The Brownian fan

April 9, 2014

Martin Hairer[1] and Jonathan Weare[2]

[1] Mathematics Department, the University of Warwick
[2] Statistics Department and the James Frank Institute, the University of Chicago
Email: M.Hairer@Warwick.ac.uk, Email: weare@uchicago.edu

Abstract

We provide a mathematical study of the modified Diffusion Monte Carlo (DMC) algorithm introduced in the companion article [HW14]. DMC is a simulation technique that uses branching particle systems to represent expectations associated with Feynman-Kac formulae. We provide a detailed heuristic explanation of why, in cases in which a stochastic integral appears in the Feynman-Kac formula (e.g. in rare event simulation, continuous time filtering, and other settings), the new algorithm is expected to converge in a suitable sense to a limiting process as the time interval between branching steps goes to 0. The situation studied here stands in stark contrast to the "naïve" generalisation of the DMC algorithm which would lead to an exponential explosion of the number of particles, thus precluding the existence of any finite limiting object. Convergence is shown rigorously in the simplest possible situation of a random walk, biased by a linear potential. The resulting limiting object, which we call the "Brownian fan", is a very natural new mathematical object of independent interest.

Keywords: Diffusion Monte Carlo, quantum Monte Carlo, rare event simulation, sequential Monte Carlo, particle filtering, Brownian fan, branching process

Contents

1   Introduction

Consider a Markov chain y_{t_k} with transition probabilities $\mathbf{P}(x, dy)$ on some state space $\mathcal{X}$ and, in anticipation of developments below, indexed by a sequence of real numbers $t_0 < t_1 < t_2 < \cdots$. The nature of the state space does not matter (Polish is enough), but one should think of $\mathcal{X} = \mathbf{R}^d$ for definiteness. Given functions $\chi \colon \mathcal{X}^2 \to \mathbf{R}$ and $f \colon \mathcal{X} \to \mathbf{R}$ with sufficient integrability properties (say bounded), the Diffusion Monte Carlo (DMC) algorithm (see [HW14] for description of DMC compatible with the discussion below) computes an estimate of expectations of the form

$$\langle f \rangle_t = \mathbf{E}\left(f(y_t) \exp\left(- \sum_{t_k \leq t} \chi(y_{t_k}, y_{t_{k+1}}) \right) \right) , \tag{1.1}$$

where y a realisation of the Markov chain described by $\mathbf{P}$. Strictly for convenience we have assume that t is among the times $t_0, t_1, t_2, \ldots$. More precisely, at time t, DMC produces a collection of N_t copies of the underlying system, $\{x_t^{(i)}\}_{i=1}^{N_t}$ so that

$$\mathbf{E} \sum_{i=1}^{N_t} f(x_t^{(i)}) = \langle f \rangle_t. \tag{1.2}$$

Of course, the expectation in (1.1) can be computed by generating many independent sample trajectories of y. However, in most cases the weights $\exp(- \sum_{t_k \leq t} v(y_{t_k}, y_{t_{k+1}}))$ will quickly degenerate so that a huge number of samples is required to generate a single statistically significant sample. At each time step DMC removes samples with very low weights and replaces them with multiple copies of samples with larger weights, focusing computational effort on statistically significant samples.

    Variants of DMC are used regularly in a wide range of fields including electronic structure, rare event simulation, and data assimilation (see [HM54, dDG05, AFRtW06, JDMD06, HW14] for a small sample of these applications). The algorithm has also been the subject of significant mathematical inquiry [DM04]. Continuous time limits have been considered [Rou06] for cases in which χ scales like the time interval between branching steps. Perhaps because standard DMC does not have a limit in those cases, more general Feynman-Kac formulae do not seem to have been considered despite their appearance in applications.

    For the purposes of this article, the main example one should keep in mind is that where y is a time discretisation of a diffusion process obtained for example by applying an Euler scheme with fixed stepsize ε:

$$y_{t_{k+1}} = y_{t_k} + \varepsilon F(y_{t_k}) + \sqrt{\varepsilon} \Sigma(y_{t_k}) \xi_{k+1} , \tag{1.3}$$

where $\varepsilon = t_{k+1} - t_k$ and the ξ_k are a sequence of i.i.d. random variables (not necessarily Gaussian). Regarding the function χ, we will mostly consider the case $\chi(y, \tilde{y}) = V(\tilde{y}) - V(y)$. As described in [HW14], this choice arrises in application of DMC to rare event simulation problems. Notice that, for this choice of χ and for y in (1.3), when ε is small, $\chi(y_{t_k}, y_{t_{k+1}}) \sim \sqrt{\varepsilon}$. This causes a dramatic failure in standard DMC [HW14].

Yet, for small ε (with fixed t) the expectation $\langle f \rangle_t$ defined in (1.1) has the perfectly nice limit

$$\mathbf{E}\left(f(y_t) \exp\left(-V(y_t) + V(y_0) \right) \right)$$

where now y_t is the continuous time limit of (1.3) (the diffusion with drift F and diffusion Σ). When a stochastic integral appears in the exponential the situation is completely analogous. It is natural then to search for a modification of DMC that can handle these settings.

It was shown in [HW14] that the following algorithm provides an unbiased estimator for $\langle f \rangle_t$ (it satisfies (1.2)) and is superior to DMC (it has lower variance and equivalent expected cost).

Algorithm TDMC *Ticketed DMC*

1. Begin with M copies $x_0^{(j)} = x_0$. For each $j = 1, \ldots, M$ choose an independent random variable $\theta_0^{(j)} \sim \mathcal{U}(0,1)$.

2. At step k there are N_{t_k} samples $(x_{t_k}^{(j)}, \theta_{t_k}^{(j)})$. Evolve each of the $x_{t_k}^{(j)}$ one step to generate N_{t_k} values

$$\tilde{x}_{t_k}^{(j)} \sim \mathbf{P}(y_{t_{k+1}} \in dx \,|\, y_{t_k} = x_{t_k}^{(j)}).$$

3. For each $j = 1, \ldots, N_{t_k}$, let

$$P^{(j)} = e^{-\chi(x_{t_k}^{(j)}, \tilde{x}_{t_{k+1}}^{(j)})}. \tag{1.4}$$

   If $P^{(j)} < \theta_{t_k}^{(j)}$ then set
   $$N^{(j)} = 0.$$

   If $P^{(j)} \geq \theta_{t_k}^{(j)}$ then set
   $$N^{(j)} = \max\{\lfloor P^{(j)} + u^{(j)} \rfloor, 1\}, \tag{1.5}$$

   where $u^{(j)}$ are independent $\mathcal{U}(0,1)$ random variables.

4. For $j = 1, \ldots, N_{t_k}$, if $N^{(j)} > 0$ set

$$x_{t_{k+1}}^{(j,1)} = \tilde{x}_{t_{k+1}}^{(j)} \quad \text{and} \quad \theta_{t_{k+1}}^{(j,1)} = \frac{\theta_{t_k}^{(j)}}{P^{(j)}}$$

   and for $i = 2, \ldots, N^{(j)}$

$$x_{t_{k+1}}^{(j,i)} = \tilde{x}_{t_{k+1}}^{(j)} \quad \text{and} \quad \theta_{t_{k+1}}^{(j,i)} \sim \mathcal{U}((P^{(j)})^{-1}, 1).$$

5. Finally set $N_{t_{k+1}} = \sum_{j=1}^{N_{t_k}} N^{(j)}$ and list the $N_{t_{k+1}}$ vectors $\{x_{t_{k+1}}^{(j,i)}\}$ as $\{x_{t_{k+1}}^{(j)}\}_{j=1}^{N_{t_{k+1}}}$.

6. At time t produce the estimate

$$\widehat{f}_t = \frac{1}{M} \sum_{j=1}^{N_t} f(x_t^{(j)}).$$

The aim of this article is to argue that one expects the application of Algorithm TDMC to (1.3) to converge to a finite continuous-time limiting particle process as $\varepsilon \to 0$. Section 2.1 provides very detailed heuristic as to why we expect this to be the case and gives a precise mathematical definition for the expected limiting object.

In the special case where $d = 1$, $F = 0$, $\Sigma = 1$, and $V(x) = x$, the process in question is built recursively by successive realisations of a Poisson point process in a space of excursions of y_t. A precise definition is given in Definition 2.9 and we call this object the *Brownian fan*. It is of particular interest since, similarly to the fact that every diffusion process looks locally like a Brownian motion, one would expect the general limiting objects described in Section 2.1 to locally "look like" the Brownian fan. The Brownian fan is interesting in its own right from a mathematical perspective and does not seem to have been studied before, though it is very closely related to the "Virgin island model" introduced in [Hut09].

Loosely speaking, the Brownian fan is a branching process obtained in the following way. Start with a (usually finite) number of initial particles on $\mathbf{R}$ that furthermore come each with a tag $v \in \mathbf{R}$, such that the position x satisfies $x > v$. Each of these particles, which we call the ancestor particles, perform independent Brownian motions until they hit the barrier $x = v$, where they get killed. Each of these particles independently produces offspring according to the following mechanism. Denoting by $\mathcal{E}$ the space of excursions on $\mathbf{R}$, let $\mathcal{M}$ be the Poisson point process on $\mathbf{R} \times \mathcal{E}$ with intensity measure $a\,dt \otimes \mathbf{Q}$ for some $a > 0$, where $\mathbf{Q}$ is Itô's Brownian excursion measure [Itô72]. If (t, w) is one of the points of $\mathcal{M}$ and the corresponding ancestor particle is alive at time t and located at x_t, then it gives rise to an offspring which performs the motion described by w, translated so that the origin of the excursion lies at (t, x_t). This mechanism is then repeated recursively for each new particle created in this way.

The Brownian fan has a number of very nice properties. For example, as established by Theorem 2.14, the continuous-time analogue of the number of particles at time t, N_t, corresponding to the Brownian fan satisfies

$$\sup_{t > 0} \mathbf{E} \exp\left(\lambda_t N_t\right) < \infty \tag{1.6}$$

for some continuous decreasing function $\lambda_t > 0$. As established in Proposition 2.16, the bound in (1.6) implies that the continuous-time analogue of the workload,

$$\mathcal{W}_t = \int_0^t N_s\,ds$$

is very nearly a differentiable function of time,

$$\sup_{t \leq T} \lim_{h \to 0} \frac{|\mathcal{W}_{t+h} - \mathcal{W}_t|}{h|\log h|} < \infty.$$

This is close to optimal, since there turns out to be a dense set of exceptional times at which there are infinitely many particles alive (see the argument given at the end of Section 2.2.3 below).

To conclude this round-up of its mathematical properties, we establish in Proposition 3.3 that the Brownian fan is a Feller process in a suitable state space. When combined with the continuity of its sample paths established in Proposition 4.8, this implies that the Brownian fan is a strong Markov process. Note that the construction of such a state space is a non-trivial endeavour due to the fact that, while the number N_t of particles alive at time t has finite moments of all orders, there exists a *dense* set

of exceptional times for which $N_t = \infty$! Offspring are continuously being created at infinite rate, thus making the Brownian fan quite different from a standard branching diffusion.

The "meat" of this article, Sections 4 and 5, is then devoted to a rigorous proof of the convergence of the output of Algorithm TDMC to the Brownian fan for any sufficiently light-tailed one-step distribution for the random variables ξ_k. This result is formulated in Theorem 5.1. The usual method by which one attempts to characterise the continuous-time limit of a sequence of discrete-time processes involves studying the limit of the generators of the discrete-time processes. In our case, as described in Section 3.4, the discrete-time generators do *not* converge to the correct limit, at least when applied to a class of very natural-looking test functions. Surprisingly (and rather confusingly), they do actually converge to the generator of a Brownian fan, but unfortunately with the wrong parameters! En route to our convergence result we establish a number of important and extremely encouraging results about the behaviour of the process generated by Algorithm TDMC for finite ε. For example, in Proposition 4.2 we obtain that

$$\sup_{\varepsilon < \varepsilon_0} \mathbf{E}|N_t|^p < \infty \,,$$

for any $p \geq 0$ and $\varepsilon_0 > 0$ sufficiently small where here N_t refers to the number of particles generated in Algorithm TDMC and not to the Brownian fan (for which we have the even stronger result in (1.6)). In Corollary 4.10 we also prove a uniform (in ε) form of continuity of the processes.

1.1 Notations

For any Polish space $\mathcal{Y}$, we will denote by $\mathscr{M}_+(\mathcal{Y})$ the space of finite positive measures on $\mathcal{Y}$, endowed with the topology of weak convergence, together with the convergence of total mass. Given a random variable X, we denote its law by $\mathcal{D}(X)$, except in some cases where we introduce dedicated notations.

We will often use the notation $a \lesssim b$ as a shorthand for the inequality $a \leq Cb$ for some constant C. The dependence of C on other quantities will usually be clear from the context, and will be indicated when ambiguities may arise. We will also use the standard notations $a \wedge b = \min\{a,b\}$, $a \vee b = \max\{a,b\}$, and $\lfloor a \rfloor = \max\{i \in \mathbb{Z} : i \leq a\}$.

Acknowledgements

2 The continuous-time limit and the Brownian fan

From now on, we restrict ourselves to the analysis of the case $\chi(x,y) = V(y) - V(x)$ for some "potential" V defined on the state space of the underlying Markov process. We argue that if the underlying process is obtained by approximating a diffusion process then, unlike in the case of the naïve generalisation of DMC (see [HW14, Algorithm DMC]), our modification, Algorithm TDMC, converges to a non-trivial limiting process as the stepsize ε converges to 0.

We will first provide a heuristic argument showing what kind of limiting process one would expect to obtain. The remainder of this article will then be devoted to rigorously constructing the limiting process and proving convergence in the simple case in which the underlying Markov chain is a random walk (rescaled so that it converges to a standard Brownian motion) and the biasing potential V is linear. In this case the limiting process is a very natural object that does not seem to have been studied in the literature so far. We call this object, which is closely related to the construction in [Hut09], the *Brownian fan* (see Section 2.2). It also has a flavour very similar to the construction of the Brownian web [FINR04] and the Brownian net [SS08], although there does not seem to be an obvious transformation linking these objects.

2.1 Heuristic derivation of the continuous-time limit

Throughout this section the underlying Markov chain will be given just like in the introduction by the following approximation to a diffusion:

$$y_{(k+1)\varepsilon} = y_{k\varepsilon} + \varepsilon F(y_{k\varepsilon}) + \sqrt{\varepsilon}\Sigma(y_{k\varepsilon})\,\xi_{k+1}\,, \qquad y_{k\varepsilon} \in \mathbf{R}^n\,, \tag{2.1}$$

where the ξ_k are a sequence of i.i.d. (not necessarily Gaussian) random variables with law ν and the identity on $\mathbf{R}^n$ as their covariance matrix. The functions F and Σ are sufficiently "nice" functions, but since this section is only heuristic, we do not state specific regularity, growth or non-degeneracy assumptions.

Remark 2.1 We have slightly changed our notations by writing $y_{k\varepsilon}$ instead of y_k for the position of the Markov chain after k steps. This is in order to make explicit the fact that as $\varepsilon \to 0$, one has convergence to a continuous-time process. The corresponding notational changes in Algorithm TDMC are straightforward.

