

ONE-DIMENSIONAL BROWNIAN PARTICLE SYSTEMS WITH RANK DEPENDENT DRIFTS

BY SOUMIK PAL*, JIM PITMAN†

Cornell University and University of California, Berkeley

We study interacting systems of linear Brownian motions whose drift vector at every time point is determined by the relative ranks of the coordinate processes at that time. Our main objective has been to study the long range behavior of the spacings between the Brownian motions arranged in increasing order. For finitely many Brownian motions interacting in this manner, we characterize drifts for which the family of laws of the vector of spacings is tight, and show its convergence to a unique stationary joint distribution given by independent exponential distributions with varying means. We also study one particular countably infinite system, where only the minimum Brownian particle gets a constant upward drift, and prove that independent and identically distributed exponential spacings remain stationary under the dynamics of such a process. Some related conjectures in this direction have also been discussed.

1. Introduction. In this paper we consider systems of interacting one-dimensional Brownian motions $X = (X_i(t), i \in I, t \geq 0)$, where i ranges over an index I , which is either the finite set $\{1, \dots, N\}$, or the countable set of positive integers \mathbb{N} .

For $I = \{1, 2, \dots, N\}$, if we define ordered coordinates of any vector $x \in \mathbb{R}^N$ by

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)},$$

the locations $X_i(t)$ of the Brownian particles evolve according to the system of stochastic differential equations

$$(1) \quad dX_i(t) = \sum_{j \in I} \delta_j 1 \left\{ X_i(t) = X_{(j)}(t) \right\} dt + dB_i(t) \quad (i \in I)$$

for some sequence of *drifts* $\delta_1, \delta_2, \dots \in \mathbb{R}$. Here the B_i 's are assumed to be independent (\mathcal{F}_t) -Brownian motions for some suitable underlying filtration

*Research is partially supported by N.S.F. grant DMS-0306194 to the probability group at Cornell.

†Research is supported in part by N.S.F. Grants DMS-9970901, DMS-0405779, and DMS-0071448.

AMS 2000 subject classifications: Primary 60G07, 60G55

(\mathcal{F}_t) . Less formally, the Brownian particles evolve independently except that the i th ranked particle is given drift δ_i .

For these finite systems, with arbitrary initial distribution of $(X_i(0), i \in I)$, and arbitrary drifts δ_i , the existence and uniqueness in law of such an N particle system is guaranteed by the standard theory of SDE's. See Lemma 6 for more details. We would also like to define SDE (1) when I is the entire countably infinite collection \mathbb{N} . This is trickier since the ordered process might not remain well-defined anymore, and the existence of the solution of SDE (1) depends both on the initial distribution and the sequence of drifts. We consider one such system in Section 3 with a drift sequence $(\delta, 0, 0, \dots)$ ($\delta > 0$) for which the finite-dimensional arguments can be suitably extended. For this system to exist in the weak sense, starting from an initial vector $X_1(0) < X_2(0) < \dots$, we show that it suffices to assume

$$\liminf_{n \rightarrow \infty} \frac{(X_n(0) - X_1(0))^2}{n} > 0.$$

See Lemma 11 for the details of the proof.

Consider now the *Brownian spacings system* derived from this ordered Brownian particle system with rank dependent drifts, i.e.,

$$\Delta_k(t) := X_{(k+1)}(t) - X_{(k)}(t) \text{ for } k, k+1 \in I.$$

The ordered particle system derived from independent Brownian motions with no drift (meaning $\delta_i \equiv 0$) was studied by Harris [14], Arratia [1] and Sznitman [27],[28, p. 187]. By Donsker's theorem, this system can be interpreted as a scaling limit of ordered particle systems derived from independent symmetric nearest neighbour random walks on \mathbb{Z} . Harris [14] considers the spacings of an infinite ordered Brownian particle system defined by :

$$\Delta_i^*(t) := B_{(i+1)}(t) - B_{(i)}(t) \quad (i \in \mathbb{Z})$$

where $\{B_i\}$ is a family of independent Brownian motions with no drifts and initial states $B_i(0) = B_{(i)}(0)$ such that $B_0(0) = 0$ and the $B_i(0)$ for $i \in \mathbb{Z} \setminus \{0\}$ are points of a Poisson process of rate λ on \mathbb{R} . That is, $(B_{(i)}(t), t \geq 0)_{i \in \mathbb{Z}}$ is the almost surely unique collection of processes with continuous paths such that $B_{(i)}(t) \leq B_{(i+1)}(t)$ for all $i \in \mathbb{Z}, t \geq 0$ and the union of the graphs of these processes is identical to the union of the graphs of the Brownian paths $(B_i(t), t \geq 0)_{i \in \mathbb{Z}}$. We call $(B_{(i)}(t), i \in \mathbb{Z}, t \geq 0)$ the *Harris system of ordered Brownian motions with rate λ* and their differences $(\Delta_i^*(t), i \in \mathbb{Z}, t \geq 0)$ the *Harris system of Brownian spacings with rate λ* .

As observed by Arratia [2, §4], the corresponding stationary system of spacings between particles of the exclusion process associated with a nearest-neighbor random walk on \mathbb{Z} can be interpreted as a finite or infinite series of queues, also known as the *zero range process with constant rate*. See [15, 16, 17, 19, 18, 22] for background on systems of Brownian queues. Such connections between systems of queues and one-dimensional interacting particle systems have been exploited by a number of authors, in particular Kipnis [20], Srinivasan [26], Ferrari-Fontes [12], [13], and Seppäläinen [25]. Ferrari [11] surveys old and new results on the limiting behavior of a tagged particle in various interacting particle systems. Also see articles by Baryshnikov [4] and O’Connell-Yor [22] for some recent studies of Brownian queues in tandem, connected to the directed percolation and the directed polymer models, and the GUE random matrix ensemble.

More recently rank dependent SDEs have been considered by several authors as possible models for financial or economic data. Fernholz in [9] introduces the so-called *Atlas model* which we study in this paper. It is a model of finitely many Brownian particles, where at every time point the minimum Brownian motion gets a constant positive drift, while the rest of them get no drift. The general rank dependent interacting Brownian models, whose drifts and volatilities depend on time-varying ranks have been considered by Banner, Fernholz, and Karatzas in [3], with whom our work in this paper bears close resemblance. For SDE (1), they work under a specific condition on the drift sequence required for stability of the solution process. We prove in Theorem 8 (condition (13)) that this condition is indeed necessary and sufficient. Although their method is mostly based on a beautiful analysis of the local times of intersections of different Brownian motions, they also note the connection with the Harrison-Williams theory of reflected Brownian motions which we use extensively in this paper. See Lemma 4 for a complete statement.

In that same article [3] the authors establish marginal convergence of spacings Δ_k to exponential distributions. They leave the question of joint convergence open, which we settle in this paper in Theorem 8 by proving that the vector of spacings converge jointly to independent exponentials with different means. They also study ergodic properties of such processes including a demonstration of the exchangeability of the indices of the Brownian particles under rank-dependent drifts and volatilities. In their later papers, Fernholz and Karatzas also consider generalized versions of (1) where the drifts and the volatilities depend on both the index and the rank of a Brownian particle. As expected, in most such cases, explicit descriptions of their properties become very difficult to obtain. However, many interesting results

can still be recovered. A good source for what has been done so far can be found in the recent survey article by Fernholz and Karatzas [10].

McKean and Shepp, in [21], also consider Brownian motions interacting via their ranks. They start with two Brownian motions, and their objective is to find the optimal drift under constraints (as a control) such that the probability that both Brownian motions never hit zero is maximized. The solutions is, as they establish, the Atlas model for the two particle system.

An interesting related model, studied by Rost and Vares in [24], replaces the ordered particles in the Harris model by linear Brownian motions repelled by their nearest neighbors through a potential. The authors study stationary measures for the spacings of such processes, and show that rescaled combinations of spacings converge to an Ornstein-Uhlenbeck process.

Our purpose here is to draw attention to the general class of Brownian particle systems with rank dependent drifts, as considered in [3]. Many natural questions about these systems remain open. We are particularly interested in an infinite version of the Atlas model, with a drift sequence $(\delta, 0, 0, \dots)$ with *Atlas drift* $\delta > 0$. One result we obtain for this system is the following:

THEOREM 1. *For each $\delta > 0$, a sequence of independent Exponential(2δ) variables provides an equilibrium distribution of spacings for the infinite Atlas model with the Atlas drift δ .*

Theorem 1, which essentially follows from Theorem 14, suggests a number of interesting conjectures and open problems. In particular, we can immediately formulate:

CONJECTURE 2. *For each $\delta > 0$, Theorem 1 describes the unique equilibrium distribution of spacings for the infinite Atlas particle system with Atlas drift δ .*

Let $(\Delta_1(t), \Delta_2(t), \dots)_{t \geq 0}$ denote the equilibrium state of spacings of the infinite Atlas Brownian particle system described by Theorem 1. This process has some subtle features. For each $k = 1, 2, \dots$ and each $t > 0$,

$$(2) \quad (\Delta_k(t), \Delta_{k+1}(t), \dots) \stackrel{d}{=} (\Delta_1(t), \Delta_2(t), \dots)$$

and the common distribution of these sequences is that of independent exponential (2δ) variables. But while both sides of (2) define stationary sequence-valued processes as t varies, these processes do not have the same law for all k . In particular, the finite-dimensional distributions of non-negative stationary process $(\Delta_k(t), t \geq 0)$ depend on k .

Harris [14, eqn. (7.1)] gave an explicit formula for the law of $B_{(0)}(t)$, the location at time t of particle initially at 0 in the Harris system of ordered Brownian motions, from which he deduced for $2\lambda = 1$ that

$$(3) \quad \frac{B_{(0)}(t)}{t^{1/4}} \xrightarrow{d} \left(\frac{2}{\pi}\right)^{1/4} \frac{1}{\sqrt{2\lambda}} B(1) \text{ as } t \rightarrow \infty$$

where $B(1)$ is standard Gaussian. As remarked by Arratia [1, p. 71], the conclusion for general $\lambda > 0$ follows from the case $2\lambda = 1$ by Brownian scaling. See also De Masi, Ferrari [5], Rost, Vares [24], and Arratia [1] where variants (or generalizations) of (3) is proved for a tagged particle in the exclusion process on \mathbb{Z} associated with a simple symmetric random walk. Harris conjectured that the process $B_{(0)}$ is not Markov, and left open the problem of describing the long-run behaviour of paths of $B_{(0)}$. These questions were answered by Dürr, Goldstein, and Lebowitz in [6, Thm 7.1] where they prove that the *tagged* process $B_{(0)}(t)$, suitably rescaled, converges to fractional Brownian motion with Hurst parameter $1/4$. In fact, they show a general theorem where convergence to fractional Brownian motion holds for systems with processes which have stationary increments and perform *elastic collisions* as in the Harris model. Also see [7] where the same authors generalize their results in the case where an external potential is present.

