensemble. Commun. Math. Phys. **332**, 437–447 (2014). arXiv:1306.6347
23. Gorin, V., Panova, G.: Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. arXiv:1301.0634
24. Gorin, V., Shkolnikov, M.: Limits of multilevel TASEP and similar processes. Annales de l'Institut Henri Poincare (to appear). arXiv:1206.3817
25. Han, W.: Simulations of continuous time dynamics for $\theta = 0.5$ and $\theta = 2$. (2013) Available at http://stat.berkeley.edu/mshkolni/X.jar and http://mccme.ru/vadicgor/Jack_dynamics.jar
26. Johansson, K., Nordenstam, E.: Eigenvalues of GUE minors. Electron. J. Probab. **11**(50), 1342–1371 (2006). (electronic)
27. Kerov, S., Okounkov, A., Olshanski, G.: The boundary of Young graph with Jack edge multiplicities. Intern. Math. Res. Notices (4), 173–199 (1998). arXiv:q-alg/9703037
28. Kalenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, Berlin (2002)
29. Karatzas, I., Shreve, S.: Brownian Motion and Stochastic Calculus, 2nd edn. Springer, New York (1991)
30. Liggett, T.M., Schonmann, R.H., Stacey, A.M.: Domination by product measures. Ann. Probab. **25**(1), 71–95 (1997)
31. Macdonald, I.G.: Symmetric Functions and Hall Polynomials, 2nd edn. Oxford University Press, Oxford (1999)
32. Mehta, M.L.: Random Matrices. Elsevier, Amsterdam (2004)
33. Neretin, YuA: Rayleigh triangles and non-matrix interpolation of matrix beta integrals. Sbornik: Math. **194**(4), 515–540 (2003)
34. Nordenstam, E.: On the shuffling algorithm for domino tilings. Electron. J. Probab., **15**, 75–95 (2010), Paper no. 3. arXiv:0802.2592
35. Okounkov, A., Olshanski, G.: Shifted Jack polynomials, binomial formula, and applications. Math. Res. Lett. **4**, 69–78 (1997). arXiv:q-alg/9608020
36. Okounkov, A., Olshanski, G.: Asymptotics of Jack polynomials as the number of variables tends to infinity. Intern. Math. Research Notices 1998, no. 13, 641–682. arXiv:q-alg/9709011
37. Okounkov, A.: Infinite wedge and random partitions. Selecta Math. **7**, 57–81 (2001). arXiv:math/9907127
38. Okounkov, A.Yu., Reshetikhin, N.Yu.: The birth of a random matrix. Mosc. Math. J. **6**(3), 553–566 (2006)
39. Pal, S., Shkolnikov, M.: Intertwining diffusions and wave equations. arXiv:1306.0857
40. Pardoux, E., Williams, R.J.: Symmetric reflected diffusions. Annales de l'Insitut Henri Poincare, section B **30**, 13–62 (1994)
41. Ramirez, J., Rider, B., Virag, B.: Beta ensembles, stochastic Airy spectrum, and a diffusion. J. Am. Math. Soc. **24**(4), 919–944 (2011). arXiv:math/0607331
42. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 3rd edn. Springer, New York (1999)
43. Sekiguchi, Jiro: Zonal spherical functions on some symmetric spaces. Publ. RIMS. Kyoto Univ. **12**(Suppl), 455–464 (1977)
44. Selberg, A.: Bemerkninger om et multipelt integral. Norsk. Mat. Tidsskr. **24**, 71–78 (1944)
45. Soucaliuc, F., Toth, B., Werner, W.: Reflection and coalescence between independent one-dimensional Brownian paths. Annales de l'Institut Henri Poincare (B) Probability and Statistics **36**, 509–545 (2000)
46. Warren, J.: Dyson's Brownian motions, intertwining and interlacing, Electron. J. Probab., **12**, 573–590 (2007), Paper no. 19. arXiv:math/0509720