Concerning the function χ, we take $\chi(x,y) = V(y) - V(x)$ for some regular potential $V\colon \mathbf{R}^n \to \mathbf{R}$. Recall now that, as long as a particle is alive, its ticket θ evolves under Algorithm TDMC as

$$\theta_{k\varepsilon}^{(j)} = \theta_{k\varepsilon}^{(j)} \exp(V(\tilde{x}_{(k+1)\varepsilon}^{(j)}) - V(x_{k\varepsilon}^{(j)}))\,.$$

It is natural therefore to replace θ by the quantity v given by

$$\exp(-v_t^{(j)}) = \theta_t^{(j)} \exp(-V(x_t^{(j)}))\,.$$

In this way, the new "tag" v does not change over time, but is assigned to a particle at the moment of its birth. Translating Steps 3 and 4 of the algorithm into this slightly different setting, we see that if a particle performs a step from x to y such that $V(y) < V(x)$, then it can potentially spawn one or more descendants. The tags v of the descendants are then distributed according to

$$e^{-v} \stackrel{\text{law}}{=} \mathcal{U}(e^{-V(x)}, e^{-V(y)})\,, \tag{2.2}$$

and a particle with tag v lives as long as it stays within the region $\{x \,:\, V(x) \le v\}$.

2.1.1 Description of the limit

For very small values of ε, the process described above has the following features. Taking the limit $\varepsilon \to 0$ in (2.1), we observe that each particle follows a diffusion process, solving the equation

$$dy_t = F(y_t)\,dt + \Sigma(y_t)\,dB_t\,, \tag{2.3}$$

where B_t is a standard d-dimensional Brownian motion. If the particle has tag v, then this process is killed as soon as it exits the sublevel set $\{x \,:\, V(x) \le v\}$.

Consider the following representation of the object produced by Algorithm TDMC. Denote by $\mathbf{Q}^\varepsilon_{x,v}$ the law of the ε-discretization of (2.3) generated by (2.1) starting at x and killed upon exiting the set $\{y \,:\, V(y) \le v\}$. Let τ and $\{w_{k\varepsilon}\}_{k=0}^{\tau/\varepsilon}$ be, respectively, the lifetime and trajectory of the original particle. The trajectories of the offspring of this initial trajectory are very nearly given (it will be true in the small ε limit) by a realisation $\mu^{\varepsilon,1}$ of a Poisson point process with intensity

$$\mathcal{Q}^\varepsilon(w, \cdot) = \sum_{k=0}^{\tau/\varepsilon-1} A^\varepsilon(k)\, \Theta^\star_{k\varepsilon} \int \sqrt{\varepsilon} \mathbf{Q}^\varepsilon_{w_{k\varepsilon}, V(w_{k\varepsilon})+\delta}\, \eta(k, d\delta)\,, \qquad (2.4)$$

where, according to the rule for generating new offspring in Algorithm TDMC,

$$A^\varepsilon(k) = \begin{cases} \frac{1}{\sqrt{\varepsilon}}\big(e^{-(V(w_{(k+1)\varepsilon})-V(w_{k\varepsilon}))} - 1\big) & \text{if } V(w_{(k+1)\varepsilon}) < V(w_{k\varepsilon}) \\ 0 & \text{if } V(w_{(k+1)\varepsilon}) \ge V(w_{k\varepsilon}) \end{cases}$$

and, according to the rule for generating offspring tickets in Algorithm TDMC,

$$\int f(\delta)\eta(k, d\delta) = \int_0^1 f\big(-\log(1 + u(e^{-(V(w_{(k+1)\varepsilon})-V(w_{k\varepsilon}))} - 1))\big)\, du.$$

Here Θ_t is the map that shifts trajectories forward by time t. Since each offspring behaves independently just like the original particle, this suggests that the nth generation $\mu^{\varepsilon,n}$ of offspring is obtained recursively as a realisation of the Poisson point process with intensity given by

$$\mathcal{G}^{\varepsilon,n}(\cdot) = \int \mathcal{Q}^\varepsilon(\tilde{w}, \cdot)\, \mu^{\varepsilon,n-1}(d\tilde{w})\,.$$

At each "microscopic" step, the probability of creating a descendent is of order $\sqrt{\varepsilon}$ so that, in the limit $\varepsilon \to 0$, each particle spawns descendants at infinite rate. However, any such descendant is created at distance $\mathcal{O}(\sqrt{\varepsilon})$ of the "barrier" $V(x) = v$. As a consequence, the probability that it survives for a time of order 1 before being killed is itself only of order $\sqrt{\varepsilon}$. Therefore, the rate at which a particle creates descendants that actually survive for a time τ of order 1 is finite, but tends to infinity as $\tau \to 0$.

Now we will consider the small ε-limit of the object we have constructed. The trajectory w_t becomes a sample path of (2.3) exiting the set $\{y \,:\, V(y) \le v\}$ at time τ. Denote by $\mathbf{Q}_{x,v}$ the law of the diffusion (2.3) starting at x and killed upon exiting the set $\{y \,:\, V(y) \le v\}$, which is a probability measure on some space of excursions in $\mathbf{R}^n$. The characterisation of the standard Itô excursion measure (see for example [RY91, Theorem 4.1] and [PY82]) then suggests that, for every $x \in \mathbf{R}^n$ such that $\nabla V \ne 0$ and Σ is non-degenerate, the limit

$$\mathbf{Q}_x = \lim_{\delta \to 0^+} \frac{1}{\delta} \mathbf{Q}_{x, V(x)+\delta}\,,$$

exists as a σ-finite measure in the sense that $\frac{1}{\delta}\mathbf{Q}_{x,V(x)+\delta}$ restricted to the set of excursions longer than a fixed length converges weakly to $\mathbf{Q}_x$ restricted to the same set.

The discussion so far suggests that for the limiting object, the trajectories of the first generation of offspring are given by a realisation μ^1 of the Poisson point process with intensity measure

$$\mathcal{Q}(w, \cdot) = \int_0^\tau A(w_t)\, \Theta^\star_t \mathbf{Q}_{w_t}\, dt\,, \qquad (2.5)$$

for some intensity $A \colon \mathbf{R}^n \to \mathbf{R}_+$ yet to be determined and that the nth generation μ^n of offspring is obtained recursively as a realisation of a Poisson point process with intensity given by

$$\mathcal{G}^n(\cdot) = \int \mathcal{Q}(\tilde{w}, \cdot)\,\mu^{n-1}(d\tilde{w})\,,$$

with $\mathcal{Q}$ as in (2.5). In order to fully characterise the limiting object, it remains to provide an expression for the intensity function A.

Let us start by replacing $\mathbf{Q}_{x,v}^\varepsilon$ in equation (2.4) by $\mathbf{Q}_{x,v}$, i.e. by assuming that for small ε, excursions of the discrete process are very similar to excursions of its continuous time limit. We then apply the relations

$$\mathbf{Q}_{x,V(x)+\delta} \approx \delta \mathbf{Q}_x \tag{2.6}$$

and

$$V(w_{(k+1)\varepsilon}) - V(w_{k\varepsilon}) \approx \sqrt{\varepsilon}\langle\nabla V(w_{k\varepsilon}), \Sigma(w_{k\varepsilon})\xi_{k+1}\rangle$$

with the ξ_{k+1} as in (2.1). We then formally obtain

$$\mathcal{Q}^\varepsilon(w, \cdot) \approx \varepsilon \sum_{k=0}^{\tau/\varepsilon-1} \mathbf{1}_{\langle\nabla V(x), \Sigma(x)\xi_{k+1}\rangle<0}\langle\nabla V(w_{k\varepsilon}), \Sigma(w_{k\varepsilon})\xi_{k+1}\rangle$$

$$\times \int_0^1 u\langle\nabla V(w_{k\varepsilon}), \Sigma(w_{k\varepsilon})\xi_{k+1}\rangle du\,\Theta_{k\varepsilon}^\star\mathbf{Q}_{w_{k\varepsilon}}.$$

Our arguments so far therefore suggest that

$$A(x) = \frac{1}{2}\int_{\langle\nabla V(x), \Sigma(x)z\rangle<0}\langle\nabla V(x), \Sigma(x)z\rangle^2\,\nu(dz)\,, \tag{2.7}$$

where the distribution ν has mean 0 and identity covariance matrix. Assuming that ν is symmetric this becomes

$$A(x) = c\langle\nabla V(x), \Sigma(x)\Sigma^T(x)\nabla V(x)\rangle\,, \tag{2.8}$$

where $c = \frac{1}{4}$. If ν is not symmetric, one might even expect a prefactor c that depends on x.

In fact, as we will see in a specific case in the remainder of this section, the correct value is $c = \frac{1}{2}$, whether the law of ξ is symmetric or not. The reason for this discrepancy is that the relation

$$\mathbf{Q}_{x,V(x)+\delta}^\varepsilon \approx \delta \mathbf{Q}_x$$

used in our derivation is only valid if $\delta \gg \sqrt{\varepsilon}$. In our case however, one precisely has $\delta \sim \sqrt{\varepsilon}$, which introduces a correction factor that eventually gives rise to the value $c = \frac{1}{2}$. The aim of the next subsection is to show in more detail how this factor $\frac{1}{2}$ arises in the simplest situation where $F = 0$ and $\Sigma = 1$.

2.1.2 The case of Brownian motion

We now consider the one-dimensional case, where the limiting underlying process is simple Brownian motion. Regarding the underlying discrete problem, we consider the Markov chain defined recursively by

$$y_{(k+1)\varepsilon} = y_{k\varepsilon} + \sqrt{\varepsilon}\xi_{k+1}\,, \tag{2.9}$$

for an i.i.d. sequence of centred random variables ξ_k with law ν and variance 1. For the potential function V, we choose $V(x) = -ax$ for some $a > 0$.

In order to show that the constant c appearing in (2.8) is equal to $\frac{1}{2}$, we will now argue that if we denote by $\mathbf{Q}$ the standard Itô excursion measure (which we normalise in such a way that $\mathbf{Q} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathbf{Q}_\varepsilon$, where $\mathbf{Q}_\varepsilon$ is the law of a standard Brownian motion starting at ε and killed when it hits the origin) and by $\mathbf{Q}_z^\varepsilon$ the law of the random walk (2.9) starting at $\sqrt{\varepsilon} z$ and stopped as soon as it takes negative values, then there exists a function G such that

$$\mathbf{Q}_z^\varepsilon \approx \sqrt{\varepsilon}\, G(z)\mathbf{Q} \,, \tag{2.10}$$

as $\varepsilon \to 0$ when both sides are restricted to excursions that survive for at least some fixed amount of time. We will see that the function G behaves like $G(z) \approx z$ for large values of z, but has a non-trivial behaviour for values of order 1. In terms of our notation from the previous subsection (since y_t is spatially homogeneous and V is linear) this implies that the approximation (2.6) should really have been replaced by

$$\mathbf{Q}_{x,V(x)+\delta}^\varepsilon \approx a\sqrt{\varepsilon}\, G\!\left(\frac{\delta}{a\sqrt{\varepsilon}}\right)\mathbf{Q} \,.$$

Since we assumed $a > 0$, our process creates offspring only when it performs a step towards the right, i.e. when $\xi_{k+1} > 0$. The probability that a new particle is created in the k-th step is approximately $a\xi_{k+1}$. Furthermore, the small ε rule for the generation of tags implied by Algorithm TDMC is

$$\frac{\delta}{a\sqrt{\varepsilon}} = \frac{a^{-1}v + x}{\sqrt{\varepsilon}} \sim \mathcal{U}(0, \xi_{k+1}). \tag{2.11}$$

As a consequence, once we have identified the function G in (2.10), our arguments in the previous section lead to the formula

$$A(x) = \int_0^\infty (az)\left(\frac{1}{z}\int_0^z aG(y)\,dy\right)\nu(dz)$$

where ν is the law of the steps ξ_k. If we can show that

$$\int_0^\infty \int_0^z G(y)\,dy\,\nu(dz) = \frac{1}{2} \,, \tag{2.12}$$

then we will have

$$A(x) = \frac{a^2}{2} \,,$$

a formula consistent with a choice of $c = \frac{1}{2}$ in (2.8).

In order to identify G, we note that if a random walk starting from $\sqrt{\varepsilon}\delta$ survives for some time of order 1 before becoming negative then, with overwhelming probability, it will have reached a height of at least $\varepsilon^{1/4}$ (say). Furthermore, if we condition the random walk $\mathbf{Q}_z^\varepsilon$ to reach a level $\sqrt{\varepsilon}\gamma$ with $1 \ll \gamma \ll \varepsilon^{-1/2}$, one would expect its law to be well approximated by $\sqrt{\varepsilon}\gamma\mathbf{Q}$ when restricted to excursions that survive for a time of order 1.

As a consequence, we expect that

$$\mathbf{Q}_z^\varepsilon \approx \bar{P}_{z,\gamma}\sqrt{\varepsilon}\gamma\mathbf{Q} \,, \qquad \gamma \gg 1 \,,$$

where $\bar{P}_{z,\gamma}$ denotes the probability that the simple random walk (2.9) with $\varepsilon = 1$ started at z reaches the level γ before becoming negative.

The remainder of this section is devoted to the proof of the fact that if we define $\bar{P}_{z,\gamma}$ in this way, then under some integrability assumptions for the one-step probability ν, the limit

$$G(z) = \lim_{\gamma \to \infty} \gamma \bar{P}_{z,\gamma} \,,$$

exists and does indeed satisfy (2.12), independently of the choice of ν. Actually, we will prove these statements for the quantity $P_{z,\gamma} = \bar{P}_{z,\gamma+z}$, which we interpret as the probability that the random walk starting at the origin reaches $[\gamma, \infty)$ before reaching $(-\infty, -z]$. Our first result is as follows:

Proposition 2.2 *Assume that the law ν satisfies $\nu(\{|x| \geq K\}) \leq C \exp(-cK^{\beta})$ for all $K \geq 0$ and some strictly positive constants c, C and β. Then, the limit*

$$G(s) = \lim_{\gamma \to \infty} \gamma P_{s,\gamma} \tag{2.13}$$

exists and satisfies the relations

$$G(s) = \int_{-s}^{\infty} G(s+z)\,\nu(dz)\,, \quad s \geq 0\,, \qquad \lim_{s \to \infty} \frac{G(s)}{s} = 1\,. \tag{2.14}$$

Furthermore, for every $\delta > 0$ there exists C such that the bound

$$|(\gamma + s)P_{s,\gamma} - G(s)| \leq C\frac{1+s}{\gamma^{\frac{1}{2}-\delta}}\,, \tag{2.15}$$

holds uniformly for all $s \geq 0$ and $\gamma \geq 1 \vee s$.

Proof. Denote by y_k the kth step of the random walk starting at the origin. Our main tool is the quantitative convergence result [Fra73], which states that the supremum distance between a Wiener process and the diffusively rescaled random walk over n steps is of order $n^{-1/4}$.

As a consequence we claim first that, for every $\delta > 0$ there exists a constant C such that, for every $a \in [\frac{1}{3}, 3]$, we have the bound

$$\left| P_{a\gamma,\gamma} - \frac{a}{1+a} \right| \leq \frac{C}{\gamma^{\frac{1}{2}-\delta}}\,, \tag{2.16}$$

valid for every $\gamma \geq 1$. Indeed, for any $n \geq 1$, it follows from the previously quoted convergence result that there exists a Brownian motion B such that $|B_t - y_{\lfloor t \rfloor}| \leq n^{1/4+\delta}$ for all $t \in [0, n]$ with probability greater than $1 - C/n^q$. Here, $\delta > 0$ and $q \geq 1$ are arbitrary, but the constant C of course depends on them.