In this connection, we have the following conjecture:

CONJECTURE 3. *For each $\delta > 0$, in the infinite Atlas model with Atlas drift $\delta > 0$, and initially ordered values $X_{(k)}(0), k = 1, 2, \dots$ which are the points of a Poisson process with rate 2δ on $(0, \infty)$,*

$$\frac{X_{(k)}(t) - X_{(k)}(0)}{t^{1/4}} \xrightarrow{d} \left(\frac{2}{\pi}\right)^{1/4} \frac{c_k}{\sqrt{2\delta}} B(1) \text{ as } t \rightarrow \infty$$

for some sequence of constants $c_k > 0$ with $c_k \rightarrow 1$ as $k \rightarrow \infty$.

The paper is organized as follows. In the next section, we analyze the finite Brownian particle system with rank dependent drifts. The main result is Theorem 8, which describes the convergence in total variation of the laws of the spacings to that of independent exponential distributions. The precise condition needed on the drift sequence for such stability has also been proved. In Section 3 we look at countably infinite Brownian particles with the dynamics of the Atlas model. The main result, Theorem 1, follows readily from Theorem 14.

2. The finite Brownian particle system. We first present some results regarding the asymptotic behavior of spacings for the N particle system with arbitrary rank-dependent drifts $\delta_i, 1 \leq i \leq N$. We start by recording the following two lemmas, which clarify the issues of existence and uniqueness of the N particle system with arbitrary drifts δ_i , and characterize the associated ordered particle system. See the book by Revuz and Yor [23] for a background and the definitions of concepts from the calculus of continuous semimartingales.

LEMMA 4. *Let $X_i, 1 \leq i \leq N$ be a solution of the SDE (1) defined on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, for some arbitrary initial condition and arbitrary drifts $\{\delta_j\}$. Then for each $1 \leq j \leq N$ the j th ordered process $X_{(j)}$ is a continuous semimartingale with decomposition*

$$(4) \quad dX_{(j)}(t) = d\beta_j(t) + \frac{1}{\sqrt{2}}(dL_{(j-1,j)}(t) - dL_{(j,j+1)}(t))$$

where the β_j 's for $1 \leq j \leq N$ are independent (\mathcal{F}_t) -Brownian motions with unit variance coefficient and drift coefficients δ_j , and where $L_{(0,1)} = L_{(N,N+1)} = 0$, and for $1 \leq j \leq N - 1$

$$(5) \quad L_{(j,j+1)}(t) = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t 1((X_{(j+1)}(s) - X_{(j)}(s))/\sqrt{2} \leq \varepsilon) ds \quad (t \geq 0)$$

which is half the continuous increasing local time process at 0 of the semimartingale $(X_{(j+1)} - X_{(j)})/\sqrt{2}$. Moreover, the ordered system is a Brownian motion in the domain

$$(6) \quad \{(x_{(i)}, 1 \leq i \leq N) \in \mathbb{R}^N : x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)}\}$$

with constant drift vector $(\delta_j, 1 \leq j \leq N)$ and normal reflection at each of the $N - 1$ boundary hyperplanes $\{x_{(i)} = x_{(i+1)}\}$ for $1 \leq i \leq N - 1$.

PROOF. Sznitman [27, page 594], [28, Lemma 3.7] gave these results in the case of zero drifts. As he observed, they follow from Tanaka's formula [23, page 223] and the definition of Brownian motion in a polyhedron with normal reflection. See, for example, Varadhan and Williams [29]. Sznitman further proves that the Brownian motions $\{\beta_j\}$ are independent, which is essentially due to Knight's theorem [23, page 183]. The factors of $\sqrt{2}$ are most easily easily checked in the case $N = 2$. That they are the same for all N follows by a localization argument. The results of the lemma for general drifts δ_j are deduced from the results with no drift by application of the next two lemmas. \square

LEMMA 5. *Let X_1, X_2, \dots be a sequence of independent Brownian motions indexed by a finite set I . Consider the ordered process $X_{(1)}, X_{(2)}, \dots$. Then, the following holds almost surely:*

$$(7) \quad \sum_{i=1}^{|I|} \int_0^t 1 \left\{ X_i(s) = X_{(j)}(s) \right\} ds = t, \quad \forall j < |I| + 1, \forall t \in [0, \infty).$$

Moreover, the points of increments of the finite variation processes $L_{(j,j+1)}$ are almost surely disjoint.

When I is countable, equation (7) still holds true as long as the assumptions on the initial values $X_1(0) < X_2(0) < \dots$ are sufficient to guarantee the existence of the ordered processes for all finite times.

PROOF. For any two indices $k < l$, the lebesgue measure of the set $\{t \in [0, \infty) : X_k = X_l\}$ is zero. This follows because the zero set of Brownian motion is of lebesgue measure zero almost surely.

Now, the event $\sum_{i=1}^{|I|} \int_0^t 1 \left\{ X_i(s) = X_{(j)}(s) \right\} ds > t$ for some j and some t implies that there exists some pair (k, l) such that the lebesgue measure of the set $\{0 \leq s \leq t : X_k(s) = X_{(j)}(s) = X_l(s)\}$ is positive. This is of measure zero according to the previous paragraph and by countable additivity.

The other possibility of $\sum_{i=1}^{|I|} \int_0^t 1 \left\{ X_i(s) = X_{(j)}(s) \right\} ds < t$ is trivially ruled out by our assumption that the ordered processes are always achieved.

For the second assertion, note that, according to the general theory of semi-martingale local times [23], the process $L_{(j,j+1)}$ increases only on the random closed set of times t when $X_{(j)}(t) = X_{(j+1)}(t)$. These random sets are almost surely disjoint as j varies. This is because, with probability 1, there is *no triple collision*, meaning a time $t > 0$ at which $X_i(t) = X_j(t) = X_k(t)$. It follows from the fact that the bivariate process $(X_i - X_j, X_j - X_k)$ is a linear transformation of a standard planar Brownian motion which does not hit points. \square

Call $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ the *canonical setup* if $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ is the usual space of continuous paths in \mathbb{R}^N , with the usual right-continuous filtration, and X_i is the i th coordinate process.

LEMMA 6. *Let $\delta = (\delta_i, 1 \leq i \leq N) \in \mathbb{R}^N$ and let μ be a probability distribution on \mathbb{R}^N .*

(i) *In the canonical setup there is a unique probability measure $\mathbb{P}^{\delta, \mu}$ under which the coordinate processes $(X_i, 1 \leq i \leq N)$ solve the system of SDE's (1) with initial distribution μ . In particular, for $\delta = 0$, the law $\mathbb{P}^{0, \mu}$ is the*

Wiener measure governing standard Brownian motion in \mathbb{R}^N with initial distribution μ .

(ii) In the canonical setup, for each $t > 0$, the law $\mathbb{P}^{\delta, \mu}$ is absolutely continuous with respect to $\mathbb{P}^{0, \mu}$ on \mathcal{F}_t , with density

$$(8) \quad \exp \left(\sum_{j=1}^N \delta_j \beta_j(t) - \frac{1}{2} \sum_{j=1}^N \delta_j^2 t \right)$$

where β_j , which is the same as in (4), can be defined by the expression

$$(9) \quad \beta_j(t) = \sum_{i=1}^N \int_0^t 1 \{ X_i(s) = X_{(j)}(s) \} dX_i(s), \quad 1 \leq j \leq N.$$

Under $\mathbb{P}^{\delta, \mu}$ the β_j 's are independent Brownian motions on \mathbb{R} with drift coefficients $\{\delta_j\}$ and unit diffusion coefficients.

(iii) If $(X_i, 1 \leq i \leq N)$ is a realization of the N particle system with drifts $\{\delta_i\}$ and initial distribution μ on an arbitrary probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then the \mathbb{P} joint distribution of the processes X_i is identical to the $\mathbb{P}^{\delta, \mu}$ distribution of the coordinate processes on the canonical space, as specified in (ii).

PROOF. This is an instance of a well known general construction of the solution of an SDE with drift terms from one with no drift terms [23, Chapter IX, Theorem (1.10)].

Under $\mathbb{P}^{0, \mu}$, the fact that $\{\beta_j\}$ is a collection of independent Brownian motions also follows from expression (9) and equation (7). Under $\mathbb{P}^{\delta, \mu}$, the process β_j is a stochastic integral with respect to Brownian motions with drifts. By expanding the drift term, it is obvious that β_j is a Brownian motion itself with drift δ_j . \square

Note that the SDE defining the N particle system with drifts is a typical example of an SDE for which there is uniqueness in law, but not pathwise uniqueness.

For a solution X of SDE (1), let

$$\bar{X}(t) := \frac{1}{N} \sum_{i=1}^N X_i(t) = \frac{1}{N} \sum_{j=1}^N X_{(j)}(t)$$

which is the *center of mass* of the particle system, where we regard each particle as having mass $1/N$. We call the \mathbb{R}^N -valued process

$$(X_i - \bar{X}, 1 \leq i \leq N)$$

the *centered system*. Note that the $N - 1$ spacings defined by taking differences of order statistics of the original system are identical to the $N - 1$ spacings defined by differences of the order statistics of the centered system. So the N order statistics of the centered system, which are constrained to have average 0, are an invertible linear transformation of the $N - 1$ spacings of original system.

LEMMA 7. *For the N particle system, with arbitrary drifts and initial distribution,*

(i) *the center of mass process is a Brownian motion with drift*

$$(10) \quad \bar{\delta}_N := N^{-1} \sum_{j=1}^N \delta_j$$

and diffusion coefficient $1/N$; explicitly

$$(11) \quad \bar{X}(t) = \bar{X}(0) + \bar{\delta}_N t + \frac{B(t)}{\sqrt{N}}$$

where $B(t) := N^{-\frac{1}{2}} \sum_{j=1}^N B_j(t)$ is a standard BM on \mathbb{R} ; consequently

$$\frac{\bar{X}(t)}{t} \rightarrow \bar{\delta}_N \text{ almost surely as } t \rightarrow \infty.$$

(ii) *The shifted center of mass process $\bar{X}(t) - \bar{X}(0)$ is independent of the centered system $(X_i(\cdot) - \bar{X}(\cdot))$, $1 \leq i \leq n$.*

PROOF. Part (i) follows plainly from the SDE (1) by summing over i .

For part (ii), note that if σ_1 and σ_2 are two independent sub- σ -algebras of a probability space (A, \mathcal{A}, P) , and we change P to another probability Q via defining $dQ/dP = fg$, where $f \in \sigma_1$, $g \in \sigma_2$, then σ_1 and σ_2 remain independent under Q .