Take n such that $n^{1/4+\delta} \leq \gamma$. If y hits $[\gamma, \infty)$ before $(-\infty, -a\gamma]$, then either $\sup_{t \leq n} |B_t - y_{\lfloor t \rfloor}| > n^{1/4+\delta}$, or $\sup_{t \leq n} |B_t| \leq 3\gamma$, or B hits $[\gamma - n^{1/4+\delta}, \infty)$ before it hits $(-\infty, -a\gamma - n^{1/4+\delta}]$. As a consequence,

$$P_{a\gamma,\gamma} \leq \frac{a\gamma + n^{1/4+\delta}}{(1+a)\gamma} + \frac{C}{n^q} + \exp(-cn/\gamma^2)\,.$$

Reversing the roles of γ and $a\gamma$, we thus obtain the bound

$$\left| P_{a\gamma,\gamma} - \frac{a}{1+a} \right| \lesssim \frac{n^{1/4+\delta}}{\gamma} + \frac{1}{n^q} + \exp(-cn/\gamma^2)\,.$$

Choosing δ small enough and $n = \gamma^{2+\delta}$, the claim (2.16) then follows.

In order to obtain the convergence of the right hand side in (2.13), we make use of the fact that, for $\bar{\gamma} > \gamma$, one has the identity

$$P_{s,\bar{\gamma}} = P_{s,\gamma} \int_0^\infty P_{s+\gamma+z,\bar{\gamma}-\gamma-z}\,\nu_\gamma(dz)\,, \qquad (2.17)$$

where ν_γ is the law of the "overshoot" $y_n - \gamma$ at the first time n such that $y_n \geq \gamma$, conditioned on never reaching below the level $-s$. Since $P_{s+\gamma+z,\bar{\gamma}-\gamma-z}$ is an increasing function of z, we immediately obtain the lower bound

$$P_{s,\bar{\gamma}} \geq P_{s,\gamma} P_{s+\gamma,\bar{\gamma}-\gamma}\,,$$

If we choose $\gamma = a\bar{\gamma}$ for $a \in [\frac{1}{4}, \frac{1}{2}]$, it then follows from (2.16) that

$$P_{s,\bar{\gamma}} \geq P_{s,\gamma}\Big(\frac{\gamma+s}{\bar{\gamma}+s} - \frac{C}{\gamma^{\frac{1}{2}-\delta}}\Big)\,,$$

for all γ sufficiently large and uniformly over all $s \in [0,\gamma]$. Setting $Q_{s,\gamma} = (\gamma+s)P_{s,\gamma}$, it thus follows that one has the bound

$$Q_{s,\bar{\gamma}} \geq Q_{s,\gamma}\Big(1 - \frac{C}{\gamma^{\frac{1}{2}-\delta}}\Big)\,, \qquad (2.18)$$

possibly for a different constant C. Let $\gamma_0 \geq 1$ be such that the factor on the right of this equation is greater than $1/2$. By (2.16), there then exists $s_0 \geq \gamma_0$ such that for $s \geq s_0$ and $\gamma \in [s, 2s]$ one has $Q_{s,\gamma} \geq C(1+s)$. Furthermore, for $s \leq s_0$ and $\gamma \in [(1 \vee s), 2(\gamma_0 \vee s)]$, there exists a non-zero constant such that $Q_{s,\gamma} \geq C$. Iterating (2.18), we then conclude that there exists a constant $C > 0$ such that the bound

$$Q_{s,\gamma} \geq C(1+s)\,,$$

holds uniformly over all $s > 0$ and all $\gamma \geq 1 \vee s$.

On the other hand, for arbitrary $\alpha > 0$, one has from (2.17) the lower bound

$$P_{s,\bar{\gamma}} \leq (\nu_\gamma(\{x > \gamma^\alpha\}) + P_{s+\gamma+\gamma^\alpha,\bar{\gamma}-\gamma-\gamma^\alpha})P_{s,\gamma}\,.$$

In order to bound $\nu_\gamma(\{x > \gamma^\alpha\})$, we note that this event can happen only if either one of the first γ^3 increments exceeds γ^α, or the random walk never exceeds the value γ within these γ^3 steps. Similarly to before, it then follows that

$$\nu_\gamma(\{x > \gamma^\alpha\}) \lesssim \frac{1}{P_{s,\gamma}}(\gamma^3 \exp(-c\gamma^{\alpha\beta}) + \gamma^{-q} + \exp(-c\gamma))\,,$$

for every $q > 0$ and uniformly over $s \leq \gamma$. It follows from the lower bound on $P_{s,\gamma}$ obtained previously that $\nu_\gamma(\{x > \gamma^\alpha\}) \lesssim \gamma^{-q}$ for any power $q > 0$, so that we obtain the upper bound

$$P_{s,\bar{\gamma}} \leq P_{s,\gamma}\Big(\frac{\gamma+s}{\bar{\gamma}+s} + \frac{C}{\gamma^{\frac{1}{2}-\delta}}\Big)\,, \qquad (2.19)$$

with the same domain of validity as before. Using a very similar argument as before, we obtain a constant $\bar{C}$ such that $\gamma P_{s,\gamma} \leq \bar{C}(1+s)$ uniformly over $s > 0$ and $\gamma \geq 1 \vee s$.

Combining the bounds we just obtained, we obtain

$$|\bar{\gamma}P_{s,\bar{\gamma}} - \gamma P_{s,\gamma}| \leq \frac{C}{\gamma^{\frac{1}{2}+\delta}}\,,$$

uniformly over $\bar{\gamma} > \gamma > s$, from which it follows immediately that the sequence $\{\gamma P_{s,\gamma}\}_{\gamma \geq 1}$ is Cauchy, so that it has a limit $G(s)$.

It remains to show that G has the desired properties. The first one follows immediately from the identity

$$\gamma P_{s,\gamma} = \int_{\mathbf{R}} \gamma P_{s+z,\gamma-z} \, \nu(dz) \,,$$

which holds provided that we define the integrand to be 1 for $z > \gamma$ and 0 for $z < -s$.

In order to show that $G(s)/s \to 1$, we fix some (large) value s and choose $\gamma_n = 2^n s$. It then follows from (2.16) that

$$|Q_0 - s| \lesssim s^{\frac{1}{2}+\delta} \,,$$

where we used the notation $Q_n = (\gamma_n + s)P_{s,\gamma_n}$ as a shorthand. Furthermore, it follows immediately from (2.19) that there exists a constant C independent of s such that $|Q_n| \leq Cs$ uniformly in n. As a consequence, we obtain the recursive bound

$$|Q_n - Q_{n-1}| \leq \frac{Cs}{\gamma_n^{\frac{1}{2}-\delta}} \,.$$

Summing over n yields $|Q_n - s| \lesssim s^{\frac{1}{2}+\delta}$, uniformly in n, so that the claim follows. The quantitative error bound (2.15) follows in the same way. $\square$

Corollary 2.3 *In the same setting as above, one has the bound*

$$\left| P_{s,\gamma} - \frac{s}{\gamma+s} \right| \lesssim \frac{1 + s^{\frac{1}{2}+\delta}}{\gamma} \,,$$

uniformly for all $s \geq 0$ and $\gamma \geq 1 \vee s$.

Proof. Combine (2.15) with the bounds on $G(s) - s$ obtained at the end of the proof above. $\square$

Somewhat surprising is the fact that the function G obtained in the Proposition 2.2 does indeed satisfy (2.12), independently of the choice of transition probability ν, provided that we assume that ν has some exponential moment.

Proposition 2.4 *Let G be as in Proposition 2.2 and assume that the law ν satisfies*

$$\int_{\mathbf{R}} e^{c|z|} \nu(dz) < \infty \,,$$

for some $c > 0$. Then, one has the identity $\int_0^\infty \nu([s,\infty)) \, G(s) \, ds = \frac{1}{2}$.

Remark 2.5 Note that, by Fubini's theorem,

$$\int_0^\infty \nu([s,\infty)) \, G(s) \, ds = \int_0^\infty \int_0^s G(y) \, dy \, \nu(ds) \,,$$

so that we do obtain (2.12).

Proof. Integrating (2.14) from 0 to an arbitrary value $K > 0$ and applying Fubini's theorem, we obtain the identity

$$\int_0^K G(s)\,ds = \int_0^\infty G(z)\,\nu([z-K,z])\,dz \;. \tag{2.20}$$

In this proof we denote by $\mathcal{I} = \int_0^\infty G(z)\nu([z,\infty))\,dz$ the quantity of interest. Simple algebraic manipulations then yield from (2.20)

$$\begin{aligned}
\mathcal{I} &= \int_0^\infty G(z)\,\nu([z-K,\infty))\,dz - \int_0^\infty G(z)\,\nu([z-K,z])\,dz \\
&= \int_0^\infty G(z)\,\nu([z-K,\infty))\,dz - \int_0^K G(z)\,dz \\
&= \int_0^\infty G(z)(\nu([z-K,\infty)) - \mathbf{1}_{z<K})\,dz \;.
\end{aligned}$$

Since this identity holds for every $K > 0$, it follows in particular that one has

$$\begin{aligned}
\mathcal{I} &= \varepsilon \int_0^\infty G(z) \int_0^\infty (\nu([z-K,\infty)) - \mathbf{1}_{z<K})e^{-\varepsilon K}\,dK\,dz \\
&= \int_0^\infty G(z)\Big(\varepsilon \int_0^\infty \nu([z-K,\infty))e^{-\varepsilon K}\,dK - e^{-\varepsilon z}\Big)\,dz \;, \tag{2.21}
\end{aligned}$$

for every $\varepsilon > 0$. At this stage, we note that one has the identity

$$\varepsilon \int_0^\infty \nu([z-K,\infty))e^{-\varepsilon K}\,dK = e^{-\varepsilon z}\mathbf{E}e^{\varepsilon\xi} - \varepsilon e^{-\varepsilon z} \int_z^\infty e^{\varepsilon K}\nu([K,\infty))\,dK \;,$$

where ξ denotes an arbitrary random variable with law ν. Since ν has some exponential moment by assumption, $\nu([K,\infty))$ decays exponentially so that the second term in this identity satisfies

$$\left|\varepsilon e^{-\varepsilon z} \int_z^\infty e^{\varepsilon K}\nu([K,\infty))\,dK\right| \le C\varepsilon e^{-\gamma z} \;,$$

for some constants $\gamma, C > 0$, provided that ε is small enough. Inserting this into (2.21), it follows that

$$\mathcal{I} = \int_0^\infty G(z)e^{-\varepsilon z}\mathbf{E}(e^{\varepsilon\xi} - 1)\,dz + \mathcal{O}(\varepsilon) \;.$$

At this stage, we use again the fact that ξ has exponential moments to deduce that

$$\mathbf{E}e^{\varepsilon\xi} - 1 = \frac{\varepsilon^2}{2} + \mathcal{O}(\varepsilon^3) \;,$$

where we used the fact that $\mathbf{E}\xi = 0$ and $\mathbf{E}\xi^2 = 1$, so that

$$\mathcal{I} = \frac{\varepsilon^2}{2} \int_0^\infty G(z)e^{-\varepsilon z}\,dz + \mathcal{O}(\varepsilon) \;.$$

It then follows from the fact that $\lim_{s\to\infty} G(s)/s = 1$ and the dominated convergence theorem that

$$\mathcal{I} = \lim_{\varepsilon\to 0} \frac{\varepsilon^2}{2} \int_0^\infty ze^{-\varepsilon z}\,dt = \frac{1}{2} \;,$$

which is precisely the desired expression. $\qquad\square$

Remark 2.6 It is clear that these results should hold under much weaker integrability conditions on ν. However, since we need some exponential moments on ν at several places in the sequel, we did not try to improve on this.

2.2 Some properties of the limiting process

In this section, we provide a rigorous definition of the limiting process loosely defined in Section 2.1.1, and we study some of its properties. In order to be able to use existing results on Brownian excursions, we restrict ourselves to the same situation as in Section 2.1.2, namely the case where the underlying diffusion is a Brownian motion and the potential $V(x) = -ax$ is linear. We call the resulting object the *Brownian fan*.

2.2.1 Recursive Poisson point processes

Before we give a formal definition of the Brownian fan, we define a "recursive Poisson point process". Loosely speaking, this is a Crump-Mode-Jagers process [Jag75] with Poisson distributed offspring, but where the number of offspring of any given individual is allowed to be almost surely infinite. Note again that our construction is very similar to the one given in [Hut09]. Given a Polish space $\mathcal{X}$ and a function $F\colon \mathcal{X} \to \mathbf{R}_+$, we denote throughout this section by $\mathscr{M}_+^F(\mathcal{X})$ the space of σ-finite measures μ on $\mathcal{X}$ such that

$$\mu(F^{-1}(0)) = 0 \,, \qquad \mu(\{x\,:\,F(x) > \varepsilon\}) < \infty \,,$$

for all $\varepsilon > 0$. We endow this with the topology of convergence in total variation on each set of the form $\{x\,:\,F(x) > 1/n\}$. Given a (measurable) map $\mathcal{Q}$ from $\mathcal{X}$ to $\mathscr{M}_+^F(\mathcal{X})$, we can then build for every $x \in \mathcal{X}$ a point process as follows.

Define $\mu_x^0 = \delta_x$ and, for $n \geq 1$, define μ_x^n recursively as a (conditionally independent of the μ_x^ℓ with $\ell < n$) realisation of a Poisson point process with intensity measure

$$\mathcal{Q}_n = \int_{\mathcal{X}} \mathcal{Q}(y)\,\mu_x^{n-1}(dy) \,,$$

where we view μ_x^n as a random σ-finite positive integer-valued measure on $\mathcal{X}$. (In principle, it may happen that $\mathcal{Q}_n(\{x\,:\,F(x) > \varepsilon\}) = \infty$ for some $\varepsilon > 0$. In this case, our construction stops there.)

When then set

$$\mu_x^{[n]} = \sum_{\ell=0}^n \mu_x^n \,, \tag{2.22}$$

and we call $\mu_x^{[n]}$ the *recursive Poisson point process* of depth n with kernel $\mathcal{Q}$. We will occasionally need to refer the Brownian fan spawned by an initial Brownian motion w. For this purpose we will use the symbol $\mu_w^{[n]}$ (or μ_w^n for a specific generation) and rely on the context to differentiate $\mu_x^{[n]}$ and $\mu_w^{[n]}$. If, in these symbols we omit the subscript entirely then it is assumed that $x = 0$. In general, there is no reason to expect the sequence $\mu_x^{[n]}$ to converge to a finite limit. However, one has the following simple criterion ensuring that this is the case:

Lemma 2.7 *Let F and $\mathcal{Q}$ be as above and assume that there exists $c < 1$ such that $\int F(y)\mathcal{Q}(x, dy) \leq cF(x)$ for every $x \in \mathcal{X}$. Then, for every $x \in \mathcal{X}$, there exists a random σ-finite measure $\mu_x^{[\infty]}$ on $\mathcal{X}$ such that $\lim_{n\to\infty} \mathbf{E} \int F(y)(\mu_x^{[\infty]} - \mu_x^{[n]})(dy) = 0$.*

Proof. Fix $x \in \mathcal{X}$ and denote by $\mathscr{F}_n$ the σ-algebra generated by $\mu_x^{[n]}$. It then follows from the definition of the $\mu_x^{[n]}$ that

$$\mathbf{E}\Big(\int_{\mathcal{X}} F(y)\,\mu_x^{\ell+1}(dy)\,\Big|\,\mathscr{F}_\ell\Big) = \int_{\mathcal{X}}\int_{\mathcal{X}} F(y)\mathcal{Q}(z,dy)\,\mu_x^\ell(dz) \le c\int_{\mathcal{X}} F(z)\,\mu_x^\ell(dz)\ .$$

As a consequence, one has $\mathbf{E}\int_{\mathcal{X}} F(y)\,\mu_x^\ell(dy) \le c^\ell F(x)$, and the claim follows. $\qquad\square$

Remark 2.8 A useful identity is the following. Denote by $\{\tilde{\mu}_y^{[\infty]}\}_{y\in\mathcal{X}}$ a collection of *independent* copies of recursive Poisson point processes with "initial conditions" y. Then one has the identity in law

$$\mu_x^{[\infty]} \stackrel{\text{law}}{=} \delta_x + \int_{\mathcal{X}} \tilde{\mu}_y^{[\infty]}\,\mu_x^1(dy)\ , \tag{2.23}$$

where, as before, μ_x^1 is a realisation of a Poisson point process with intensity $\mathcal{Q}(x,\cdot)$, which is itself independent of the $\tilde{\mu}_y^{[\infty]}$. This identity makes sense since the integral on the right is really just a countable sum.

2.2.2 Construction of the Brownian fan

We now denote by $\mathcal{E}$ the set of excursions with values in $\mathbf{R}$. We consider elements of $\mathcal{E}$ as triples (s,t,y) Where $s < t \in \mathbf{R} \cup \{+\infty\}$, and $y \in \mathcal{C}(\mathbf{R},\mathbf{R})$ has the property that $y_\tau = y_t$ for $\tau \ge t$ and $y_\tau = y_s$ for $\tau \le s$. We also write $\mathcal{E}_0$ for the subset of those triples (s,t,y) such that $s = 0$.

Denoting a generic excursion by w, we write $\mathfrak{s}(w)$ for its starting time and $\mathfrak{e}(w)$ for its end time, *i.e.*, $\mathfrak{s}(s,t,y) = s$, and $\mathfrak{e}(s,t,y) = t$. We also denote by $\mathfrak{l}(w)$ the lifetime of the excursion, which is the interval $\mathfrak{l}(w) = [\mathfrak{s}(w), \mathfrak{e}(w)]$. In order to keep notations compact, we will also identify an excursion with its path component, making the abuse of notation $w_t = y_t$. There is a natural metric on $\mathcal{E}$ given by

$$d(\tilde{w}, w) = d_{\mathfrak{l}}(\tilde{w}, w) + \sum_{k\ge 1} 2^{-k}\Big(1 \wedge \sup_{|t|\le 2^k} |\tilde{w}_t - w_t|\Big)\ , \tag{2.24}$$

where the distance $d_{\mathfrak{l}}$ between the supports is given by

$$d_{\mathfrak{l}}(\tilde{w}, w) = 1 \wedge \big(|\mathfrak{s}(\tilde{w}) - \mathfrak{s}(w)| + |\tanh\mathfrak{e}(\tilde{w}) - \tanh\mathfrak{e}(w)|\big)\ .$$

The reason for this particular choice of metric is that it ensures that $\mathcal{E}$ is a Polish space, while still allowing for infinite excursions.

For $\tau \in \mathbf{R}$ and $v \in \mathcal{E}$, we denote by $\Theta_{v,\tau}\colon \mathcal{E} \to \mathcal{E}$ the shift map given by

$$\Theta_{v,\tau}\colon (s,t,w) \mapsto (s+\tau, t+\tau, w_{\cdot+\tau} + v_\tau)\ ,$$

which essentially changes the coordinate system so that the origin $(0,0)$ is mapped to (τ, v_τ). Denoting as before by $\mathbf{Q}$ the standard Itô excursion measure, we now give the following definition:

Definition 2.9 The *Brownian fan* with intensity $a > 0$ is the recursive Poisson point process on $\mathcal{E}$ with kernel

$$\mathcal{Q}(w,\cdot) = \frac{a}{2}\int_{\mathfrak{s}(w)}^{\mathfrak{e}(w)} \Theta_{w,\tau}^\star \mathbf{Q}\,d\tau\ , \tag{2.25}$$
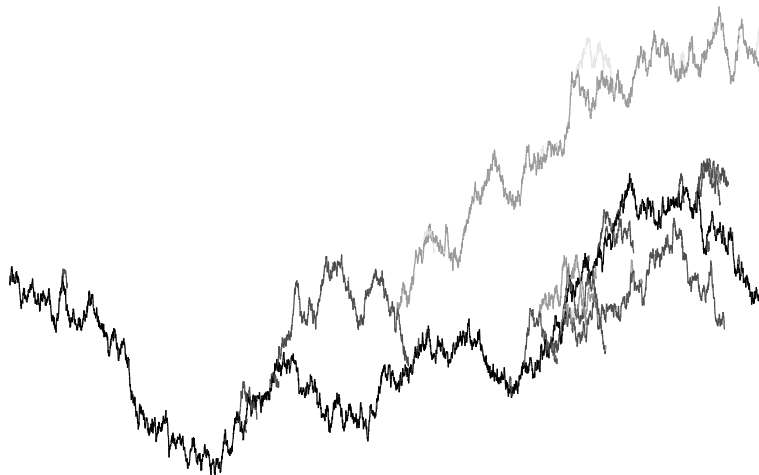
Figure 1: A typical realisation of the Brownian fan with intensity $a = \frac{3}{2}$ over the time interval $[0, 1]$. Successive generations are draw in lighter shades of grey.

and initial condition given by a realisation of Brownian motion, starting at the origin and killed when it reaches the level $-L$, where L is exponentially distributed with mean a.

Remark 2.10 The reason for killing the original Brownian motion at this particular level is natural, due to the distribution of the initial tag in Step 1 of the algorithm. It is however essentially irrelevant to the mathematical construction.

Remark 2.11 Formally, the Brownian fan is a particular case of the Virgin Island Model [Hut09] with a playing the same role in both models, $h = a$, and $g = 1/2$. The differences are twofold. First, the case of constant non-vanishing a actually doesn't fall within the framework of [Hut09] since the author there uses the standing assumption that $a(0) = 0$. The other difference is mostly one of perspective. while we have so far defined the Brownian fan as a point process on a space of excursions, one of the purposes of this article is to show that it is also well-behaved as an actual Markov process with values in a suitable state space of (possibly infinite) point configurations.

Remark 2.12 By only keeping track of the genealogy of the particles and not their precise locations, one can construct a "real tree" on top of which the Brownian fan is constructed (loosely speaking) by attaching a Brownian excursion to each branch. This is very similar in spirit to Le Gall's construction of the Brownian map [LG07, LG10], starting from Aldous's continuous random tree [Ald91]. The scaling properties of the Brownian fan however are quite different. In particular, Theorem 2.14 below implies that the underlying tree has Hausdorff dimension 1, as opposed to the CRT which has Hausdorff dimension 2.

Before we proceed, let us show that it is possible to verify the assumptions of Lemma 2.7, so that this object actually exists for every $a > 0$:

Proposition 2.13 *The kernel $\mathcal{Q}$ defined in (2.25) satisfies the assumption of Lemma 2.7 with the choice*

$$F(w) = e^{-\eta \mathfrak{s}(w)}(1 - e^{-\eta |\mathfrak{l}(w)|}) \, ,$$

provided that η is large enough.

Proof. It follows from the properties of **Q** that there exists a constant C independent of γ and a such that

$$\int_{\mathcal{E}} F(\tilde{w}) \mathcal{Q}(w, d\tilde{w}) = \frac{a}{\sqrt{2\pi}} \int_{\mathfrak{s}(\tilde{w})}^{\mathfrak{e}(\tilde{w})} e^{-\eta \tau} \int_0^\infty (1 - e^{-\eta s}) s^{-3/2} \, ds \, d\tau$$

$$\leq \frac{Ca}{\sqrt{\eta}} (e^{-\eta \mathfrak{s}(\tilde{w})} - e^{-\eta \mathfrak{e}(\tilde{w})}) = \frac{Ca}{\sqrt{\eta}} F(v) \, ,$$

where C is a constant independent of a and η. The claim then follows by choosing $\sqrt{\eta} > Ca$. □

The remainder of this section is devoted to a study of the basic properties of the Brownian fan. In particular, we will show that there exists a suitable space $\mathcal{X}$ of particle configurations such that it can be viewed as a $\mathcal{X}$-valued Markov process with continuous sample paths that satisfies the Feller property.

2.2.3 Number of particles and workload rate

Define the set $\mathcal{N}_t \subset \mathcal{E}$ of excursions that are "alive at time t" by

$$\mathcal{N}_t = \{w \in \mathcal{E} \, : \, \mathfrak{s}(w) < t < \mathfrak{e}(w)\} \, .$$

With this notation, the number of particles alive at time t for the Brownian fan is given by

$$N_t = \mu^{[\infty]}(\mathcal{N}_t) \, ,$$

which is in principle allowed to be infinite.

Theorem 2.14 *There exist a constant $C > 0$ and a strictly positive continuous decreasing function $\lambda \colon \mathbf{R}_+ \to \mathbf{R}_+$ such that*

$$\mathbf{E} \exp(\lambda_t N_t) \leq C \, , \tag{2.26}$$

holds uniformly over all $t > 0$.

Remark 2.15 We will see in the proof that one can choose λ of the form

$$\lambda_t = K^{-1} e^{-Kt} \, , \tag{2.27}$$

for K sufficiently large.

Proof. For $\lambda > 0$ and $s, t \geq 0$, set

$$N_{s,t}^\lambda = \log \mathbf{E} \exp(\lambda \mu_w^{[\infty]}(\mathcal{N}_t)) \, ,$$

where w is any excursion starting from 0 with lifetime s. Since, by the definition (2.25), the value of $N_{s,t}^\lambda$ does not depend on the precise choice of excursion, we do not include it

in the notation. It also follows from the construction of $\mu_w^{[\infty]}$ that the function $t \mapsto N_{t,t}^\lambda$ is increasing in t and that $N_{s,t}^\lambda = N_{t,t}^\lambda$ for $s \geq t$.

We also define M_t^λ by

$$M_t^\lambda = \log \mathbf{E} \exp\Big(\lambda \int_{\mathcal{E}} \mu_w^{[\infty]}(\mathcal{N}_t) \,\mathscr{M}(dw)\Big) \,, \tag{2.28}$$

where $\mathscr{M}$ is a Poisson random measure with intensity measure $\mathbf{Q}$ and the realisations $\mu_w^{[\infty]}$ are independent of $\mathscr{M}$ and of each other. While N^λ measures the total number of offspring alive at time t due to an excursion starting at time 0, M^λ measures the rate at which these offspring are created.

Indeed, combining (2.23) with the definition of μ_w^1 and the superposition principle for Poisson point processes, we have the identity

$$N_{s,t}^\lambda = \lambda \mathbf{1}_{s \geq t} + \frac{a}{2} \int_0^{s \wedge t} M_{t-r}^\lambda \, dr \,. \tag{2.29}$$

It therefore remains to obtain suitable bounds on M_t^λ.

It follows from (2.29) and standard properties of Poisson point processes (see for example [PZ07, Theorem 6.3]) that one has the identity

$$M_t^\lambda = \int_{\mathcal{E}} (\mathbf{E} \exp(\lambda \mu_w^{[\infty]}(\mathcal{N}_t)) - 1) \, \mathbf{Q}(dw) = \int_{\mathcal{E}} (e^{N_{\mathfrak{e}(w),t}^\lambda} - 1) \, \mathbf{Q}(dw)$$

$$\leq \int_{\mathcal{E}} \Big(\exp\Big(\lambda \mathbf{1}_{\mathcal{N}_t}(w) + \frac{a}{2} \int_0^{\mathfrak{e}(w) \wedge t} M_{t-s}^\lambda \, ds \Big) - 1 \Big) \, \mathbf{Q}(dw) \,.$$

At this stage, we note that the integrand appearing in this expression depends on w only through $\mathfrak{e}(w)$. It is then convenient to break the integral into a contribution coming from $\mathfrak{e}(w) > t$, as well as its complement. Since, under $\mathbf{Q}$, the quantity $\mathfrak{e}(w)$ is distributed according to the measure $\frac{s^{-3/2}}{\sqrt{2\pi}} \, ds$, this yields

$$M_t^\lambda \leq \Big(\exp\Big(\lambda + \frac{a}{2} \int_0^t M_r^\lambda \, dr \Big) - 1 \Big) \, \mathbf{Q}(\mathfrak{e}(w) \geq t)$$

$$+ \int_0^t \Big(\exp\Big(\frac{a}{2} \int_0^s M_{t-r}^\lambda \, dr \Big) - 1 \Big) \frac{s^{-3/2}}{\sqrt{2\pi}} \, ds \,.$$

We now assume that both λ and t are sufficiently small so that

$$\lambda + \frac{a}{2} \int_0^t M_r^\lambda \, dr \leq 1 \,. \tag{2.30}$$

This assumption allows us to use the bound $e^t - 1 \leq 2t$, so that we obtain the more manageable expression

$$M_t^\lambda \leq \frac{4t^{-1/2}}{\sqrt{2\pi}} \Big(\lambda + \frac{a}{2} \int_0^t M_r^\lambda \, dr \Big) + a \int_0^t \int_0^s M_{t-r}^\lambda \, dr \, \frac{s^{-3/2}}{\sqrt{2\pi}} \, ds$$

$$= \frac{4\lambda t^{-1/2}}{\sqrt{2\pi}} + \frac{2at^{-1/2}}{\sqrt{2\pi}} \int_0^t M_r^\lambda \, dr + a \int_0^t \int_s^t M_r^\lambda \, dr \, \frac{(t-s)^{-3/2}}{\sqrt{2\pi}} \, ds$$

$$= \frac{4\lambda t^{-1/2}}{\sqrt{2\pi}} + \frac{2at^{-1/2}}{\sqrt{2\pi}} \int_0^t M_r^\lambda \, dr + \frac{2a}{\sqrt{2\pi}} \int_0^t (t-s)^{-1/2} M_s^\lambda \, ds$$

$$\leq \frac{4\lambda t^{-1/2}}{\sqrt{2\pi}} + \frac{4a}{\sqrt{2\pi}} \int_0^t (t-s)^{-1/2} M_s^\lambda \, ds \, .$$

Writing $H_t^\lambda = t^{1/2} M_t^\lambda$, we thus obtain the bound

$$H_t^\lambda \leq \frac{4\lambda}{\sqrt{2\pi}} + \frac{4a}{\sqrt{2\pi}} t^{1/2} \int_0^t (t-s)^{-1/2} s^{-1/2} H_s^\lambda \, ds \, . \tag{2.31}$$

We can now apply the fractional version of Gronwall's lemma [NR02, Lemma 7.6] (with $b = \frac{4a}{\sqrt{2\pi}}$, $a = \frac{4\lambda}{\sqrt{2\pi}}$, and $\alpha = \frac{1}{2}$), so that there exists a constant $C > 0$ depending on a but independent of λ such that

$$H_t^\lambda \leq C\lambda \exp(Ct) \, .$$

From this, we immediately deduce from (2.29) and the definition of H^λ a similar bound on $N_{t,t}^\lambda$. Choosing $\lambda = K^{-1} e^{-Kt}$ for sufficiently large t then allows to satisfy (2.30) and to obtain $N_{t,t}^\lambda \leq 2$, thus completing the proof. □

As a corollary, we obtain a rather sharp bound on the modulus of continuity of the total workload process $\mathcal{W}_t = \int_0^t N_s \, ds$. One has

Proposition 2.16 *For every $T > 0$, one has*

$$\sup_{t \leq T} \lim_{h \to 0} \frac{|\mathcal{W}_{t+h} - \mathcal{W}_t|}{h|\log h|} < \infty \, , \tag{2.32}$$

almost surely.