Now as in Lemma 6, when (X_i) denotes N independent Brownian motions with initial distribution μ , the shifted average $\bar{X}(t) - \bar{X}(0)$ and the centered process $Y = X - \bar{X}\mathbf{1}$ are independent. This follows from the fact that conditionally on X_0 , the processes are Gaussian with zero covariance, and that $\bar{X}(t) - \bar{X}(0)$ is independent of X_0 . Now to get to $\mathbb{P}^{\delta, \mu}$, the Radon-Nikodym derivative is given by (8). Note that, from expression (9) we get

$$\beta_j(t) = \sum_{i=1}^N \int_0^t 1 \{Y_i(s) = Y_{(j)}(s)\} dY_i(s) + \bar{X}(t) - \bar{X}(0).$$

Thus, from (8), it is clear that $d\mathbb{P}^{\delta, \mu}/d\mathbb{P}^{0, \mu}$ can be written as fg , where $f \in \sigma(Y)$ and $g \in \sigma(\bar{X}(\cdot) - \bar{X}(0))$. Now, by first localizing at finite time, and using the argument in the preceding paragraph, we establish the claim in part (ii). \square

THEOREM 8. *For $1 \leq k \leq N$ let*

$$(12) \quad \alpha_k := \sum_{i=1}^k (\delta_i - \bar{\delta}_N)$$

where $\bar{\delta}_N$ is the average drift as in (10). For each fixed initial distribution of the N particle system with drifts $\{\delta_i, i = 1, 2, \dots, N\}$, the collection of laws of $X_{(N)}(t) - X_{(1)}(t)$ for $t \geq 0$ is tight if and only if

$$(13) \quad \alpha_k > 0 \text{ for all } 1 \leq k \leq N - 1,$$

in which case the following results all hold:

- (i) The distribution of the spacings system $(X_{(j+1)} - X_{(j)}, 1 \leq j \leq N - 1)$ at time t converges in total variation norm as $t \rightarrow \infty$ to a unique stationary distribution for the spacings system, which is that of independent exponential variables Y_j with rates $2\alpha_j, 1 \leq j \leq N - 1$. Moreover, the spacings system is reversible at equilibrium.
- (ii) The distribution of the centered system at time t converges in total variation norm as $t \rightarrow \infty$ to a unique stationary distribution for the centered system, which is the distribution of

$$(S_{\pi(i)-1} - \bar{S}, 1 \leq i \leq N)$$

where $S_0 := 0$ and $S_i := Y_1 + \dots + Y_i$ for $1 \leq i \leq N - 1$, where π a uniform random permutation of $\{1, \dots, N\}$ which is independent of the Y_i , and

$$\bar{S} := \frac{1}{N} \sum_{i=1}^N S_{\pi(i)-1} = \frac{1}{N} \sum_{i=1}^{N-1} S_i = \frac{1}{N} \sum_{i=1}^{N-1} (N - i)Y_i.$$

Moreover, the centered system is reversible at equilibrium.

- (iii) As $t \rightarrow \infty$

$$X_i(t)/t \rightarrow \bar{\delta}_N \text{ almost surely for each } 1 \leq i \leq N,$$

and the same is true for $X_{(i)}(t)/t$ instead of $X_i(t)/t$.

Remark. Regard the system as split into *left-hand particles* of rank 1 to k and *right-hand particles* of rank $k + 1$ to N . If these two parts of the system are started at some strictly positive distance from each other, they evolve independently like copies of the k -particle system and the $(N - k)$ -particle system respectively, until the first time there is a collision between a left-hand particle and a right-hand particle. It follows by summing up over the corresponding drifts that the centers of mass of the two subsystems left to themselves would have almost sure asymptotic speeds $\bar{\delta}_k$ and $\hat{\delta}_k$ respectively, where

$$\bar{\delta}_k := \frac{1}{k} \sum_{i=1}^k \delta_i \text{ and } \hat{\delta}_k := (N - k)^{-1} \sum_{i=k+1}^N \delta_i.$$

Since

$$\alpha_k = k\bar{\delta}_k - k\bar{\delta}_N = \frac{k(N - k)}{N} (\bar{\delta}_k - \hat{\delta}_k)$$

we see that $\alpha_k > 0$ iff $\bar{\delta}_k > \hat{\delta}_k$, which ensures that the righthand system cannot avoid an eventual collision with the lefthand system. According to (12), for arbitrary prescribed $\bar{\delta}_N \in \mathbb{R}$, and $\alpha_k > 0$, the unique drift vector determining an ergodic N -particle system whose average drift is $\bar{\delta}_N$ and whose asymptotic spacings are independent exponential variables with rates $2\alpha_k$, is given by

$$(14) \quad \delta_i = \bar{\delta}_N + \alpha_i - \alpha_{i-1} \quad (1 \leq i \leq N)$$

where $\alpha_0 := \alpha_N := 0$. Given an arbitrary cumulative probability distribution function F on the line, and arbitrary $\bar{\delta} \in \mathbb{R}$ and $\varepsilon > 0$, it is clear that by taking N suitably large, we can choose $(\alpha_k, 1 \leq k \leq N - 1)$ and hence $(\delta_j, 1 \leq j \leq N)$ so that for all sufficiently large t

$$\mathbb{P} \left(\sup_x \left| \frac{1}{N} \sum_{i=1}^N 1(X_i(t) - \bar{X}(t) \leq x) - F(x) \right| > \varepsilon \right) < \varepsilon$$

and $\bar{X}(t)/t \rightarrow \bar{\delta}$ almost surely as $t \rightarrow \infty$. Thus no matter what its initial distribution, the N particle system looks asymptotically like a cloud of particles with mass distribution close to F , drifting along the line at speed $\bar{\delta}$.

PROOF OF THEOREM 8. According to Lemma 4, the ordered N particle system is a Brownian motion $(X_{(k)}(t), 1 \leq k \leq N)_{t \geq 0}$ in the domain (6) with identity covariance matrix, constant drift vector $(\delta_j, 1 \leq j \leq N)$ and normal reflections at each of the $N - 1$ boundary hyperplanes $\{x_{(i)} = x_{(i+1)}\}$

for $1 \leq i \leq N - 1$. Note that the vector $(1, 1, \dots, 1)$ of all ones is in the intersection of these boundary hyperplanes.

Let β be an independent one-dimensional Brownian motion ($\beta_0 = 1$) with a negative drift $-\theta$ ($\theta > 0$) which is reflected at the origin. The process defined by

$$Y'_i(t) := X_i(t) - \bar{X}(t) + \frac{\beta}{\sqrt{N}}, \quad i = 1, 2, \dots, N,$$

has the same spacings as the original process X . Moreover, by the independence of the centered system and the center of mass process established in Lemma 7, and the fact that $\bar{Y}' = \beta/\sqrt{N}$ is the projection of Y' on the subspace generated by the vector $(1, 1, \dots, 1)$, it follows that the vector $(Y'_1, Y'_2, \dots, Y'_N)$ is a Brownian motion with identity covariance matrix which is normally reflected in the wedge

$$\left\{ y \in \mathbb{R}^N : \sum_{i=1}^N y_i \geq 0, y_1 \leq y_2 \leq \dots \leq y_N \right\}.$$

The drift vector in this wedge can be written as a linear combination of spacings by summation by parts:

$$(15) \quad \sum_{i=1}^N \left(\delta_i - \bar{\delta}_N - \frac{\theta}{\sqrt{N}} \right) y_i = - \sum_{k=1}^{N-1} \left(\sum_{i=1}^k (\delta_i - \bar{\delta}_N) \right) (y_{k+1} - y_k) - \frac{\theta}{\sqrt{N}} \sum_{i=1}^N y_i.$$

Noting that $Y'_{(k+1)} - Y'_{(k)} \equiv X_{(k+1)} - X_{(k)}$ for all k , the condition (13) for stability, and part (i) of the Lemma, are now read from the general result about equilibrium distributions of reflecting Brownian motions stated in the following lemma, and standard theory of Harris recurrent Markov processes (see, e.g., [8]).

Part (ii) is established by showing that the centered system $(X_i(t) - \bar{X}(t), 1 \leq i \leq N), t \geq 0$ is a Harris positive-recurrent diffusion with the indicated invariant measure. The recurrence property has been proved in detail in [3]. The invariance is evident because a uniform randomization of labels relative to the centered order statistics is clearly invariant for the centered motion. This convergence in distribution, combined with part (i) of Lemma 6 gives convergence in probability of $X_i(t)/t$ to $\bar{\delta}_N$ for each $1 \leq i \leq N$. Almost sure convergence can now be justified by appeal to an ergodic theorem for the Harris recurrent centered diffusion. \square

The proof of Theorem 8 is completed by the following lemma, which we deduce from the general theory of stationary distributions for reflecting Brownian motions in polyhedra due to R. Williams [30]:

LEMMA 9. Let $R := (R_t, t \geq 0)$ be a Brownian motion in the domain

$$\{x \in \mathbb{R}^K : b_i(x) \geq 0 \text{ for } i = 1, \dots, K\}$$

for some collection of K linearly independent linear functionals b_i , with X having identity covariance matrix, normal reflection at the boundary, and constant drift δ with

$$(16) \quad \sum_{i=1}^K \delta_i x_i = - \sum_{i=1}^K a_i b_i(x) \quad (x_1, \dots, x_K) = x \in \mathbb{R}^K.$$

This process R has a stationary probability distribution iff $a_i > 0$ for all $i = 1, \dots, K$, in which case in the stationary state the $b_i(R_t)$ are independent exponential variables with rates $2a_i$, and the process in its stationary state is reversible.

PROOF. This is read from the particular case of [30, Theorem 1.2] when the matrix Q is identically 0. According to that theorem, R is in duality with itself relative to the measure ρ on the domain whose density function with respect to Lebesgue measure at x is $\exp(2\delta \cdot x)$. Using (16), the linear change of variables to $y_i = b_i(x)$ shows that the ρ distribution of the $b_i(x)$ has joint density at (y_1, \dots, y_K) equal to $c \prod_{i=1}^K \exp(-2a_i y_i)$ for some $c > 0$. It follows that

$$(17) \quad \rho \text{ has finite total mass iff } a_i > 0 \text{ for all } i$$

in which case, when ρ is normalized to be a probability, the ρ distribution of the $b_i(x)$ is that of independent exponential variables with rates $2a_i$. The “if” part of the conclusion is now evident. For the “only if” part we argue that if a stationary probability distribution ρ' existed, it would obviously have a strictly positive density on the domain. Then R sampled at time $0, 1, 2, \dots$ would be an irreducible Harris recurrent Markov process with respect ρ' , hence $\rho = c\rho'$ for some $c > 0$ by the uniqueness of the invariant measure of a Harris recurrent Markov chain, and then $a_i > 0$ for all i by (17). \square

From Theorem 8 we immediately deduce:

COROLLARY 10. For each $\delta > 0$ the N -particle Atlas system with drift vector $(\delta, 0, \dots, 0)$ is ergodic with average speed δ/N . The stationary distribution of

$$(X_{(j+1)} - X_{(j)}, 1 \leq j \leq N-1)$$

is that of independent exponentials $(\zeta_j, 1 \leq j \leq N-1)$, where the rate of ζ_j is $2\delta(1 - j/N)$.

3. The infinite Atlas model. The infinite Atlas model can be described loosely as a countable collection of linear Brownian motions such that at every time point the minimum Brownian motion is given a positive drift of $\delta > 0$, and the rest are left untouched. This is an example of (1) where $I = \mathbb{N}$, $\delta_1 = \delta$ and all other $\delta_i = 0$. Throughout this section we take $\delta = 1$, since for our purpose here, the general case follows from the case when $\delta = 1$ by scaling.