Proof. It follows from the generalised Young inequality that, for every $a, b \in \mathbf{R}_+$, and every $\lambda, \eta > 0$, one has the inequality

$$ab \leq \frac{\eta}{\lambda}(e^{\lambda a} - 1 - \lambda a + (1 + b/\eta)\log(1 + b/\eta) - b/\eta) \, ,$$

so that

$$N_t \leq \frac{\eta}{\lambda}(e^{\lambda N_t} + (1 + 1/\eta)\log(1 + 1/\eta)) \, .$$

It follows immediately that

$$|\mathcal{W}_{t+h} - \mathcal{W}_t| \leq \frac{\eta}{\lambda} \int_t^{t+h} e^{\lambda N_s} \, ds + \frac{h(1 + \eta)}{\lambda} \log(1 + 1/\eta) \, .$$

Setting $\eta = h$, we obtain the bound

$$|\mathcal{W}_{t+h} - \mathcal{W}_t| \leq \frac{h}{\lambda} \int_0^{T+1} e^{\lambda N_s} \, ds + C_\lambda h|\log h| \, ,$$

uniformly over all $h \leq 1$ and all $t \in [0, T]$. The claim now follows immediately from Theorem 2.14. □

Although the number of particles alive at any deterministic time has exponential moments, there exists a *dense* set of exceptional times for which $N_t = \infty$. For one, this follows from the fact that, under $\mathbf{Q}$, $\mathfrak{e}(w)$ is distributed proportionally to $s^{-3/2} ds$, so that every particle creates an infinite number of offspring in every time interval.

Actually, one has the even stronger statement that there is a dense set of exceptional times at which the number of particles belonging to the *first* generation of offspring is infinite. Indeed, if we denote by $\mathscr{M}$ a Poisson random measure on $\mathbf{R}_+^2$ with density $cs^{-3/2}\,dr\,ds$ for a suitable constant c, then the number N_t^1 of particles in the first generation of offspring is given by

$$N_t^1 = \mathscr{M}(A_t)\,, \qquad A_t = \{(r,s) \in [0,t] \times \mathbf{R}_+ \,:\, s \geq t - r\}\,.$$

For $k \geq 0$ and $d \in \{1, \ldots 2^k\}$, we then set

$$A_{k,d} = [(d-1)2^{-k}, d2^{-k}] \times [4^{-k}, 4^{1-k}]\,,$$

so that, by the scaling properties of $\mathscr{M}$, the random variables $N_{k,d} = \mathscr{M}(A_{k,d})$ form a sequence of i.i.d. Poisson random variables. For any given point (r,s), we set $B_{(r,s)} = [r, r+s]$, which is the set of times t such that $(r,s) \in A_t$, and we set

$$D_{(r,s)} = \{(k,d) \,:\, B_{(r',s')} \subset B_{(r,s)} \,\forall (r',s') \in A_{k,d}\}\,.$$

Since the set $D_{(r,s)}$ is infinite for every (r,s), we can then build a sequence (r_n, s_n) recursively in the following way. Start with $(r_0, s_0) = (0,1)$ and then, given (r_n, s_n) for some $n \geq 0$, define (k_n, d_n) as the first (in lexicographic order) element $(k,d) \in D_{(r_n,s_n)}$ such that $N_{k,d} \geq 1$. We then set (r_{n+1}, s_{n+1}) to be one of the points of $\mathscr{M}$ located in A_{k_n,d_n}. By construction, one then has $\cap_{n\geq 1} B_{(r_n,s_n)} = \{t\}$ for some $t \in [0,1]$, and $\mathscr{M}(A_t) \geq \sum_n \mathscr{M}(A_{k_n,d_n}) = \infty$, as stated. Of course, the interval $[0,1]$ in this procedure is arbitrary. If we want to show that there exists an exceptional time within any deterministic time interval $[t_0, t_1]$, it suffices to start the algorithm we just described with $r_0 = t_0$ and $s_0 = t_1 - t_0$.

3   The Brownian fan as a Markov process

In this section, we slightly shift our perspective. We no longer consider the Brownian fan as a point process of excursions, but we consider it as an evolving system of particles. Our system will therefore be described by a Markov process in some space of integer-valued measures on a subset of $\mathbf{R}^2$ corresponding to the admissible combinations of "position + tag". The problem is that, as we have seen in the previous section, there are exceptional times at which the limiting process consists of infinitely many particles. The first challenge is therefore to construct a space $\mathcal{X}$ of integer-valued measures with a sensible topology which can still accommodate these "bad" configurations in such a way that the limiting process is continuous both as a function of time and as a function of its initial configuration.

Once this space is defined, we show that the Brownian fan possesses the Feller property in $\mathcal{X}$ (i.e. the corresponding Markov semigroup leaves the space of bounded continuous functions invariant). In fact, we will show that it preserves the space of Lipschitz continuous functions. The continuity property of the discrete-time process established below in Proposition 4.8 also implies the time continuity of the Brownian fan (see Theorem 5.1 below). These properties allow us to conclude that the Brownian fan $t \mapsto \mu_t$ is in fact a strong Markov process.

In Section 3.4 below, we will furthermore compute its generator A on a class of "nice" test functions. We will also present a "negative" result showing that if we denote by T_ε the one-step Markov operator corresponding to the evolution of Algorithm TDMC, then one has $A \neq \lim_{\varepsilon \to 0} \varepsilon^{-1}(T_\varepsilon - 1)$. This is in stark contrast with, for example, Euler approximations to stochastic differential equations, where such an equality would hold, at least when applied to sufficiently regular test functions.

3.1 State space

Our construction is essentially the Wasserstein-1 analogue of the construction given in [FG10]. Let $\mathcal{M} \subset \mathbf{R}^n$ be a convex open set with boundary $\partial\mathcal{M}$. For $p \in (0, 1]$, we then denote by $\ell^p(\mathcal{M})$ the set of all integer-valued measures μ on $\mathcal{M}$ such that

$$\|\mu\|_p = \int_{\mathcal{M}} d^p(y, \partial\mathcal{M})\, \mu(dy) < \infty\,, \tag{3.1}$$

where $d(y, \partial\mathcal{M})$ denotes the (Euclidean) distance from y to the boundary of $\mathcal{M}$. Note that since this quantity vanishes at the boundary, there are elements $\mu \in \ell^p(\mathcal{M})$ such that $\mu(\mathcal{M}) = \infty$.

We endow $\ell^p(\mathcal{M})$ with a slight modification of the Wasserstein-1 metric by setting:

$$\|\mu - \nu\|_p = \sup_{f \in \mathrm{Lip}_p^0(\mathcal{M})} \left(\int f(y)\mu(dy) - \int f(y)\nu(dy) \right)\,, \tag{3.2}$$

where we denoted by $\mathrm{Lip}_p^0(\mathcal{M})$ the set of all functions $f \colon \mathcal{M} \to \mathbf{R}$ such that

$$|f(x) - f(y)| \le |x - y|^p\,, \tag{3.3}$$

for all $x, y \in \mathcal{M}$, and $f(y) = 0$ for all $y \in \partial\mathcal{M}$.

Remark 3.1 Our notation is consistent in the sense that if we take for ν the null measure in (3.2), then we precisely recover (3.1). This can be seen by taking $f(x) = d^p(x, \partial\mathcal{M})$, which is optimal by (3.3) and the triangle inequality.

If μ and ν happen to have the same (finite) mass, then the expression (3.2) does not change when one adds a constant to f. In this case, we are thus reduced to the usual Wasserstein-1 distance between μ and ν, but with respect to the modified distance function

$$d_p(x, y) = |x - y|^p \wedge (d^p(x, \partial\mathcal{M}) + d^p(y, \partial\mathcal{M}))\,.$$

Note that the completion of $\mathcal{M}$ under the distance function d_p consists of $\mathcal{M} \cup \{\Delta\}$, where Δ is a single "point on the boundary" such that $d_p(x, \Delta) = d^p(x, \partial\mathcal{M})$ for every $x \in \mathcal{M}$.

If one has $\mu(\mathcal{M}) < \nu(\mathcal{M}) < \infty$, then the distance $\|\cdot\|_p$ reduces to the Wasserstein-1 distance (again with respect to d_p) between $\bar{\mu}$ and ν, where $\bar{\mu}$ is obtained from μ by placing a mass $\nu(\mathcal{M}) - \mu(\mathcal{M})$ on the boundary Δ. The following alternative characterisation of (3.2) in the case of purely atomic measures is a version of the Monge-Kantorovich duality in this context:

Lemma 3.2 *Consider a situation where $\mu = \sum_{i=1}^{N} \delta_{x_i}$ and $\nu = \sum_{i=1}^{M} \delta_{y_i}$. Then,*

$$\|\mu - \nu\|_p = \inf_{\sigma \in S_{N+M}} \sum_{i=1}^{N+M} d^p(x_i, y_{\sigma(i)})\,,$$

where S_{N+M} is the group of permutations of $N + M$ elements and we set $x_j = \Delta$ for $j > N$ and $y_j = \Delta$ for $j > M$.

Proof. See for example [Pra08]. $\qquad\square$

This characterisation suggests the following "interpolation" procedure between elements in $\ell^p(\mathcal{M})$. Let $\mu = \sum_{i=1}^N \delta_{x_i}$ and $\nu = \sum_{i=1}^N \delta_{y_i}$, where we assumed that both measures charge the same number of points (this is something that we can always achieve by possibly adding points on Δ). assume furthermore that these points are ordered in such a way that

$$\|\mu - \nu\|_p = \sum_{i=1}^N |x_i - y_i|^p .$$

Again, this can always be enforced by suitably reordering the points and possibly adding points on the boundary. We then define, for $t \in (0,1)$, the "linear interpolation" $L_t(\mu, \nu)$ by

$$L_t(\mu, \nu) = \sum_{i=1}^N \delta_{z_i} , \qquad z_i = ty_i + (1-t)x_i . \tag{3.4}$$

Note that this procedure is not necessarily unique, but it is easy to resolve this ambiguity by optimising over the possible pairings $\{(x_i, y_i)\}$ realising the above construction, according to some arbitrary criteria.

In any case, one can check that this construction has the property that

$$\|L_s(\mu, \nu) - L_t(\mu, \nu)\|_p \le |t - s|^p \|\mu - \nu\|_p , \tag{3.5}$$

for any $s, t \in [0, 1]$, which will be a useful fact in the sequel.

3.2 Definition of the process

For the remainder of this section, we set

$$\mathcal{M} = \{(x, v) \in \mathbf{R}^2 \, : \, v > -ax\} ,$$

which is the natural configuration space for our process. We will use capital letters to distinguish points in $\mathcal{M}$ from points in $\mathbf{R}$. By Theorem 2.14, we already know that, for any fixed time t, the Brownian fan almost surely has only finitely many particles alive at time t. Define now the evaluation map $E_t \colon \mathcal{E} \to \mathcal{M} \cup \{\Delta\}$ by

$$E_t(w) = \begin{cases} (w_t, -aw_{\mathfrak{s}(w)}) & \text{if } t \in \mathfrak{l}(w), \\ \Delta & \text{otherwise.} \end{cases}$$

For a given "initial condition" $(x, v) \in \mathcal{M}$, we then set

$$\mu_t = E_t^\star \mu_w^{[\infty]} ,$$

which is an $\ell^p(\mathcal{M})$-valued random variable. Here, w is a realisation of a Brownian motion starting at x and killed when it hits $-v/a$, and $\mu_w^{[\infty]}$ is the corresponding realisation of the Brownian fan. (Just so that E_t has the correct effect on w, one can for example set $\mathfrak{s}(w) = -1$ and make sure that $w(-1) = -v/a$.) As a consequence of Theorem 2.14 and of our definition of the Brownian fan, we then indeed have $\mu_t \in \ell^p(\mathcal{M})$ for every $p \le 1$.

Note at this stage that we can simply discard the, typically infinite, mass on Δ by identifying measures that only differ on Δ. As already mentioned earlier, this is consistent with the identification $\Delta \sim \partial \mathcal{M}$ already made in the interpretation of the construction of $\ell^p(\mathcal{M})$.

This construction can be extended to any initial condition in $\ell^p(\mathcal{M})$ with finite total mass, by considering independent Brownian fans for each particle. As a consequence of the Markov property of the Brownian excursion and the independence properties of Poisson point processes, it is then straightforward to verify that $t \mapsto \mu_t$ is indeed a Markov process.

Actually, by Proposition 4.2, we know that for any fixed collection of deterministic times $\{t_1, \ldots, t_k\}$, one has $\mu_{t_k}(\mathcal{M}) < \infty$ almost surely, so that our construction determines a probability measure on $(\ell^p(\mathcal{M}))^{\mathbf{R}_+}$ by Kolmogorov's extension theorem. At this stage however, we know absolutely nothing about the continuity properties of this process, and this is the subject of the remainder of this section.

3.3 Feller property

We now show that the Brownian fan constructed in Sections 2.2.2 and 3.2 has the Feller property in $\ell^p(\mathcal{M})$ for every $p \le 1$.

As in [FG10], we could have defined spaces $\ell^p(\mathcal{M})$ in a natural way for $p > 1$. However, the Feller property would fail in this case because of the following simple heuristic argument. For any p, we can change the initial condition by an amount less than δ in ℓ^p by creating N particles at distance $\varepsilon = (\delta/N)^{1/p}$ from the boundary of $\mathcal{M}$. For ε small, the probability that any such particle survives up to time 1 (say) is bounded from below by $c\varepsilon$ for some $c > 0$. On average, the number of survivors will thus be on the order of $\varepsilon N \sim \delta^{1/p} N^{(p-1)/p}$. Furthermore, at time $1/2$, each of these surviving particles will be at a distance of order 1 of the boundary of $\mathcal{M}$. As a consequence, by increasing N but keeping δ fixed (or even sending δ to 0 sufficiently slowly), the law of the process at time 1 with an initial condition arbitrarily close to 0 can be at arbitrarily large distance of 0, so that the Feller property fails.