It can be constructed rigorously in the weak solution framework in the following way. Start with the canonical setup (as in Lemma 6) of Brownian path-space:

$$(18) \quad (C[0, \infty], \{ \mathcal{F}_t \}_{0 \leq t < \infty}, \mathbb{W}^x),$$

where $\{ \mathcal{F}_t \}$ is the right continuous filtration generated by the coordinates which satisfy the usual conditions, and \mathbb{W}^x is the law of the Brownian motion starting from $x \in \mathbb{R}$. We now look at the countable product of a sequence of spaces like (18):

$$(19) \quad \Omega = C[0, \infty]^{\mathbb{N}}, \quad \mathcal{G}_t = \bigotimes_1^{\infty} \mathcal{F}_t(i), \quad P^x = \bigotimes_1^{\infty} \mathbb{W}^{x_i},$$

where the natural coordinate mapping is a countable collection of independent Brownian motions under P^x starting from the sequence $x = (x_1, x_2, \dots)$.

Let x be a sequence such that $x_{(1)} > -\infty$ and let $X = (X_1, X_2, \dots)$ denote the sequence of infinite independent Brownian motions starting from x . We have the following lemma whose proof will follow later.

LEMMA 11. *Assume that the initial sequence x is arranged in increasing order $x_1 \leq x_2 \leq \dots$ and satisfies*

$$(20) \quad \liminf_{n \rightarrow \infty} \frac{(x_n - x_1)^2}{n} > 0.$$

Then, P^x -almost surely, $X_{(1)}(t) > -\infty$ for all $t \geq 0$, and the process

$$(21) \quad N_t := \sum_{i=1}^{\infty} \int_0^t 1_{\{X_i(s) = X_{(1)}(s)\}} dX_i(s).$$

is a $\{ \mathcal{G}_t \}$ -martingale whose quadratic variation $\langle N \rangle_t \equiv t$. The stochastic exponential of N , given by

$$(22) \quad D_t = \exp(N_t - t/2).$$

is hence again a non-negative $\{ \mathcal{G}_t \}$ -martingale.

We change the measure P^x by using the martingale D , i.e., define Q by

$$Q^x|_{\mathcal{G}_t} := D_t \cdot P^x|_{\mathcal{G}_t}, \quad t \geq 0.$$

By Girsanov's theorem the probability measure Q^x exists and is well-defined, and the coordinate process under Q^x is a solution of the infinite Atlas model (1). Hence Q^x is the *unique* law of the Atlas model starting at x .

Our aim is the following: suppose the initial points $X_1(0) < X_2(0) < X_3(0) < \dots$ are spread according to the Poisson process with rate two on the positive half-line, and we run the infinite Atlas model starting from these points. We shall prove that the product law of independent Exponential(2) is invariant under the dynamics of the vector process Δ of spacings given by

$$(23) \quad \Delta_i(t) = X_{(i+1)}(t) - X_{(i)}(t), \quad i = 1, 2, \dots$$

The proof is achieved through a series of lemmas, the main argument being comparing the infinite Atlas model with the finite Atlas model and suitably passing to the limit. Throughout the proof the probability space is given by (19).

We start with the following lemma whose proof follows directly from Lemma 5. We will find it convenient to use the following notation for the operator which sorts a given finite vector. If $x \in \mathbb{R}^n$, $1 \leq n < \infty$, define

$$(24) \quad \mathcal{S}(x) = (x_{(1)}, x_{(2)}, \dots, x_{(n)}).$$

LEMMA 12. *Every $\omega \in \Omega$ comprises of a sequence of processes $\omega(t) = (\omega_1(t), \omega_2(t), \dots)$. For any $N \in \mathbb{N}$, let us denote the ordered values of the processes with the first N indices by*

$$Z^N(\omega) = (Z_1^N, Z_2^N, \dots, Z_N^N)(\omega) = \mathcal{S}((\omega_1, \omega_2, \dots, \omega_N)).$$

Then

$$(25) \quad D_N(t) = \exp \left(\sum_{i=1}^N \int_0^t 1_{\{X_i(s)=Z_i^N(s)\}} dX_i(s) - t/2 \right)$$

is a \mathcal{G}_t -martingale. We denote by Q_N^x the probability measure obtained by the changing the measure P^x by the martingale D_N .

Since, under every P^x , each of the Brownian motions are independent, it follows, by applying Girsanov's Theorem, that, under Q_N^x , the first N

coordinates $(\omega_1, \dots, \omega_N)$ evolve according to the finite Atlas model, while the rest of the coordinates are independent Brownian motions with the corresponding initial starting points.

Let μ denote the probability measure whereby the points $0 = X_1(0) < X_2(0) < \dots$ are such that the spacings $X_{i+1}(0) - X_i(0)$ are iid Exponential(2).

LEMMA 13. *For μ -almost every x , the measure Q^x exists and we can define*

$$Q \cdot \mu = \int Q^x d\mu(x).$$

PROOF. The proof follows from Lemma 11 and the law of large numbers.

■

We now have our main theorem in this section which proves that μ is an invariant measure for the spacings of the infinite Atlas model.

THEOREM 14. *For any $K \in \mathbb{N}$, and any function $F : \mathbb{R}^K \rightarrow \mathbb{R}$, which is smooth and has compact support, and for any time t , we have*

$$(26) \quad \mathbb{E}^{Q \cdot \mu} [F(\Delta_1, \Delta_2, \dots, \Delta_K)(t)] = \mathbb{E}^\mu [F(\Delta_1, \Delta_2, \dots, \Delta_K)(0)].$$

Here, as defined in (23), Δ_i is the i th spacing $X_{(i+1)} - X_{(i)}$.

PROOF OF THEOREM 1. Since the previous result holds for a class of functions which determines the marginal distributions of a sequence valued process, Theorem 1 follows readily for $\delta = 1$. The theorem for the infinite Atlas model with a general $\delta > 0$ follows by scaling. ■

PROOF OF LEMMA 11. Fix an arbitrary time $T > 0$, and consider x as in the assumption of the lemma. We have the following claim.

CLAIM 15.

$$P^x \left(\omega : \exists M(\omega) \text{ s.t. } \forall n \geq M, \inf_{0 \leq s \leq T} (X_n(s) - X_1(s)) > 0 \right) = 1.$$

- We prove the above claim by establishing the following

$$(27) \quad P^x \left(\omega : \exists M(\omega) \text{ s.t. } \forall n \geq M, \inf_{0 \leq s \leq T} (X_n(s) - X_1(s)) > 0 \right) = 1.$$

This follows by defining the events

$$(28) \quad A_i := \left\{ \inf_{0 \leq s \leq T} (X_i(s) - X_1(s)) \leq 0 \right\}, \quad i = 1, 2, \dots$$

Since $X_i - X_1$ is a Brownian motion starting from $(x_i - x_1)$, which for sufficiently large i is strictly positive (by (20)), by Bernstein's inequality [23, page 153], one can easily estimate $P^x(A_i) \leq \exp(-(x_i - x_1)^2/2T)$ for all large enough i . Again by assumption (20), we get

$$\begin{aligned} \limsup_{n \rightarrow \infty} (P^x(A_n))^{1/n} &\leq \limsup_n \exp(-(x_n - x_1)^2/2Tn) \\ &= \exp\left(-\frac{1}{2T} \liminf_n \frac{(x_n - x_1)^2}{n}\right) < 1. \end{aligned}$$

Thus, by Cauchy's root test, it follows that the series sum $\sum_{i=1}^{\infty} P^x(A_i) < \infty$. One can now apply the Borel-Cantelli lemma to obtain that $P^x(\limsup_i A_i) = 0$, which proves (27), and hence the required claim.

We now prove that the process N exists and is a martingale in time interval $[0, T]$. Since T is arbitrary, this proves Lemma 11. Define the finite approximations

$$B_k(t) = \sum_{i=1}^k \int_0^t 1_{\{X_i(s)=X_{(1)}(s)\}} dX_i(s), \quad 0 \leq t \leq T, \quad k = 1, 2, \dots$$

Each B_k is a stochastic integral with bounded, progressively measurable integrands. It is clear that they are martingales with quadratic variations

$$\begin{aligned} \langle B_k \rangle_t &= \sum_{i=1}^k \int_0^t 1_{\{X_i(s)=X_{(1)}(s)\}} ds \\ &= \mathbb{L} \left\{ 0 \leq s \leq t : \min_{1 \leq i \leq k} X_i(s) = X_{(1)}(s) \right\}, \end{aligned}$$

where \mathbb{L} refers to the Lebesgue measure on the line. We shall show that the sequence of martingales $\{B_k\}$ is a cauchy sequence in the \mathbb{H}^2 -norm, and hence has a limit which is denoted by N , as in (21).

To see this, observe that for any $n, k \in \mathbb{N}$, we get

$$\langle B_{n+k} - B_n \rangle_T = \mathbb{L} \left\{ 0 \leq s \leq T : \min_{n+1 \leq i \leq n+k} X_i(s) = X_{(1)}(s) \right\}.$$

By Claim 15, we see that $P^x(\lim_{n,k \rightarrow \infty} \langle B_{n+k} - B_n \rangle_T = 0) = 1$. It is also clear that $\langle B_{n+k} - B_n \rangle_T \leq T$, for all $n, k \in \mathbb{N}$. It follows from the Dominated Convergence Theorem that

$$\lim_{n,k \rightarrow \infty} E^{P^x}(\langle B_{n+k} - B_n \rangle_T) = 0.$$

This, by definition, shows that the sequence of martingales $\{B_n\}$ is a Cauchy sequence in the \mathbb{H}^2 norm. Since the space of continuous martingales under that norm is complete, the sequence $\{B_n\}$ converges to a limiting martingale which we denote N . It also follows that $\langle N \rangle_t = \lim_{n \rightarrow \infty} \langle B_n \rangle_t = t$. Thus, N is actually a $\{\mathcal{G}_t\}$ -Brownian motion. This completes the proof of Lemma 11. \blacksquare

Remark. The condition in (20) is clearly loose. In fact, if we consider A_i as in (28), all we require is that, for every $T > 0$, one should have

$$\sum_{i=1}^{\infty} P^x(A_i) \leq \sum_{i=1}^{\infty} \exp\left(-\frac{(x_i - x_1)^2}{2T}\right) < \infty,$$

which is a much weaker condition than required by (20).

The proof of Theorem 14 relies on probability estimates proved in the next three Lemmas 17, 18, and 19. The second one is actually a generalization of the first. But, we choose to do the first separately since it is simpler and more transparent. But first we will do some basic computations.