For $p \le 1$ on the other hand, we have

Proposition 3.3 *For any $p \le 1$, the Brownian fan gives rise to a Feller process in $\ell^p(\mathcal{M})$. Even more, the corresponding Markov semigroup preserves the space of bounded Lipschitz continuous functions.*

Proof. For any two initial conditions μ_0 and $\bar{\mu}_0$, write as before

$$\mu_0 = \sum_{j=1}^{N} \delta_{X_0^{(j)}}, \qquad \bar{\mu}_0 = \sum_{j=1}^{N} \delta_{\bar{X}_0^{(j)}},$$

with the $X_0^{(j)} \in \mathcal{M}$ and $\bar{X}_0^{(j)} \in \mathcal{M}$ chosen in such a way that

$$\|\mu_0 - \bar{\mu}_0\|_p = \sum_{j=1}^{N} d_p(X_0^{(j)}, \bar{X}_0^{(j)}). \tag{3.6}$$

Our aim now is to show that there exists a constant C such that

$$\mathbf{E}\|\mu_t - \bar{\mu}_t\|_p \le C\|\mu_0 - \bar{\mu}_0\|_p,$$

independently of $t \le 1$, where the pair $(\mu_t, \bar{\mu}_t)$ is a particular coupling between the Brownian fans starting from μ_0 and $\bar{\mu}_0$ respectively. Denote by $\mu_t^{(j)}$ the contribution to μ_t originating from the initial particle $X_0^{(j)}$ and similarly for $\bar{\mu}_t^{(j)}$. Then, by the triangle inequality, one obtains the bound

$$\mathbf{E}\|\mu_t - \bar{\mu}_t\|_p \le \sum_{j \ge 1} \mathbf{E}\|\mu_t^{(j)} - \bar{\mu}_t^{(j)}\|_p,$$

so that the claim follows from (3.6) if we can show that

$$\mathbf{E}\|\mu_t^{(j)} - \bar{\mu}_t^{(j)}\|_p \le C d_p(X_0^{(j)}, \bar{X}_0^{(j)}) .$$

In other words, it suffices to consider the special case when both μ_0 and $\bar{\mu}_0$ consist of one single particle, which we denote by $X_0 = (x_0, v_0)$ and $\bar{X}_0 = (\bar{x}_0, \bar{v}_0)$ respectively.

One then constructs a coupling between the two processes μ_t and $\bar{\mu}_t$ by running both particles with the same Brownian motion and spawning children according to the same Poisson process (as long as the corresponding particle is alive). We denote by X_t and $\bar{X}_t$ the evolutions of the two initial particles in $\mathcal{M}$, driven by the same realisation of a Brownian motion, and stopped when they reach $\partial\mathcal{M}$. We can assume without loss of generality that $v_0 + ax_0 < \bar{v}_0 + a\bar{x}_0$, so that the particle X dies before the particle $\bar{X}$. Denoting by τ and $\bar{\tau}$ the respective lifetimes of these particles, one thus has $\tau \le \bar{\tau}$.

Denote now by $\mathscr{M}_t$ the (random) measure on $[0, t] \times \mathcal{M}$ which is such that, for $I \subset [0, t]$ and $A \subset \mathcal{M}$, $\mathscr{M}_t(I \times A)$ is the number of particles in A at time t that are offspring of a particle created from the "ancestor particle" $\bar{X}$ at some time $s \in I$. With this notation, if we denote by $\Xi \colon \mathcal{M} \to \mathcal{M}$ the map

$$\Xi(x, v) = (x + x_0 - \bar{x}_0, v - a(x_0 - \bar{x}_0)) ,$$

then one has the decompositions

$$\mu_t(A) = \mathbf{1}_{\tau \ge t}\delta_{X_t}(A) + \Xi^\star \mathscr{M}_t([0, \tau \wedge t] \times A) ,$$
$$\bar{\mu}_t(A) = \mathbf{1}_{\bar{\tau} \ge t}\delta_{\bar{X}_t}(A) + \mathscr{M}_t([0, \bar{\tau} \wedge t] \times A) .$$

Denote now by δ the *Euclidean* distance between the two initial particles, so that their ℓ^p-distance is δ^p. It then follows immediately from the above decomposition that one has the bound

$$\|\mu_t - \bar{\mu}_t\|_p \le \delta^p + \mathbf{1}_{t \in [\tau, \bar{\tau}]}d_p(\bar{X}_t, \partial\mathcal{M}) + \delta^p \mathscr{M}_t([0, \tau \wedge t] \times \mathcal{M})$$
$$+ \int_{\mathcal{M}} d_p(Y, \partial\mathcal{M})\mathscr{M}_t([\tau \wedge t, \bar{\tau} \wedge t] \times dY) .$$

Since $d(\bar{X}(\tau), \partial\mathcal{M}) = \delta$ by the definition of τ and δ, it follows from Jensen's inequality and the Martingale property of (stopped) Brownian motion that one has the bound $\mathbf{E}\mathbf{1}_{t \in [\tau, \bar{\tau}]}d_p(\bar{X}_t, \partial\mathcal{M}) \le \delta^p$. It also follows from Theorem 2.14 that $\mathbf{E}\mathscr{M}_t([0, \tau \wedge t] \times \mathcal{M}) < \infty$, independently of δ. Finally, it follows from an argument very similar to the proof of Theorem 2.14 that

$$\mathbf{E}\int_{\mathcal{M}} d_p(Y, \partial\mathcal{M})\mathscr{M}_t(ds \times dY) \le C\,ds ,$$

uniformly over $s \in [0, t]$. It follows that

$$\mathbf{E}\int_{\mathcal{M}} d_p(Y, \partial\mathcal{M})\mathscr{M}_t([\tau \wedge t, \bar{\tau} \wedge t] \times dY) \le C\mathbf{E}|\bar{\tau} \wedge t - \tau \wedge t| \le C\delta ,$$

where we used the fact that if τ_δ is the first hitting time of 0 by a Brownian motion starting at δ, then $\mathbf{E}(\tau_\delta \wedge 1) \le C\delta$. Combining these bounds completes the proof. $\qquad\square$

3.4 Lack of convergence of the generators

One standard method to prove convergence of a sequence of Markov processes to a limiting process, once tightness has been established, is to show that the corresponding generators converge in a suitable sense. In our situation, one actually does *not* expect the generator of the approximate process to converge to that of the limiting process, when testing it on "nice" test functions. We first argue at a technical level why this is the case, before providing an intuitive explanation.

Inspired by [EK86, Daw93], we consider test functions of the form

$$F(\mu) = \exp(\langle \log f, \mu \rangle) , \tag{3.7}$$

where $f \colon \mathcal{M} \to \mathbf{R}_+$ is a sufficiently smooth function such that $f(x, v) = 1$ for $(x, v) \in \partial \mathcal{M}$. This boundary condition is required since elements $\mu \in \ell^p(\mathcal{M})$ can have infinite mass (and, as we have already seen, the limiting process really does acquire infinite mass at some exceptional times) accumulating near $\partial \mathcal{M}$. Being in $\ell^p(\mathcal{M})$ for $p \leq 1$ does however ensure that smooth functions such as above are integrable.

Exploiting the independence structure of the process as well as its space homogeneity, we can reduce ourselves to the case of an initial condition of the form $\mu_0 = \delta_{(x,v)}$ for some $v < -ax$. In this case, for sufficiently small $\varepsilon > 0$, the probability that the original particle dies within the time interval ε is of the order ε^p for any $p > 0$. We therefore only need to take into account the possibility of creating some descendant(s), with the killing mechanism being taken care of by the boundary condition of f. While the average number of "second generation" descendants is of order ε, any such descendant will typically have travelled to a distance of order $\sqrt{\varepsilon}$ from $\partial \mathcal{M}$, so that only the first generation has a chance of contributing to the generator.

We then have

$$\frac{1}{\varepsilon}(\mathbf{E}F(\mu_\varepsilon) - F(\mu_0)) \approx A_0 f(x, v) + \frac{f(x, v)}{\varepsilon} \mathbf{E}\langle f - 1, \mathscr{M}_\varepsilon \rangle ,$$

where

$$A_0 = \frac{1}{2}\partial_x^2 ,$$

is the generator of Brownian motion, and where $\mathscr{M}_\varepsilon$ is the (projection to time ε of the) Poisson point process yielding the first generation of offsprings. Note now that since these offspring will be created near $\partial \mathcal{M}$ and since $f = 1$ there, we can further approximate this expression by

$$\frac{1}{\varepsilon}(\mathbf{E}F(\mu_\varepsilon) - F(\mu_0)) \approx A_0 f(x, v) + \frac{f(x, v)}{\varepsilon} f'(x) \mathbf{E} \int_{\mathcal{M}} (\bar{x} + a^{-1}\bar{v}) \mathscr{M}_\varepsilon(d\bar{x}, d\bar{v}) .$$

Here, we wrote $f'(x)$ as a shortcut for $\partial_x f(x, v)|_{v=-ax}$. Denoting by e_s the position, relative from its starting point, of an excursion of length s and making use of the formula (2.25) for the intensity measure of $\mathscr{M}_\varepsilon$, we obtain for the last term in this equation the expression

$$\frac{1}{\varepsilon}\mathbf{E} \int_{\mathcal{M}} (\bar{x} + a^{-1}\bar{v}) \mathscr{M}_\varepsilon(d\bar{x}, d\bar{v}) = \frac{a}{2\sqrt{2\pi}\varepsilon} \int_0^\infty \int_0^{s \wedge \varepsilon} \mathbf{E}e_s(t)\, dt\, s^{-3/2}\, ds . \tag{3.8}$$

At this stage we note that, for a Brownian excursion of length s, we have for $t \leq s$ the identity

$$\mathbf{E}e_s(t) = \sqrt{\frac{8t}{\pi s}(s - t)}$$

which can be computed using the explicit formula given for the transition probabilities of the Brownian excursion on page 59 of [**?**]. Inserting this into (3.8), a tedious but straightforward calculation then yields

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathbf{E} \int_{\mathcal{M}} (\bar{x} + a^{-1}\bar{v}) \mathscr{M}_\varepsilon(d\bar{x}, d\bar{v}) = \frac{a}{2} \,,$$

so that we finally obtain for the generator A the expression

$$AF(\mu_0) = A_0 f(x, v) + \frac{a}{2} f(x, v) f'(x) \,. \tag{3.9}$$

Recall that this is for the particular case where $\mu_0 = \delta_{(x,v)}$. In the general case, we can use the independence structure of the process to obtain

$$AF(\mu_0) = F(\mu_0) \int_{\mathcal{M}} \left(\frac{A_0 f(x, v)}{f(x, v)} + \frac{a}{2} f'(x) \right) \mu_0(dx, dv) \,.$$

Remark 3.4 Compare this with the generator of a usual branching diffusion, where the term $a f'(x)/2$ would be replaced by $a(f(x) - 1)$, with a the branching rate.

On the other hand, if we denote by T_ε the Markov operator describing one step of Algorithm TDMC, we might expect that one also obtains A as the limit $\varepsilon^{-1}(T_\varepsilon - 1)$ as $\varepsilon \to 0$. This would indeed be the case if there was no branching or if branching did only occur at a finite rate. Considering again initial conditions of the form $\mu_0 = \delta_{(x,v)}$ for some $v < -ax$. By reasoning similar to that in the beginning of Section 2, we obtain

$$\varepsilon^{-1}(T_\varepsilon F - F)(\mu_0) \approx A_0 f(x, v) + \frac{a}{2} f(x, v) f'(x) \int_0^\infty y^2 \, \nu(dy) \,, \tag{3.10}$$

which is always *different* from (3.9), and is actually what we would have obtained from the wrong guess (2.7).

A possible reason for this discrepancy is that, while the Markov semigroup of the Brownian fan does indeed preserve test functions of the type (3.7), we do not expect this to be true of the Markov operator T_ε. Instead, the "correct" space of test functions for T_ε is of the same type, but the function f should have a "boundary layer" near $\partial \mathcal{M}$.

Remark 3.5 Another reason why the generator of the Brownian fan is not such a useful object is that many seemingly innocent observables, like for example the total number $\mathcal{N}$ of particles, do *not* belong to its domain. This follows from the fact that if we consider again a simple initial condition μ_0 as above, then $\mathbf{E}\mathcal{N}(\mu_\varepsilon) - 1 \approx \sqrt{\varepsilon}$ for small ε. The reason the total number of particles nevertheless remains finite (at least for fixed times) is that there are exceptional states where one or more particles are very near $\partial \mathcal{M}$ and for which $\mathbf{E}\mathcal{N}(\mu_\varepsilon) - 1 \approx -\mathcal{O}(1)$.

The remainder of the article is devoted to providing a rigorous proof of the fact that, in the situation of the previous two sections, the process given by Algorithm TDMC converges to the Brownian fan in $\mathcal{C}([0, T], \ell^p(\mathcal{M}))$ for any $T > 0$ and $p \le 1$. The overall strategy of the proof is classical: we first prove a tightness result in Section 4 and then show that finite-dimensional marginals converge to those of the Brownian fan in Section 5.

Difficulties arise on two fronts. First, to prove the tightness result, it is convenient to have uniform moment bounds on the number of particles at fixed time for the approximating system. These turn out to be much more difficult to obtain for the approximating

system than for the Brownian fan, which is mainly due to a lack of uniform exponential bounds. A second difficulty arises in the proof that finite-dimensional distributions converge to those of the Brownian fan. While it is intuitively clear that those excursions that survive for times of order $\mathcal{O}(1)$ do converge to suitably normalised Brownian excursions, this result is rather technical and, surprisingly, does not seem to appear in the literature. Furthermore, no convergence result holds for the typical excursions which die very early. We therefore also need to argue that, both at the level of Algorithm TDMC and at the level of the Brownian fan, these small excursions do not matter in the limit.

4   Tightness

As in the previous section, we restrict ourselves to the particular case when the underlying Markov process is given by a rescaled random walk, namely

$$y_{(k+1)\varepsilon} = y_{k\varepsilon} + \sqrt{\varepsilon}\xi_{k+1} \, , \tag{4.1}$$

where the ξ_k are i.i.d. random variables with distribution ν having some exponential moments, and where the potential V is given by a linear function, $V(x) = -ax$. Our aim is to show that as $\varepsilon \to 0$, the sequence of birth and death processes obtained by running Algorithm TDMC is tight in a state space $\mathcal{X}$, which we will now describe.

4.1   Formulation of the tightness result

With the construction of the previous section in mind, we choose as our state space $\mathcal{X} = \ell^p(\mathcal{M})$, where $\mathcal{M} = \{(x, v, n) \in \mathbf{R}^2 \times \mathbf{N} \ : \ v > -ax\}$, and $p \leq 1$ is arbitrary. Here, the coordinate n is used to keep track of the generation of a particle: direct offspring of a particle from the nth generation belong to the $(n + 1)$st generation. We extend the Euclidean distance to $\mathbf{R}^2 \times \mathbf{N}$ by additionally postulating that the distance between particles belonging to different generations is given by the sums of the distances of the two particles to $\partial\mathcal{M}$. The boundary $\partial\mathcal{M}$ is given as before by $\partial\mathcal{M} = \{(x, v, n) \ : \ v = -ax\}$. We will use capital letters for elements of $\mathcal{M}$ to differentiate them from elements of $\mathbf{R}$.

In order to formulate our result, we will make use of the following notation. For $t = k\varepsilon$ with k an integer, we denote by μ_t^ε the empirical measure of the particles alive at time t, and by N_t the number of such particles. Sometimes, it will be convenient to consider the particles instead as a collection of elements $X_t^{(j)} \in \mathcal{M}$, so that we write

$$\mu_t^\varepsilon = \sum_{j=1}^{N_t} \delta_{X_t^{(j)}} \, .$$

We do not specify how exactly we order the particles, as this is completely irrelevant for our purpose. For $t \in (k\varepsilon, (k + 1)\varepsilon)$, we define μ_t^ε by using the "linear interpolation" procedure (3.4), setting

$$\mu_t^\varepsilon = L_s(\mu_{k\varepsilon}^\varepsilon, \mu_{(k+1)\varepsilon}^\varepsilon) \, , \qquad s = \varepsilon^{-1}t - k \, .$$

The interpolation procedure L_s is a very minor modification from the one described above, in the sense that we only connect particles belonging to the same generation. In this way, the process $t \mapsto \mu_t^\varepsilon$ has continuous trajectories for every ε. The main result of this section is as follows:

Theorem 4.1 *Let $p \leq 1$ and denote by $\mathcal{L}^\varepsilon$ the law of the process $t \mapsto \mu_t^\varepsilon$ described above, viewed as a family of probability measures on $\mathcal{C}([0,1], \mathcal{X})$. Assume furthermore that there exists $c > 0$ such that $\int e^{c|y|}\nu(dy) < \infty$.*

Then, for any single particle initial condition $\mu_0^\varepsilon = \delta_{X_0}$ with $X_0 \in \mathcal{M}$, there exists $\varepsilon_0 > 0$ such that the family $\{\mathcal{L}^\varepsilon\}_{\varepsilon \leq \varepsilon_0}$ is tight.