LEMMA 16. *If $Y \sim \text{Gamma}(r, \lambda)$ for some $r \geq 1$, then, for any $t > 0$, we have*

$$\begin{aligned} \mathbb{E}\left(e^{-Y^2/2t}\right) &\leq e^{t\lambda^2/2} \left(2\lambda^2 t\right)^{r/2} \frac{\Gamma(r/2)}{2\Gamma(r)} \\ &= \sqrt{\pi} \left(\frac{\lambda^2 t}{2}\right)^{r/2} \frac{e^{t\lambda^2/2}}{\Gamma((r+1)/2)}. \end{aligned}$$

PROOF.

$$\begin{aligned} \mathbb{E}\left(e^{-Y^2/2t}\right) &= \int_0^{\infty} e^{-y^2/(2t)} \frac{\lambda^r}{\Gamma(r)} y^{r-1} e^{-\lambda y} dy \\ (29) \quad &= \frac{\lambda^r}{\Gamma(r)} e^{t\lambda^2/2} \int_0^{\infty} y^{r-1} e^{-(y+t\lambda)^2/(2t)} dy \\ &= \frac{\lambda^r}{\Gamma(r)} e^{t\lambda^2/2} \int_{\sqrt{t}\lambda}^{\infty} (\sqrt{t}z - t\lambda)^{r-1} e^{-z^2/2} \sqrt{t} dz, \quad z = \frac{y + t\lambda}{\sqrt{t}}. \end{aligned}$$

For $z \geq \sqrt{t}\lambda$, since $r \geq 1$, one has

$$(\sqrt{tz} - t\lambda)^{r-1} \leq (\sqrt{tz})^{r-1}.$$

Thus one can bound the (29) by

$$\begin{aligned} \mathbb{E} \left(e^{-Y^2/2t} \right) &\leq \frac{\lambda^r}{\Gamma(r)} e^{t\lambda^2/2} \int_{\sqrt{t}\lambda}^{\infty} (\sqrt{tz})^{r-1} e^{-z^2/2} \sqrt{t} dz \\ &= \frac{(\lambda^2 t)^{r/2}}{\Gamma(r)} e^{t\lambda^2/2} \int_{\sqrt{t}\lambda}^{\infty} z^{r-1} e^{-z^2/2} dz \\ &= \frac{(\lambda^2 t)^{r/2}}{\Gamma(r)} e^{t\lambda^2/2} \int_{t\lambda^2/2}^{\infty} (2w)^{(r-1)/2} e^{-w} (2w)^{-1/2} dw, \quad w = z^2/2 \\ &= \frac{(\lambda^2 t)^{r/2}}{\Gamma(r)} e^{t\lambda^2/2} 2^{r/2-1} \int_{t\lambda^2/2}^{\infty} w^{r/2-1} e^{-w} dw \\ &\leq \frac{(2\lambda^2 t)^{r/2}}{2\Gamma(r)} e^{t\lambda^2/2} \Gamma(r/2). \end{aligned}$$

The final identity in the lemma is due to the *duplication formula*:

$$\frac{\Gamma(r/2)}{\Gamma(r)} = \sqrt{\pi} \frac{2^{1-r}}{\Gamma((r+1)/2)}.$$

■

LEMMA 17. *For any $t > 0$, and all positive integers N satisfying $N+1 \geq 16et$, under $P \cdot \mu$, we have the following estimate of the probability of the event that during time $[0, t]$, the globally lowest ranked process is in fact the lowest ranked process among the processes with the first N indices.*

$$P \cdot \mu \left\{ X_{(1)}(s) = Z_1^N(s), \quad 0 \leq s \leq t \right\} \geq 1 - C_1 e^{2t} \left(\sqrt{\frac{4et}{N+1}} \right)^N,$$

where C_1 is a positive constant. To remind the reader, the processes $(Z_k^N, 1 \leq k \leq N)$ have been defined in Lemma 12.

PROOF. Note that the complement of the event has the following upper bound.

$$\begin{aligned} &1 - P \cdot \mu \left\{ X_{(1)}(s) = Z_1^N(s), \quad 0 \leq s \leq t \right\} \\ &\leq P \cdot \mu \left(\cup_{i \geq N+1} \{ X_i(s) \leq Z_1^N(s), \text{ for some } s \in [0, t] \} \right) \\ (30) \quad &\leq P \cdot \mu \left(\cup_{i \geq N+1} \{ X_i(s) \leq X_1(s), \text{ for some } s \in [0, t] \} \right) \\ &\leq \sum_{i=N+1}^{\infty} P \cdot \mu \left(X_i(s) \leq X_1(s), \text{ for some } s \in [0, t] \right). \end{aligned}$$

The final bound above is the so-called union bound.

Now we will use exponential bounds for Brownian supremums to estimate $P \cdot \mu (X_i(s) \leq X_1(s), \text{ for some } s \in [0, t])$ which is the same as $P \cdot \mu (\inf_{0 \leq s \leq t} (X_i(s) - X_1(s)) \leq 0)$. Note that under $P \cdot \mu$, the process $X_i - X_1$ is a Brownian motion whose initial distribution is the law of $X_i(0) - X_1(0)$. This law is $\text{Gamma}(i-1, 2)$, since it is the sum of $(i-1)$ iid $\text{Exp}(2)$. Conditional on $X(0) = x$, such that $\{X_i(0) - X_1(0) = y\}$, it follows from Bernstein's inequality that

$$P^x \left(\inf_{0 \leq s \leq t} (X_i(s) - X_1(s)) \leq 0 \right) \leq \exp(-y^2/2t).$$

Thus, for $i \geq 2$, we define $Y = X_i(0) - X_1(0) \sim \text{Gamma}(i-1, 2)$, and we use Lemma 16 to get

$$(31) \quad \begin{aligned} P \cdot \mu \left(\inf_{0 \leq s \leq t} (X_i(s) - X_1(s)) \leq 0 \right) &\leq \mathbb{E}(\exp(-Y^2/2t)) \\ &\leq \sqrt{\pi}(2t)^{(i-1)/2} \frac{e^{2t}}{\Gamma(i/2)}. \end{aligned}$$

Plugging the estimate in (30), we get

$$(32) \quad 1 - P \cdot \mu \left\{ X_{(1)} = Z_1^N(s), 0 \leq s \leq t \right\} \leq \sqrt{\pi} e^{2t} \sum_{i=N+1}^{\infty} \frac{(2t)^{(i-1)/2}}{\Gamma(i/2)}.$$

By Stirling's approximation, $\exists C$, a positive constant, such that

$$(33) \quad \Gamma(z) \geq C^{-1} e^{-z} z^{z-1/2}, \quad \forall z \in \mathbb{R}^+.$$

Thus, from (32) we get

$$\begin{aligned} 1 - P \cdot \mu \left\{ X_{(1)} = Z_1^N(s), 0 \leq s \leq t \right\} &\leq C \sqrt{\pi} e^{2t} \sum_{i=N+1}^{\infty} \frac{(2t)^{(i-1)/2} e^{i/2}}{(i/2)^{(i-1)/2}} \\ &= C \sqrt{\pi} e^{2t} \sum_{i=N}^{\infty} \left(\frac{4et}{i+1} \right)^{i/2} e^{1/2} \leq \sqrt{\pi} C e^{2t+1/2} \sum_{i=N}^{\infty} \left(\sqrt{\frac{4et}{N+1}} \right)^i \\ &\leq 2C \sqrt{\pi} e^{2t+1/2} \left(\sqrt{\frac{4et}{N+1}} \right)^N, \quad \text{when } \sqrt{4et/(N+1)} \leq 1/2. \end{aligned}$$

This proves the estimate. ■

LEMMA 18. *For any any $t > 0$ and all positive integers $1 \leq k < N \in \mathbb{N}$ satisfying $N - k + 2 \geq 16et$, under $P \cdot \mu$, we have the following estimate of the probability of the event that in the time interval $[0, t]$, the globally lowest k ranked processes are in fact the lowest k ranked processes among the ones with the first N indices.*

$$(34) \quad \begin{aligned} P \cdot \mu \left\{ \left(X_{(1)}, \dots, X_{(k)} \right) (s) = \left(Z_1^N, \dots, Z_k^N \right) (s), \quad 0 \leq s \leq t \right\} \\ \geq 1 - C_2(k) e^{2t} (N - 1)^{k-1} \left(\sqrt{\frac{4et}{N - k + 2}} \right)^{N-k+1}, \end{aligned}$$

for some positive constant C_2 depending on k . For $k = 1$, we get back the bound in Lemma 17.

PROOF. As in the previous lemma, we bound the probability of complement of the event.

$$(35) \quad \begin{aligned} 1 - P \cdot \mu \left\{ \left(X_{(1)}, \dots, X_{(k)} \right) (s) = \left(Z_1^N, \dots, Z_k^N \right) (s), \quad 0 \leq s \leq t \right\} \\ \leq P \cdot \mu \left(\cup_{i \geq N+1} \left\{ X_i(s) \leq Z_k^N(s), \quad \text{for some } s \in [0, t] \right\} \right) \\ \leq \sum_{i=N+1}^{\infty} P \cdot \mu \left\{ X_i(s) \leq Z_k^N(s), \quad \text{for some } s \in [0, t] \right\}. \end{aligned}$$

Now, if $[N]$ denotes the set $\{1, 2, \dots, N\}$, let us note that

$$Z_k^N = \max_{i_1 < i_2 < \dots < i_{k-1}} \min_{l \in [N] \setminus \{i_1, i_2, \dots, i_{k-1}\}} X_l.$$

Thus, to get an upper bound on $P \cdot \mu \left\{ X_i(s) \leq Z_k^N(s), \quad \text{for some } s \in [0, t] \right\}$, one can apply the union bound one more time to obtain

$$(36) \quad \begin{aligned} P \cdot \mu \left\{ X_i(s) \leq Z_k^N(s), \quad \text{for some } s \in [0, t] \right\} \leq \\ \sum_{i_1 < \dots < i_{k-1}} P \cdot \mu \left(X_i(s) \leq \min_{l \in [N] \setminus \{i_1, \dots, i_{k-1}\}} X_l(s), \quad \text{for some } s \in [0, t] \right) \leq \\ \sum_{i_1 < \dots < i_{k-1}} P \cdot \mu (X_i(s) \leq X_{i^*}, \text{ for some } s \in [0, t], i^* = \min\{[N] \setminus \{i_1, \dots, i_{k-1}\}\}) \end{aligned}$$

Now, we can count the frequency with which i^* takes its possible values as the choice $i_1 < i_2 < \dots < i_{k-1}$ varies in $\{1, 2, \dots, N\}$. Let $g(i)$ be the

number of ways to pick $i_1 < i_2 < \dots < i_{k-1}$ such that $i^* = i$. Then, it is straightforward to see

$$\begin{aligned} g(1) &= \#\{ \{i\} : i_1 > 1 \} = \binom{N-1}{k-1}, \\ g(2) &= \#\{ \{i\} : i_1 = 1, i_2 > 2 \} = \binom{N-2}{k-2} \\ &\vdots \\ g(l) &= \#\{ \{i\} : i_1 = 1, i_2 = 2, \dots, i_{l-1} = l-1, i_l > l \} = \binom{N-l}{k-l}, \quad l \leq k, \\ g(l) &= 0, \quad \forall l > k. \end{aligned}$$

Thus, by (36), we get

$$\begin{aligned} (37) \quad & \sum_{i_1 < i_2 < \dots < i_{k-1}} P \cdot \mu \left(X_i(s) \leq \min_{l \in [N] \setminus \{i_1, \dots, i_{k-1}\}} X_l(s), \quad \text{for some } s \in [0, t] \right) \\ & \leq \sum_{l=1}^k \binom{N-l}{k-l} P \cdot \mu (X_i(s) \leq X_l(s), \text{ for some } s \in [0, t]). \end{aligned}$$

Now, again as we did when deriving the last key estimate, $X_i - X_l$ for $i > l$ is a Brownian motion under $P \cdot \mu$ with the initial distribution $\text{Gamma}(i-l, 2)$. Thus, by Lemma 16, we can bound

$$P \cdot \mu (X_i(s) \leq X_l(s), \text{ for some } s \in [0, t]) \leq \sqrt{\pi} \frac{(2t)^{(i-l)/2} e^{2t}}{\Gamma((i-l+1)/2)}.$$

Combining the above inequality with estimates (35), (36), and (37), we get

$$\begin{aligned} (38) \quad & 1 - P \cdot \mu \left\{ (X_{(1)}, \dots, X_{(k)}) (s) = (Z_1^N, \dots, Z_k^N) (s), \quad 0 \leq s \leq t \right\} \\ & \leq \sum_{i=N+1}^{\infty} \sum_{l=1}^k \binom{N-l}{k-l} \sqrt{\pi} e^{2t} \frac{(2t)^{(i-l)/2}}{\Gamma((i-l+1)/2)}. \end{aligned}$$

We will again give a loose but good enough upper bound for the infinite sum. But, first we need to note that, for any $n \geq k$, we have

$$\binom{n}{k} / \binom{n-1}{k-1} = \frac{n}{k} \geq 1.$$

Thus, it follows that

$$\binom{N-l}{k-l} \leq \binom{N-1}{k-1}, \quad \forall 1 \leq l \leq k.$$

We can use this to simplify (38):

$$\begin{aligned} 1 - P \cdot \mu \left\{ (X_{(1)}, \dots, X_{(k)}) (s) = (Z_1^N, \dots, Z_k^N) (s), \quad 0 \leq s \leq t \right\} \\ \leq \sum_{i=N+1}^{\infty} \sum_{l=1}^k \binom{N-1}{k-1} \sqrt{\pi} e^{2t} \frac{(2t)^{(i-l)/2}}{\Gamma((i-l+1)/2)} \end{aligned}$$

We again use Stirling's approximation (33) to get

$$\begin{aligned} (39) \quad & \sum_{i=N+1}^{\infty} \sum_{l=1}^k \binom{N-1}{k-1} \sqrt{\pi} e^{2t} \frac{(2t)^{(i-l)/2}}{\Gamma((i-l+1)/2)} \\ & \leq C \binom{N-1}{k-1} \sqrt{\pi} e^{2t} \sum_{i=N+1}^{\infty} \sum_{l=1}^k \frac{(2t)^{(i-l)/2} e^{(i-l+1)/2}}{((i-l+1)/2)^{(i-l)/2}} \\ & = C \binom{N-1}{k-1} \sqrt{\pi} e^{2t+1/2} \sum_{i=N+1}^{\infty} \sum_{l=1}^k \left(\sqrt{\frac{4et}{i-l+1}} \right)^{i-l}. \end{aligned}$$

Since $i-l \geq N+1-k$ for $l \leq k < N+1 \leq i$, it is clear that $4et/(i-l+1) \leq 4et/(N+2-k)$. And hence

$$\sum_{l=1}^k \left(\sqrt{\frac{4et}{i-l+1}} \right)^{i-l} \leq \sum_{l=1}^k \left(\sqrt{\frac{4et}{N-k+2}} \right)^{i-l}.$$

By our assumption

$$(40) \quad \sqrt{\frac{4et}{N-k+2}} \leq \frac{1}{2}$$

Thus

$$\left(\sqrt{\frac{4et}{N-k+2}} \right)^{i-l} \leq \left(\sqrt{\frac{4et}{N-k+2}} \right)^{i-k}, \quad l \leq k < i,$$

and consequently

$$\sum_{l=1}^k \left(\sqrt{\frac{4et}{i-l+1}} \right)^{i-l} \leq k \left(\sqrt{\frac{4et}{N-k+2}} \right)^{i-k}.$$

Plugging this bound in (39) we get

$$\begin{aligned}
(41) \quad & \sum_{i=N+1}^{\infty} \sum_{l=1}^k \binom{N-1}{k-1} \sqrt{\pi} e^{2t} \frac{(2t)^{(i-l)/2}}{\Gamma((i-l+1)/2)} \\
& \leq C \binom{N-1}{k-1} \sqrt{\pi} e^{2t+1/2} \sum_{i=N+1}^{\infty} k \left(\sqrt{\frac{4et}{N-k+2}} \right)^{i-k} \\
& \leq Ck \binom{N-1}{k-1} \sqrt{\pi} e^{2t+1/2} \left(\sqrt{\frac{4et}{N-k+2}} \right)^{N+1-k} \sum_{j=0}^{\infty} \left(\sqrt{\frac{4et}{N-k+2}} \right)^j \\
& \leq C_2(k) e^{2t} (N-1)^{k-1} \left(\sqrt{\frac{4et}{N-k+2}} \right)^{N-k+1}, \text{ by (40),}
\end{aligned}$$

for some positive constant C_2 depending on k . This proves the lemma. ■

LEMMA 19. *For any $N \in \mathbb{N}$, define μ_N to be the law under which*

$$\begin{aligned}
(42) \quad & X_1(0) = 0, \quad X_{i+1}(0) - X_i(0) \sim \text{Exp}(2(1 - i/N)), \quad i = 1, 2, \dots, N-1, \\
& \text{and } X_{i+1}(0) - X_i(0) \sim \text{Exp}(2), \quad i = N, N+1, \dots,
\end{aligned}$$

and all these spacings are independent.

For any $t > 0$, and any three integers $k < J < N$ satisfying

$$(43) \quad J - k + 2 \geq 16et,$$

we have the following bound on the probability that under $P \cdot \mu_N$, during the time interval $[0, t]$, the lowest k ranked processes among the processes with the first N indices are in fact the lowest k among those with the first J indices.

$$\begin{aligned}
(44) \quad & P \cdot \mu_N \left\{ \left(Z_1^J, \dots, Z_k^J \right) (s) = \left(Z_1^N, \dots, Z_k^N \right) (s), \quad 0 \leq s \leq t \right\} \\
& \geq 1 - C_2(k) e^{2t} (J-1)^{k-1} \left(\sqrt{\frac{4et}{J-k+2}} \right)^{J-k+1} \left[1 - \left(\sqrt{\frac{4et}{J-k+2}} \right)^{N-J} \right].
\end{aligned}$$

PROOF. We follow the same line of argument as in the last two estimates. Thus

$$\begin{aligned} & 1 - P \cdot \mu_N \left\{ \left(Z_1^J, \dots, Z_k^J \right) (s) = \left(Z_1^N, \dots, Z_k^N \right) (s), \quad 0 \leq s \leq t \right\} \\ & \leq \sum_{i=J+1}^N P \cdot \mu_N \left(X_i(s) \leq Z_k^J(s), \quad \text{for some } s \in [0, t] \right). \end{aligned}$$

Following similar counting arguments as in (36) and (37), we can bound

$$(45) \quad \begin{aligned} & 1 - P \cdot \mu_N \left\{ \left(Z_1^J, \dots, Z_k^J \right) (s) = \left(Z_1^N, \dots, Z_k^N \right) (s), \quad 0 \leq s \leq t \right\} \\ & \leq \sum_{i=J+1}^N \sum_{l=1}^k \binom{J-l}{k-l} P \cdot \mu_N \left(X_i(s) \leq X_l(s), \quad \text{for some } s \in [0, t] \right). \end{aligned}$$

Now, under μ_N , the gap $X_i(0) - X_l(0) = \sum_{j=l}^{i-1} Y_j$, where the Y_j 's are independent and Y_j is distributed as $\text{Exp}(2(1 - j/N))$. Thus, by the exponential bound used before, we get

$$(46) \quad \begin{aligned} & P \cdot \mu_N \left(X_i(s) \leq X_l(s), \quad \text{for some } s \in [0, t] \right) \\ & \leq \mathbb{E} \left[\exp \left(-\frac{1}{2t} \left(\sum_{j=l}^{i-1} Y_j \right)^2 \right) \right]. \end{aligned}$$

Now, since each Y_j is $\text{Exp}(2(1 - j/N))$, the random variables

$$Y_j^* = (1 - j/N)Y_j$$

are iid $\text{Exp}(2)$, and each $Y_j^* \leq Y_j$ since $j \leq N$. Thus, $\sum_{j=l}^{i-1} Y_j^* \leq \sum_{j=l}^{i-1} Y_j$, and hence

$$\mathbb{E} \left[\exp \left(-\frac{1}{2t} \left(\sum_{j=l}^{i-1} Y_j \right)^2 \right) \right] \leq \mathbb{E} \left[\exp \left(-\frac{1}{2t} \left(\sum_{j=l}^{i-1} Y_j^* \right)^2 \right) \right].$$

But, each Y_j^* is $\text{Exp}(2)$, and hence

$$\mathbb{E} \left[\exp \left(-\frac{1}{2t} \left(\sum_{j=l}^{i-1} Y_j^* \right)^2 \right) \right] = \mathbb{E} \left[e^{-Y^2/2t} \right],$$

where Y is a $\text{Gamma}(i-l, 2)$ random variable. Thus, plugging in the bound from Lemma 16 into (46), we derive

$$P \cdot \mu_N \left(X_i(s) \leq X_l(s), \quad \text{for some } s \in [0, t] \right) \leq \sqrt{\pi} \frac{e^{2t}(2t)^{(i-l)/2}}{\Gamma((i-l+1)/2)}.$$

Now, we follow approximations similar to (39) and (41) for the right-hand-side on (45) to obtain

$$\begin{aligned}
& 1 - P \cdot \mu_N \left\{ \left(Z_1^J, \dots, Z_k^J \right) (s) = \left(Z_1^N, \dots, Z_k^N \right) (s), \quad 0 \leq s \leq t \right\} \\
& \leq \sum_{i=J+1}^N \sum_{l=1}^k \binom{J-1}{k-1} \sqrt{\pi} \frac{e^{2t} (2t)^{(i-l)/2}}{\Gamma((i-l+1)/2)} \\
& = C \binom{J-1}{k-1} \sqrt{\pi} e^{2t+1/2} \sum_{i=J+1}^N \sum_{l=1}^k \left(\sqrt{\frac{4et}{(i-l+1)}} \right)^{i-l} \\
& \leq C \binom{J-1}{k-1} \sqrt{\pi} e^{2t+1/2} \sum_{i=J+1}^N k \left(\sqrt{\frac{4et}{(J-k+2)}} \right)^{i-k}, \text{ by (43),} \\
& \leq C_2(k) e^{2t} (J-1)^{k-1} \left(\sqrt{\frac{4et}{J-k+2}} \right)^{J-k+1} \sum_{j=0}^{N-J-1} \left(\sqrt{\frac{4et}{J-k+2}} \right)^j.
\end{aligned}$$