Proof. Combining [Daw93, Theorem 3.6.4] and [Bil99, Theorem 8.3], we see that, in order to obtain tightness, it is sufficient to show that:

- For every $\delta > 0$, there exists a compact set $K_\delta \subset \mathcal{X}$ such that $\mathbf{P}(\mu_t^\varepsilon \in K_\delta) > 1 - \delta$, uniformly over $t \in [0,1]$ and $\varepsilon < \varepsilon_0$.

- There exists $\alpha > 0$ and $C > 0$ such that $\mathbf{E}\|\mu_t^\varepsilon - \mu_s^\varepsilon\|_p^q \leq C|t - s|^{\alpha q}$, uniformly over all $s, t \in [0,1]$ and $\varepsilon < \varepsilon_0$.

The first claim then follows from Proposition 4.7 below, while the second claim is the content of Proposition 4.8. □

The proof of this result is the content of the remainder of this subsection and goes roughly as follows. In Section 4.2, we obtain a moment bound on the number of particles alive at any fixed time $t \in [0,1]$, which is uniform in $\varepsilon > 0$. This then allows to obtain the compactness at fixed time in Section 4.3. The verification of Kolmogorov's continuity criterion is the content of Section 4.4.

4.2 Moment bounds on the number of particles

In the sequel, we will denote by $\mathbf{K}_p(X)$ the pth cumulant of a random variable X and by $\mathbf{K}_p(X \mid \mathcal{F})$ the same cumulant, conditioned on the σ-field $\mathcal{F}$. We will use the important property that $\mathbf{K}_p(X + Y \mid \mathcal{F}) = \mathbf{K}_p(X \mid \mathcal{F}) + \mathbf{K}_p(Y \mid \mathcal{F})$, provided that X and Y are independent, conditionally on $\mathcal{F}$. We will also use the fact that if X is a positive random variable, then $\mathbf{K}_p(X) \leq \mathbf{E}X^p$ and there exists a constant C such that the bound

$$\mathbf{E}X^p \leq C \sum_{q \leq p} (\mathbf{K}_q(X))^{p/q} , \tag{4.2}$$

holds.

Our aim now is to obtain a bound on the cumulants of the number of particles alive at time t which is independent of ε. We start with an initial configuration containing only one particle, which belongs to generation 0, and we set $N_t^0 \in \{0, 1\}$, depending on whether or not this particle is still alive at some subsequent time t. We also define $N_t^n = \mu_t^\varepsilon(\mathcal{M}^n)$, where $\mathcal{M}^n = \{(x, v, \ell) \in \mathcal{M} : \ell = n\}$, which is the number of particles in the nth generation that are alive at time t. We also denote by $N_{s,t}^n$ the number of such particles that were created at time $s \leq t$.

For any $X_0 = (x, v)$ with $v > -ax$, we write $\mathbf{E}_{X_0}$ for expectations of observables for the process generated by starting Algorithm TDMC with underlying dynamic (4.1), started with one single initial particle in generation 0 at location x with tag v. Finally, for $x \in \mathbf{R}$, we write $\mathbf{E}_x$ for the same expectation, but where the initial particle has tag $v = +\infty$, meaning that it is "immortal". We then have the following result:

Proposition 4.2 *Consider the situation of Theorem 4.1. For every $p \geq 0$, there exist $C_p > 0$ and $\varepsilon_0 > 0$ such that, under the rules of Algorithm TDMC, the number $N_t = \mu_t^\varepsilon(\mathcal{M})$ of particles alive at time t satisfies $\mathbf{E}_x|N_t|^p \leq 2\exp(C_p t)$, uniformly over all $\varepsilon \leq \varepsilon_0$ and $x \in \mathbf{R}$. Furthermore, there exists $\varrho > 0$ such that*

$$\mathbf{E}_x N_t^n \leq (n+1)^{t/\varrho} 2^{-n} , \tag{4.3}$$

uniformly over all $\varepsilon \le \varepsilon_0$, $n \ge 0$, and $t > 0$.

Finally, denoting by R_t^γ the number of offspring alive at time t that have never been at distance more than γ from $\partial \mathcal{M}$, we have the bound

$$\mathbf{E}_x R_t^\gamma \le C\gamma \,,$$

uniformly over $t \in [0, 1]$, $\gamma \in (0, 1]$, and $\varepsilon \le \varepsilon_0 \wedge \gamma^2$.

The proof will make use of the following elementary fact where, for $\lambda = n + p$ with $n \in \mathbf{N}$ and $p \in (0, 1]$, we denote by $\mathcal{I}(\lambda)$ the law of a random variable Y such that $Y = n$ with probability $1 - p$ and $Y = n + 1$ with probability p. (This is so that $\mathbf{E}Y = \lambda$.)

Lemma 4.3 *Let Y be a random variable with law $\mathcal{I}(\lambda)$. Then, for any $q \ge 1$ and any $\lambda > 0$, one has the bound $\mathbf{E}Y^q \le \lambda + (2\lambda)^q$.*

Proof. By inspection, one has

$$\mathbf{E}Y^q = p(n + 1)^q + (1 - p)n^q \,.$$

In the case $n = 0$, one then has $\mathbf{E}Y^q = \lambda$, so the statement is true. For $n \ge 1$, one uses the fact that both $n + 1$ and n are bounded by 2λ, and the claim follows at once. $\square$

Proof of Proposition 4.2. We restrict ourselves to times that are integer multiples of ε. Furthermore, from now on, we fix an initial condition x, so that we just write $\mathbf{E}$ instead of $\mathbf{E}_x$. We also denote by $\mathscr{F}^n$ the σ-algebra containing all information pertaining to particles in generations up to (and including) n. With this notation at hand, we obtain for N_t^n the bound

$$\mathbf{E}(N_t^n)^p \lesssim \mathbf{E} \sum_{q=1}^p (\mathbf{K}_q(N_t^n \mid \mathscr{F}^{n-1}))^{p/q}$$

$$= \mathbf{E} \sum_{q=1}^p \Big(\sum_{\varepsilon\ell \le t} \mathbf{K}_q(N_{\varepsilon\ell,t}^n \mid \mathscr{F}^{n-1}) \Big)^{p/q}$$

$$\lesssim \sum_{q=1}^p \mathbf{E} \Big(\sum_{\varepsilon\ell \le t} \mathbf{E}(|N_{\varepsilon\ell,t}^n|^q \mid \mathscr{F}^{n-1}) \Big)^{p/q} \,, \tag{4.4}$$

where we used (4.2) in the first step and the independence of the offspring in the second step. In the above expression, ℓ takes only integer values. Note now that

$$\sum_{\varepsilon\ell \le t} \frac{\varepsilon}{\sqrt{t(t + \varepsilon - \varepsilon\ell)}} \le C \,, \tag{4.5}$$

uniformly over all $t \ge \varepsilon$. As a consequence, for any positive sequence a_n and any power $r \ge 1$, one has the bound

$$\Big(\sum_{\varepsilon\ell \le t} a_\ell \Big)^r = \Big(\sum_{\varepsilon\ell \le t} \frac{\varepsilon}{\sqrt{t(t + \varepsilon - \varepsilon\ell)}} \tilde{a}_\ell \Big)^r \lesssim \sum_{\varepsilon\ell \le t} \frac{\varepsilon}{\sqrt{t(t + \varepsilon - \varepsilon\ell)}} \tilde{a}_\ell^r$$

$$= \sum_{\varepsilon\ell \le t} \frac{t^{\frac{r-1}{2}}(t + \varepsilon - \varepsilon\ell)^{\frac{r-1}{2}}}{\varepsilon^{r-1}} a_\ell^r \,,$$

where we have set $\tilde{a}_\ell = \varepsilon^{-1}\sqrt{t(t + \varepsilon - \varepsilon\ell)}\, a_\ell$ in the intermediate steps. Applying this inequality to (4.4), we obtain

$$\mathbf{E}(N_t^n)^p \lesssim \sum_{q=1}^{p} \sum_{\varepsilon\ell \leq t} \frac{t^{\frac{p-q}{2q}}(t + \varepsilon - \varepsilon\ell)^{\frac{p-q}{2q}}}{\varepsilon^{\frac{p-q}{q}}} \mathbf{E}\Big(\mathbf{E}(|N_{\varepsilon\ell,t}^n|^q \,|\, \mathscr{F}^{n-1})\Big)^{p/q}. \qquad (4.6)$$

Denote now by $M_s^{n,j}$ the number of particles in the nth generation created at time s by the jth particle from the $n-1$st generation. We write $\mathscr{G}_s$ for the σ-algebra generated by this additional data. Each of these particles yields a contribution to $N_{s,t}^n$ of either 1 or 0, depending whether it survives or not. Furthermore, these contributions, which we will denote by $S_{s,t}^{n,j,i}$, are all independent and, for the same value of j, they are also identically distributed. By definition, we thus have the identity

$$N_{\varepsilon\ell,t}^n = \sum_{j=1}^{N_{\varepsilon(\ell-1)}^{n-1}} \sum_{i=1}^{M_{\varepsilon\ell}^{n,j}} S_{\varepsilon\ell,t}^{n,j,i}.$$

We now twice make use of the inequality

$$\Big(\sum_{j=1}^{m} a_j\Big)^q \leq m^{q-1} \sum_{j=1}^{m} a_j^q, \qquad (4.7)$$

which is valid for any $q \geq 1$, $m \geq 0$ and sequence of positive numbers a_j. This yields the bound

$$(N_{\varepsilon\ell,t}^n)^q \leq (N_{(\ell-1)\varepsilon}^{n-1})^{q-1} \sum_{j=1}^{N_{(\ell-1)\varepsilon}^{n-1}} (M_{\varepsilon\ell}^{n,j})^{q-1} \sum_{i=1}^{M_{\varepsilon\ell}^{n,j}} S_{\varepsilon\ell,t}^{n,j,i}. \qquad (4.8)$$

Note that since $S_{\varepsilon\ell,t}^{n,j,i}$ can only take the values 0 or 1, raising it to the power q makes no difference. By the "gambler's ruin theorem" [LL10, Thm 5.1.7], we have the bound

$$\mathbf{E}(S_{\varepsilon\ell,t}^{n,j,i} \,|\, \mathscr{F}^{n-1} \vee \mathscr{G}_{\varepsilon\ell}) \leq C\sqrt{\varepsilon}\, \frac{\xi^j + 1}{\sqrt{t + \varepsilon - \ell\varepsilon}}, \qquad (4.9)$$

where $\sqrt{\varepsilon}\xi^j$ denotes the step performed by the jth particle of the $(n-1)$st generation between times $\varepsilon(\ell-1)$ and $\varepsilon\ell$. Regarding the number of offspring $M_{\varepsilon\ell}^{n,j}$, it follows from the definition of the algorithm that its distribution is given by $\mathcal{I}(\exp(a\sqrt{\varepsilon}\xi_+^j) - 1)$, where ξ_+^j denotes the positive part of ξ^j. Combining this with (4.8), Lemma 4.3, and (4.9), we thus obtain the bound

$$\mathbf{E}(|N_{\varepsilon\ell,t}^n|^q \,|\, \mathscr{F}^{n-1}) \lesssim |N_{(\ell-1)\varepsilon}^{n-1}|^{q-1} \sum_{j=1}^{N_{(\ell-1)\varepsilon}^{n-1}} \sqrt{\varepsilon}\, \frac{(e^{a\sqrt{\varepsilon}\xi_+^j} - 1) + (e^{a\sqrt{\varepsilon}\xi_+^j} - 1)^q}{\sqrt{t + \varepsilon - \varepsilon\ell}}(\xi^j + 1).$$

In order to simplify this expression, we use the fact that, for $x \geq 0$, there exists a constant C depending on q such that

$$e^x - 1 \leq xe^x, \qquad (e^x - 1)^q \leq Cxe^{qx},$$

for every $q \geq 1$. This yields

$$\mathbf{E}(|N_{\varepsilon\ell,t}^n|^q \,|\, \mathscr{F}^{n-1}) \lesssim |N_{(\ell-1)\varepsilon}^{n-1}|^{q-1} \sum_{j=1}^{N_{(\ell-1)\varepsilon}^{n-1}} \varepsilon\, \frac{\mathbf{1}_{\xi^j \geq 0}\, e^{aq\sqrt{\varepsilon}\xi^j}}{\sqrt{t + \varepsilon - \varepsilon\ell}}(\xi^j + 1)^2.$$

Using (4.7) once again, we get the bound

$$(\mathbf{E}(|N_{\varepsilon\ell,t}^n|^q \mid \mathscr{F}^{n-1}))^{\frac{p}{q}} \lesssim |N_{(\ell-1)\varepsilon}^{n-1}|^{p-1} \sum_{j=1}^{N_{(\ell-1)\varepsilon}^{n-1}} \varepsilon^{\frac{p}{q}} \frac{\mathbf{1}_{\xi^j \geq 0} e^{ap\sqrt{\varepsilon}\xi^j}}{(t+\varepsilon-\varepsilon\ell)^{\frac{p}{2q}}} (\xi^j+1)^{\frac{2p}{q}} \, .$$

At this stage, we note that, conditional on the state of the system at time $\varepsilon(\ell-1)$ the steps ξ^j are all independent and identically distributed with law ν. Setting

$$P_\varepsilon = \int_0^\infty e^{ap\sqrt{\varepsilon}z}(z+1)^{\frac{2p}{q}}\nu(dz) \, ,$$

it follows that

$$\mathbf{E}(\mathbf{E}(|N_{\varepsilon\ell,t}^n|^q \mid \mathscr{F}^{n-1}))^{\frac{p}{q}} \lesssim \frac{\varepsilon^{\frac{p}{q}}P_\varepsilon}{(t+\varepsilon-\varepsilon\ell)^{\frac{p}{2q}}} \mathbf{E}|N_{(\ell-1)\varepsilon}^{n-1}|^p \, .$$

Inserting this into (4.6) yields for $t \leq 1$ the bound

$$\mathbf{E}(N_t^n)^p \lesssim P_\varepsilon \sum_{\varepsilon\ell\leq t} \frac{\varepsilon}{\sqrt{t+\varepsilon-\varepsilon\ell}} \mathbf{E}|N_{(\ell-1)\varepsilon}^{n-1}|^p \, .$$

On the other hand, since we assumed that the initial particle is immortal, N_s^{n-1} is stochastically increasing as a function of s, so that we obtain from (4.5) the bound

$$\mathbf{E}(N_t^n)^p \lesssim P_\varepsilon \sqrt{t}\, \mathbf{E}(N_t^{n-1})^p \, .$$

Since, by the exponential integrability assumption on ν, there exists $C > 0$ such that, for ε small enough, $P_\varepsilon \leq C$ and since $N_t^0 = 1$, we conclude that there exists some λ_p such that, for all $n \geq 1$, one has the bound

$$\mathbf{E}(N_t^n)^p \lesssim (\lambda_p t)^{n/2} \, . \tag{4.10}$$

It follows that there exists $t_\star > 0$ (depending on p) such that $\mathbf{E}(N_t)^p \leq 2$ (say), for every $t \leq t_\star$. To show that $\mathbf{E}(N_t)^p < \infty$ for every t, we use again (4.7), combined with the Markov property of the process to conclude that $\mathbf{E}(N_{t+t_\star})^p \leq 2\mathbf{E}(N_t)^p$, from which the first claim follows.