Now, if we call $r = \sqrt{4et/(J-k+2)}$, then $r \leq 1/2$ by assumption (43). The finite geometric sum can be easily bounded as

$$\sum_{j=0}^{N-J-1} r^j = \frac{1 - r^{N-J}}{1 - r} \leq 2(1 - r^{N-J}).$$

Thus, suitably altering the constant C_2 , we get our desired bound

$$\begin{aligned}
& 1 - P \cdot \mu_N \left\{ \left(Z_1^J, \dots, Z_k^J \right) (s) = \left(Z_1^N, \dots, Z_k^N \right) (s), \quad 0 \leq s \leq t \right\} \\
& \leq C_2(k) e^{2t} (J-1)^{k-1} \left(\sqrt{\frac{4et}{J-k+2}} \right)^{J-k+1} \left[1 - \left(\sqrt{\frac{4et}{J-k+2}} \right)^{N-J} \right].
\end{aligned}$$

This proves the lemma. ■

PROOF OF THEOREM 14. For every $K \leq N \in \mathbb{N}$, let

$$Y_i^N = Z_{i+1}^N - Z_i^N, \quad 1 \leq i \leq N.$$

From Corollary 10, we know that the law of the first $(N-1)$ spacings under μ_N (defined in Lemma 19) is exactly the stationary distribution of spacings for the finite Atlas model with N particles. Since, under Q_N (see Lemma

12), the dynamics of the processes $\{X_1(t), \dots, X_N(t)\}$ is that of a finite Atlas model independent of the rest of the Brownian motions, it follows by stationarity that for any $t > 0$, we have

$$(47) \quad \mathbb{E}^{Q_N \cdot \mu_N} \left[F(Y_1^N, \dots, Y_K^N)(t) \right] = \mathbb{E}^{\mu_N} \left[F(Y_1^N, \dots, Y_K^N)(0) \right].$$

We will show that for fixed t and K , as N tends to infinity, the two sides of the above equation converges to the corresponding sides of (26). This will prove the theorem.

However, to do this we will need an intermediary stage where for an integer $J < N$, the dynamics of the process is according to Q_J while the initial distribution of the spacings is either μ or μ_N . Define, for $K < J < N$, the following quantities:

$$\begin{aligned} a &= \mathbb{E}^{Q_N \cdot \mu_N} \left[F(Y_1^N, \dots, Y_K^N)(t) \right] - \mathbb{E}^{Q_J \cdot \mu_N} \left[F(Y_1^J, \dots, Y_K^J)(t) \right], \\ b &= \mathbb{E}^{Q_J \cdot \mu_N} \left[F(Y_1^J, \dots, Y_K^J)(t) \right] - \mathbb{E}^{Q_J \cdot \mu} \left[F(Y_1^J, \dots, Y_K^J)(t) \right], \\ c &= \mathbb{E}^{Q_J \cdot \mu} \left[F(Y_1^J, \dots, Y_K^J)(t) \right] - \mathbb{E}^{Q \cdot \mu} \left[F(\Delta_1, \dots, \Delta_K)(t) \right], \text{ and} \\ d &= \mathbb{E}^{\mu} \left[F(\Delta_1, \dots, \Delta_K)(0) \right] - \mathbb{E}^{\mu_N} \left[F(Y_1^N, \dots, Y_K^N)(0) \right]. \end{aligned}$$

It is clear that a, b, c , and d , all depend on t, K, J , and N , although we choose to suppress this dependence in the notation. Also, it follows plainly from their definitions combined with equality (47) that

$$\left| \mathbb{E}^{Q \cdot \mu} \left[F(\Delta_1, \dots, \Delta_K)(t) \right] - \mathbb{E}^{\mu} \left[F(\Delta_1, \dots, \Delta_K)(0) \right] \right| \leq |a| + |b| + |c| + |d|.$$

We will now show that a, b, c, d all go to zero if we select a sequence of J and N such that J, N and N/J^2 go to infinity. This will prove the theorem.

Step 1.[Estimate of a .] For $x_1 \leq x_2 \leq \dots \leq x_{K+1}$, define

$$G(x_1, x_2, \dots, x_{K+1}) := F(x_2 - x_1, x_3 - x_2, \dots, x_{K+1} - x_K).$$

Then, clearly, G is also a continuous bounded function. Assume that $\sup_x |G(x)| \leq \alpha$. Now, by the changing the measures from Q_N and Q_J to P , we get

$$(48) \quad \begin{aligned} a &= \mathbb{E}^{Q_N \cdot \mu_N} \left[G(Z_1^N, \dots, Z_{K+1}^N)(t) \right] - \mathbb{E}^{Q_J \cdot \mu_N} \left[G(Z_1^J, \dots, Z_{K+1}^J)(t) \right] \\ &= \mathbb{E}^{P \cdot \mu_N} \left[D_N(t) G(Z_1^N, \dots, Z_{K+1}^N)(t) - D_J(t) G(Z_1^J, \dots, Z_{K+1}^J)(t) \right], \end{aligned}$$

where D_N and D_J are the Radon-Nikodým derivative processes defined in (25). Define the event

$$\Gamma := \left\{ \left(Z_1^N, \dots, Z_{K+1}^N \right) (s) = \left(Z_1^J, \dots, Z_{K+1}^J \right) (s), \forall s \in [0, t] \right\}.$$

If $\omega \in \Gamma$, by definition, $P \cdot \mu_N$ almost surely $D_N(t, \omega) = D_J(t, \omega)$. Thus (48) can be written as

$$a = \mathbb{E}^{P \cdot \mu_N} \left[\left(D_N(t)G(Z_1^N, \dots, Z_{K+1}^N) - D_J(t)G(Z_1^J, \dots, Z_{K+1}^J) \right) 1_{\Gamma^c} \right]$$

Thus, $|a|$ is bounded above by

$$\begin{aligned} & \mathbb{E}^{P \cdot \mu_N} | D_N(t)G(\cdot)1_{\Gamma^c} | + \mathbb{E}^{P \cdot \mu_N} | D_J(t)G(\cdot)1_{\Gamma^c} | \\ (49) \quad & \leq \alpha \left(\mathbb{E}^{P \cdot \mu_N} | D_N(t)1_{\Gamma^c} | + \mathbb{E}^{P \cdot \mu_N} | D_J(t)1_{\Gamma^c} | \right) \\ & \leq \alpha (\|D_N(t)\|_N + \|D_J(t)\|_N) \sqrt{P \cdot \mu_N(\Gamma^c)}. \end{aligned}$$

The final inequality is due to the Cauchy-Schwartz inequality, where the norm $\|\cdot\|_N$ refers to the \mathbf{L}^2 norm under the measure $P \cdot \mu_N$.

Now, under $P \cdot \mu_N$, both the Radon-Nikodym derivatives $D_N(t)$ and $D_J(t)$ are equal in law to $\exp(B_t - t/2)$, where B is a standard Brownian motion. Thus, it is straightforward to see

$$\|D_N(t)\|_N = \|D_J(t)\|_N = \exp(t/2).$$

Also, by Lemma 19 (put $k = K + 1$ in the statement), for large enough J and N we have $P \cdot \mu_N(\Gamma^c)$ is less than

$$C_2(K+1)e^{2t}(J-1)^K \left(\sqrt{\frac{4et}{J-K+1}} \right)^{J-K} \left[1 - \left(\frac{4et}{J-K+1} \right)^{N-J} \right].$$

If we plug everything back to (49), we see that $|a|$ goes to zero as J , N , and N/J^2 go to infinity, while keeping t and K fixed.

Step 2.[Estimate of b .] Under Q_J^x , the vector (Y_1^J, \dots, Y_K^J) depends only on the first J processes X_1, X_2, \dots, X_J , and independent of X_i , $i > J$. Thus

$$\mathbb{E}^{Q_J^x} \left[F(Y_1^J, \dots, Y_K^J) \right] = H(x_1, x_2, \dots, x_J),$$

where H is bounded function since F is bounded. Thus, we have the following equality

$$|b| = \left| \mathbb{E}^{\mu_N}(H(X_1(0), \dots, X_J(0))) - \mathbb{E}^\mu(H(X_1(0), \dots, X_J(0))) \right|$$

If ν_N and ν denote the law of the vector $(X_1(0), \dots, X_J(0))$ under μ_N and μ respectively, then, since we have assume G (or equivalently F) to be bounded in absolute value by α , we have $|b| \leq \alpha \|\nu_N - \nu\|_{TV}$. Here, $\|\cdot\|_{TV}$ refers to the total variation norm. We will now show that $\|\nu_N - \nu\|_{TV}$ goes to zero as N , J , and N/J^2 go to infinity.

Under μ_N , $X_1(0) = 0$ and each initial spacing $X_{i+1}(0) - X_i(0)$, $1 \leq i \leq N-1$, follows independent $\text{Exp}(2(1-i/N))$. While, under μ , $X_1(0) = 0$ and the spacings follow iid $\text{Exp}(2)$. Now the law of the vector $(X_i(0), 1 \leq i \leq J)$ is determined by the first J spacings $(X_{i+1}(0) - X_i(0), 1 \leq i \leq J)$ which gives us the following inequality

$$\begin{aligned} \|\nu_N - \nu\|_{TV} &\leq \int_{\mathbb{R}^J} \left| \prod_{i=1}^J 2(1-i/N)e^{-2(1-i/N)x_i} - 2^J e^{-2\sum_{i=1}^J x_i} \right| dx \\ &= \int_{\mathbb{R}^J} \left| \prod_{i=1}^J (1-i/N)e^{2ix_i/N} - 1 \right| 2^J e^{-2\sum_{i=1}^J x_i} dx. \end{aligned}$$

By an application of Cauchy-Schwartz inequality, we get

$$\begin{aligned} \|\nu_N - \nu\|_{TV}^2 &\leq \int_{\mathbb{R}^J} \left| \prod_{i=1}^J (1-i/N)e^{2ix_i/N} - 1 \right|^2 2^J e^{-2\sum_{i=1}^J x_i} dx, \\ (50) \quad &= 2^J \prod_{i=1}^J (1-i/N)^2 \int_{\mathbb{R}^J} \exp \left\{ -2 \sum_{i=1}^J (1-2i/N)x_i \right\} dx \\ &\quad - 2^{J+1} \prod_{i=1}^J (1-i/N) \int_{\mathbb{R}^J} \exp \left\{ -2 \sum_{i=1}^J (1-i/N)x_i \right\} dx + 1 \end{aligned}$$

By the standard identity $\int e^{-\lambda x} dx = \lambda^{-1}$ for $\lambda > 0$, we get

$$\begin{aligned} \|\nu_N - \nu\|_{TV}^2 &\leq 2^J \prod_{i=1}^J (1-i/N)^2 \prod_{i=1}^J (2(1-2i/N))^{-1} \\ (51) \quad &\quad - 2^{J+1} \prod_{i=1}^J (1-i/N) \prod_{i=1}^J (2(1-i/N))^{-1} + 1 \\ &= \frac{(1-1/N)^2(1-2/N)^2 \dots (1-J/N)^2}{(1-2/N)(1-4/N) \dots (1-2J/N)} - 1. \end{aligned}$$

The following inequality is straightforward to prove

$$e^{-2x} \leq 1 - x \leq e^{-x}, \quad \text{for all } 0 \leq x \leq \frac{1}{2} \log 2.$$

By our assumption J , N , and N/J^2 are going to infinity, we can assume for all sufficiently large values of J and N that $2J/N \leq \log 2/2$. By the previous inequality, we get

$$e^{-2i/N} \leq (1 - i/N) \leq e^{-i/N}, \quad 1 \leq i \leq 2J/N,$$

and consequently

$$(52) \quad \frac{(1 - 1/N)^2(1 - 2/N)^2 \dots (1 - J/N)^2}{(1 - 2/N)(1 - 4/N) \dots (1 - 2J/N)} \leq \frac{\exp\left(-2 \sum_{i=1}^J i/N\right)}{\exp\left(-2 \sum_{i=1}^J 2i/N\right)}$$

and

$$\frac{\exp\left(-4 \sum_{i=1}^J i/N\right)}{\exp\left(-\sum_{i=1}^J 2i/N\right)} \leq \frac{(1 - 1/N)^2(1 - 2/N)^2 \dots (1 - J/N)^2}{(1 - 2/N)(1 - 4/N) \dots (1 - 2J/N)}.$$

Thus, we can give upper and lower bounds

$$e^{-J(J+1)/N} \leq \frac{(1 - 1/N)^2(1 - 2/N)^2 \dots (1 - J/N)^2}{(1 - 2/N)(1 - 4/N) \dots (1 - 2J/N)} \leq e^{J(J+1)/N}.$$

Combining this with (51) we see that as N , J , and N/J^2 go to infinity, we clearly get that $|b|$ converges to zero.

Step 3.[Estimate of c .] This is similar to the methods we use to estimate a . As in Step 1, we first employ a change of measure to obtain

$$c = \mathbb{E}^{P \cdot \mu}[D_J(t)G(Z_1^J, \dots, Z_{K+1}^J)(t)] - \mathbb{E}^{P \cdot \mu}[D(t)G(X_{(1)}, \dots, X_{(K+1)})(t)].$$

Consider the event

$$(53) \quad \Gamma(t, J, K) = \left\{ \left(X_{(1)}, \dots, X_{(K+1)} \right) (s) = \left(Z_1^J, \dots, Z_{K+1}^J \right) (s), \quad 0 \leq s \leq t \right\}$$

On $\Gamma(t, J, K)$, the random processes $G(Z_1^J, \dots, Z_{K+1}^J)(s)$ are identical to $G(X_{(1)}, \dots, X_{(K+1)})(s)$ in time $s \in [0, t]$. Also, the processes Z_1^J and $X_{(1)}$ are the same in the time interval $[0, t]$. Thus the process $D(t)$ and $D_J(t)$ are also the same. Since, by our assumption, $|G|$ is bounded by α , the following upper bound on $|c|$ holds

$$(54) \quad |c| \leq \alpha \mathbb{E}^{P \cdot \mu} \left[(|D(t)| + |D_J(t)|) \mathbf{1}_{\Gamma^c(t, J, K)} \right].$$

$$\leq \alpha (\|D_t\| + \|D_J(t)\|) \sqrt{P \cdot \mu(\Gamma^c(t, J, K))},$$

where, we denote by $\|\cdot\|$ the \mathbf{L}^2 norm under the measure $P \cdot \mu$.

Now, by Lemma 18 (for $k = K + 1$ and $N = J$ in the statement), for large enough J we get

$$(55) \quad P \cdot \mu(\Gamma^c(t, J, K)) \leq C_2(K + 1)e^{2t}(J - 1)^K \left(\sqrt{\frac{4et}{J - K + 1}} \right)^{J-K}.$$

Now, since t and K are fixed, as J tends to infinity, $P \cdot \mu(\Gamma^c(t, J, K))$ goes to zero. Finally, as in the estimate of a in Step 1, note that, by (22) and (25), we can assert that

$$D(t) = \exp(X_t - t/2), \quad \text{and} \quad D_N(t) = \exp(Y_t - t/2),$$

where X and Y are Brownian motions. Thus $\|D(t)\| = e^{t/2} = \|D_N(t)\|$. If we plus back everything in (54), we get $|c|$ goes to zero as J tends to infinity.

Step 4.[Estimate of d .] At time zero the indices are arranged in the increasing order, and hence $(\Delta_1, \Delta_2, \dots, \Delta_K)$ is obviously equal to $(Y_1^N, Y_2^N, \dots, Y_K^N)$. Since $|F| \leq \alpha$, it follows that $|d|$ is bounded by α times the total variation distance between the law of $(\Delta_1, \Delta_2, \dots, \Delta_K)$ under μ and μ_N . A computation exactly similar to (50) gives us

$$\left(\frac{|d|}{\alpha} \right)^2 \leq \frac{(1 - 1/N)^2(1 - 2/N)^2 \dots (1 - K/N)^2}{(1 - 2/N)(1 - 4/N) \dots (1 - 2K/N)} - 1.$$

But, since K is fixed, and N grows to infinity, the right-hand-side above goes to zero by a logic similar to (52). This proves the estimate.

Thus, combining Steps 1, 2, 3 and 4, we have proved the theorem. ■

Acknowledgements. We thank Prof. M. E. Vares for drawing our attention to articles [5] and [24]. We are also grateful to the anonymous referee for an extremely helpful review which significantly improved the quality of the paper, including pointing out references [6] and [7].

REFERENCES

- [1] ARRATIA, R. (1983). The motion of a tagged particle in the simple symmetric exclusion system on \mathbf{Z} . *Ann. Probab.*, **11**(2) 362–373.
- [2] ARRATIA, R. (1985). Symmetric exclusion processes: a comparison inequality and a large deviation result. *Ann. Probab.*, **13**(1) 53–61.
- [3] BANNER, A. and FERNHOLZ, R. and KARATZAS, I. (2005). Atlas models of equity markets. *Ann. Appl. Probab.*, **15**(4) 2296–2330.

- [4] BARYSHNIKOV, Y. (2001). GUEs and queues. *Probab. Theory Related Fields*, **119**(2) 256–274.
- [5] DE MASI, A. and FERRARI, P. (2002). Flux fluctuations in the one dimensional nearest neighbors symmetric simple exclusion process. *J. Stat. Phys.*, **107**(3/4) 677–683.
- [6] DÜRR, D. and GOLDSTEIN, S. and LEBOWITZ, J. (1985). Asymptotics of particle trajectories in infinite one-dimensional systems with collisions. *Comm. Pure Appl. Math.*, **38**(5) 573–597.
- [7] DÜRR, D. and GOLDSTEIN, S. and LEBOWITZ, J. (1987). Self-diffusion in a non-uniform one dimensional system of point particles with collisions *Probab. Theory Related Fields*, **75**(2) 279–290.
- [8] DURRETT, R. (1996). *Stochastic Calculus: a practical introduction*. CRC Press, Inc.
- [9] FERNHOLZ, R. (2002). *Stochastic Portfolio Theory*. Springer, New York.
- [10] FERNHOLZ, R. and KARATZAS, I. (2007). A survey of stochastic portfolio theory. To appear.
- [11] FERRARI, P. A. (1996). Limit theorems for tagged particles. *Markov Process. Related Fields*, **2**(1) 17–40. Disordered systems and statistical physics: rigorous results (Budapest, 1995).
- [12] FERRARI, P. A. and FONTES, L. R. G. (1994). The net output process of a system with infinitely many queues. *Ann. Appl. Probab.*, **4**(4) 1129–1144.
- [13] FERRARI, P. A. and FONTES, L. R. G. (1996). Poissonian approximation for the tagged particle in asymmetric simple exclusion. *J. Appl. Probab.*, **33**(2) 411–419.
- [14] HARRIS, T. E. (1965). Diffusion with “collisions” between particles. *J. Appl. Probability*, **2** 323–338.
- [15] HARRISON, J. M. (1973). The heavy traffic approximation for single server queues in series. *J. Appl. Probability*, **10** 613–629.
- [16] HARRISON, J. M. (2000). Brownian models of open processing networks: canonical representation of workload. *Ann. Appl. Probab.*, **10**(1):75–103.
- [17] HARRISON, J. M. and VAN MIEGHEM, J. A. (1997). Dynamic control of Brownian networks: state space collapse and equivalent workload formulations. *Ann. Appl. Probab.*, **7**(3) 747–771.
- [18] HARRISON, J. M. and WILLIAMS, R. J. (1987). Brownian models of open queueing networks with homogeneous customer populations. *Stochastics*, **22**(2) 77–115.
- [19] HARRISON, J. M. (1978). The diffusion approximation for tandem queues in heavy traffic. *Adv. in Appl. Probab.*, **10**(4) 886–905.
- [20] KIPNIS, C. (1986). Central limit theorems for infinite series of queues and applications to simple exclusion. *Ann. Probab.*, **14**(2) 397–408.
- [21] MCKEAN, H. P. and SHEPP, L. (2005). The advantage of capitalism vs. Socialism depends on the criterion. Available at www.emis.de/journals/ZPOMI/v328/p160.ps.gz.
- [22] O’CONNELL, N. and YOR, M. (2001). Brownian analogues of Burke’s theorem. *Stochastic Process. Appl.*, **96**(2) 285–304.
- [23] REVUZ, D. and YOR, M. (1999). *Continuous martingales and Brownian motion*, 3rd ed. Springer, Berlin-Heidelberg.
- [24] ROST, H. and VARES, M. E. (1985). Hydrodynamics of a one-dimensional nearest neighbor model. *Contemporary Mathematics*, **41** 329–342.
- [25] SEPPÄLÄINEN, T. (1997). A scaling limit for queues in series. *Ann. Appl. Probab.*, **7**(4):855–872.

- [26] (1993). SRINIVASAN, R. Queues in series via interacting particle systems. *Math. Oper. Res.*, **18**(1) 39–50.
- [27] SZNITMAN, A. S. (1986). A propagation of chaos result for Burgers' equation. *Probab. Theory Relat. Fields*, **71**(4) 581–613.
- [28] SZNITMAN, A. S. (1991). Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, 165–251. Springer, Berlin.
- [29] VARADHAN, S. R. S. and WILLIAMS, R. J. (1985) Brownian motion in a wedge with oblique reflection. *Comm. Pure Appl. Math.*, **38** 405–443.
- [30] WILLIAMS, R. J. (1987) Reflected Brownian motion with skew symmetric data in a polyhedral domain. *Probab. Theory Related Fields*, **75**(4) 459–485.

506 MALOTT HALL
CORNELL UNIVERSITY
ITHACA NY 14853
E-MAIL: soumik@math.cornell.edu

367 EVANS HALL # 3860
UNIVERSITY OF CALIFORNIA
BERKELEY CA 94720
E-MAIL: pitman@stat.berkeley.edu