To get (4.3), observe that if we choose ϱ small enough so that $\lambda_1\varrho \leq 1/4$, the bound with $t \leq \varrho$ follows from (4.10). Denote now by $\mathcal{F}_t$ the filtration generated by all events up to time t. For t and ϱ that are multiples of ε it follows from the Markov property that

$$\mathbf{E}(N_t^n \mid \mathcal{F}_{t-\varrho}) \leq \sum_{\ell=0}^n N_{t-\varrho}^\ell \, \mathbf{E}N_\varrho^{n-\ell}$$

where the expectation on the right is taken with respect to an initial condition with one single immortal particle. The claimed bound for arbitrary $t > 0$ therefore follows by induction.

It remains to obtain the bound on R_t^γ. Denote by $R_{s,t}^\gamma$ the number of offspring contributing to R_t^γ that are created at time $s < t$, so that $R_t^\gamma = \sum_{\varepsilon k \leq t} R_{k\varepsilon,t}^\gamma$. Denote now by $Q_{z,t}^\gamma$ the probability that, after time t, the random walk (4.1) with initial condition $\sqrt{\varepsilon}z$ has never exited the interval $[0, \gamma]$. It then follows that

$$\mathbf{E}(R_{k\varepsilon,t}^\gamma \mid \mathcal{F}_{k\varepsilon}) = a\sqrt{\varepsilon}N_{k\varepsilon} \int_0^\infty \int_0^z e^{a\sqrt{\varepsilon}z}Q_{y,t-k\varepsilon}^\gamma \, dy \, \nu(dz) \, . \tag{4.11}$$

This is because, by Step 3 of Algorithm TDMC, the expected number of offspring created by a particle performing an upward step of size $\sqrt{\varepsilon}z$ is given by $e^{a\sqrt{\varepsilon}z} - 1$ while, if we denote by $\sqrt{\varepsilon}y$ the distance between the starting point of the offspring and the "wall" below which it is killed, the law of y is given by $c^{-1} \int_0^z e^{a\sqrt{\varepsilon}y}\,dy$ where c is a normalizing constant. Since $c = (e^{a\sqrt{\varepsilon}z} - 1)/a\sqrt{\varepsilon}$, (4.11) follows.

Using again the gambler's ruin theorem, combined with the Markov property of the random walk, we obtain for $Q_{z,t}^{\gamma}$ the bound

$$Q_{z,t}^{\gamma} \leq C \frac{(z+1)\sqrt{\varepsilon}}{\sqrt{t+\varepsilon}} \bar{Q}_{t/2}^{\gamma}\,,$$

where $\bar{Q}_t^{\gamma}$ is the probability that the random walk (4.1), starting at the origin, stays within $[-\gamma, \gamma]$ up to time t. Using the scaling properties of Brownian motion, combined with [Fra73], we conclude that one has the bound

$$\bar{Q}_t^{\gamma} \leq 1 \wedge \frac{C\delta^{2q}}{t^q}\,,$$

for every $q > 0$, so that

$$Q_{z,t}^{\gamma} \lesssim \frac{(z+1)\sqrt{\varepsilon}}{\sqrt{t+\varepsilon}} \left(1 \wedge \frac{\delta^{2q}}{t^q}\right)\,.$$

Inserting this bound into (4.11), using the previously obtained bounds on $N_{k\varepsilon}$, and summing over all k, the claim follows at once. $\square$

Remark 4.4 It follows from Theorem 4.1 in [HW14] that, in the particular case when the initial tag v is distributed according to the logarithm of a uniformly distributed random variable, one has the identity $\mathbf{E}N_t = \mathbf{E}e^{ay_t}$, where y_t is the rescaled random walk with steps ν. The (at least one-sided) exponential integrability assumption on ν is therefore a necessary assumption in order to obtain any kind of moment bounds on N_t.

Remark 4.5 Even if we assume that ν has Gaussian tails and despite the result previously obtained for the Brownian fan in Theorem 2.14, it is *not* true in general that N_t defined by Algorithm TDMC, has uniform exponential moments as $\varepsilon \to 0$. This is because even for the first step, the probability that the original particle performs a step of order ε^{-p} is of order $\exp(-\varepsilon^{-2p-1})$. If this were to happen, the number of offspring created in this way would be of order $\exp(\varepsilon^{-p})$, which immediately shows that exponential moments blow up as $\varepsilon \to 0$.

4.3   Compactness at any fixed time

We now show that we can find a compact set K_δ such that μ_t^ε belongs to K_δ with high probability, uniformly over ε and $t \in [0, 1]$. Our first ingredient for this is the following moment bound:

Proposition 4.6 *Consider the setting of Theorem 4.1. Then, there exist constants C and ε_0 such that, for every $t \in [0, 1]$ and every $\varepsilon \leq \varepsilon_0$, the bound*

$$\mathbf{E}_{X_0}\left(\int_{\mathcal{M}} |x - x_0|^{2p}\mu_t^\varepsilon(dx, dv, dk)\right) \leq Ct^p\,, \tag{4.12}$$

holds uniformly over all initial conditions $X_0 = (x_0, v_0, 0) \in \mathcal{M}$.

Proof. We restrict ourselves to the case when t is an integer multiple of ε, since the bound on the remaining times easily follows from our interpolation procedure. Furthermore, we can restrict ourselves to the case when the initial particle is immortal, which formally corresponds to setting $v = +\infty$. By translation invariance, we also restrict ourselves to the case where the initial particle is located at the origin, and we denote the corresponding expectation by $\mathbf{E}_0$.

It follows from Theorem 4.1 in [HW14] that if we choose $v = \frac{1}{a} \log u$, where u is drawn uniformly from $[0, 1]$ and denote by $\tilde{\mu}_t^\varepsilon$ the corresponding process, then one has the identity

$$\mathbf{E}_0\Big(\int_{\mathcal{M}} x^{2p} \tilde{\mu}_t^\varepsilon(dx, dv, dk)\Big) = \mathbf{E}_0\Big(e^{ay_t}(y_t)^{2p}\Big),$$

where y_t is the simple random walk (4.1) started at the origin. It follows immediately from the exponential integrability of ν that this quantity is bounded by Ct^p for $t \leq 1$ and for ε small enough.

On the other hand, one can realise the process μ_t^ε, which starts with an immortal initial particle, in the following way:

1. Consider the process $\tilde{\mu}_t^\varepsilon$, where the v-component is as above.

2. When the initial (generation 0) particle is killed, replace it instantly by an immortal particle starting at the current location.

Let x_t^0 denote the trajectory of the initial particle and let s be the time at which the initial particle is killed and replaced. Let $P_t = \mathbf{P}(s \leq t)$. By the construction just outlined, we have the recursion relation

$$F_{t,0} = E_t + P_t \mathbf{E}(F_{t-s, x_s^0} \mid s \leq t),$$

where we set

$$F_{t, x_s^0} = \mathbf{E}_0\Big(\int_{\mathcal{M}} (x + x_s^0)^{2p} \mu_t^\varepsilon(dx, dv, dk)\Big),$$

$$E_t = \mathbf{E}_0\Big(\int_{\mathcal{M}} x^{2p} \tilde{\mu}_t^\varepsilon(dx, dv, dk)\Big).$$

(Remember that the difference between E_t and $F_{t,0}$ is that in order to compute F, we start with an immortal particle.) Setting $F_t = F_{t,0}$, using the fact that, for every $\delta > 0$ one can find C_δ such that $(x + x_s^0)^{2p} \leq C_\delta (x_s^0)^{2p} + (1 + \delta)x^{2p}$, and recalling that $E_t \leq Ct^p$, we deduce that

$$F_t \leq Ct^p + C_\delta P_t \mathbf{E}((x_s^0)^{2p} \mathbf{E}_{x_s^0}(N_{t-s}) \mid s \leq t)) + (1 + \delta)P_t \mathbf{E}(F_{t-s} \mid s \leq t),$$

where N_t is the number of particles alive at time t for the system started with an immortal particle. Note now that, for $t \leq 1$, we know from Proposition 4.2 that the expected number of particles alive at any given time is bounded by some constant uniform in ε. Since this bound is also uniform in the initial condition, we have

$$P_t \mathbf{E}((x_s^0)^{2p} \mathbf{E}_{x_s^0}(N_{t-s} \mid s \leq t)) \lesssim P_t \mathbf{E}((x_s^0)^{2p} \mid s \leq t) = \mathbf{E}((x_s^0)^{2p} \mathbf{1}_{s \leq t}).$$

Defining $\tilde{s} = s \wedge t$, we then obtain the trivial bound

$$\mathbf{E}((x_s^0)^{2p} \mathbf{1}_{s \leq t}) \leq \mathbf{E}_0 |y_{\tilde{s}}|^{2p} \leq \mathbf{E}_0 |y_t|^{2p} \lesssim t^p,$$

where we made use of the fact that $|y_t|^{2p}$ is a submartingale to obtain the second inequality.

Setting now $\bar{F}_t = \sup_{s \le t} F_s$, we can combine all of these bounds to get the inequality

$$\bar{F}_t \le C_\delta t^p + (1 + \delta) P_t \bar{F}_t \;.$$

Since one can check (for example by again using the fact that the random walk approximates a Brownian motion for small ε) that $\sup_{\varepsilon \le 1} \sup_{t \le 1} P_t = \sup_{\varepsilon \le 1} P_1 < 1$, the claim follows at once by choosing δ sufficiently small. $\square$

This result can now be used to deduce the announced uniform tightness result over fixed times:

Proposition 4.7 *Consider the setting of Theorem 4.1. Then, for every $\delta > 0$ there exists a compact set $K_\delta \subset \mathcal{X}$ such that $\mathbf{P}_{X_0}(\mu_t^\varepsilon \in K_\delta) > 1 - \delta$, uniformly over $t \in [0, 1]$, $\varepsilon \le \varepsilon_0$ and $X_0 \in \mathcal{M}^0$.*

Proof. For any $n \in \mathbf{Z}_+$ and $R \in \mathbf{R}_+$, denote

$$\mathcal{M}^{[n]} = \{(x, v, \ell) \in \mathcal{M} \,:\, \ell \le n\} \;, \quad B_R = \{(x, v, \ell) \in \mathcal{M} \,:\, |x| \vee |v| \le R\} \;.$$

For any such n and R and for $m \in \mathbf{Z}_+$, we then denote by $K_{n,m,R} \subset \mathcal{X}$ the set of all integer-valued measures η on $\mathcal{M}$ such that $\eta(\mathcal{M} \setminus \mathcal{M}^{[n]}) = 0$, $\eta(\mathcal{M}) \le m$, and $\eta(\mathcal{M} \setminus B_R) = 0$. Since these sets are obviously compact, it suffices to find, for every $\delta > 0$, sufficiently large values n, R and m so that $\mathbf{P}_{X_0}(\mu_t^\varepsilon \in K_{n,m,R}) > 1 - \delta$ uniformly over all $\varepsilon \le \varepsilon_0$ and $t \in [0, 1]$.

Note that $K_{n,m,R} \subset K_n^1 \cap K_m^2 \cap K_R^x \cap K_R^v$ where, for K_n^1 and K_m^2 we only enforce the conditions involving n and m respectively. The set K_R^x consists of configurations such that the x-coordinate of every particle is less than R in absolute value, while K_R^v enforces that the v-coordinate be less than R.

It follows immediately from (4.3) that there exists $\gamma > 0$ and a constant C (depending in principle on the time we consider, but it can be chosen uniformly over $t \in [0, 1]$), such that

$$\mathbf{P}_{X_0}(\mu_t^\varepsilon \notin K_n^1) \le C e^{-\gamma n} \;,$$

for every $n \ge 0$. Similarly, it follows from the moment bounds on N_t obtained in Proposition 4.2 that, for every $p > 0$, there exists C such that

$$\mathbf{P}_{X_0}(\mu_t^\varepsilon \notin K_m^2) \le \frac{C}{m^p} \;.$$

In order to get a bound on the x-coordinate of the particles, we combine Proposition 4.6 with Chebyshev's inequality, so that

$$\mathbf{P}_{X_0}(\mu_t^\varepsilon \notin K_R^x) \le \frac{C}{R^2} \;.$$

It remains to obtain a bound on the probability of not being in K_R^v. For this, we use the fact that on the one hand, the label of a particle always satisfies $v > -ax$. On the other hand, any descendant of the initial particle will always satisfy $v \le -a \sup_{t \le 1} x_t^0$, where here we denote by x_t^0 the position of the original particle at time t. Since this particle was assumed to start at the origin, we obtain $v^2 \le a^2 x^2 + a^2 \sup_{t \le 1}(x_t^0)^2$, so that we obtain the bound

$$\mathbf{P}_{X_0}(\mu_t^\varepsilon \notin K_R^v) \le \frac{C}{R^2} \;,$$

just as above. Combining all of these bounds, the claim follows by choosing n, R and m large enough. $\square$

4.4 Kolmogorov criterion

The aim of this section is to obtain the following bound on the time regularity of our process:

Proposition 4.8 *For every $p \le 1$ and $q \ge 1$, there exists a constant C such that*

$$\mathbf{E}_{X_0}\|\mu_\delta^\varepsilon - \delta_{X_0}\|_p^q \le C\delta^{pq/2} \,, \tag{4.13}$$

where $X_0 \in \mathcal{M}^0$ is an initial condition with only one particle in the system, and $\delta < 1$ is an integer multiple of ε.

Remark 4.9 As usual, the precise location of the particle is irrelevant by translation invariance, so the above bound is uniform over all choices of X_0.

Before we proceed to the proof of Proposition 4.8, we observe that the bound (4.13) implies that Kolmogorov's criterion holds for the process over a fixed interval of time:

Corollary 4.10 *For every $p \le 1$ and $q \ge 1$, one has the bound*

$$\mathbf{E}_{X_0}\|\mu_{t+\delta}^\varepsilon - \mu_t^\varepsilon\|_p^q \le C\delta^{pq/2} \,, \tag{4.14}$$

uniformly over all $\delta, t, \varepsilon \in [0, 1]$.

Proof. Note first that we can restrict ourselves to the case when t and δ are integer multiples of ε. Indeed, it follows from the definition of $\|\cdot\|_p$ and from (3.5) that, if $k\varepsilon \le s < t \le (k+1)\varepsilon$, then

$$\|\mu_s^\varepsilon - \mu_t^\varepsilon\| \le \varepsilon^{-p}|t - s|^p\|\mu_{k\varepsilon}^\varepsilon - \mu_{(k+1)\varepsilon}^\varepsilon\| \,.$$

We then obtain from Proposition 4.8 the bound

$$
\begin{aligned}
\mathbf{E}_{X_0}\|\mu_{t+\delta}^\varepsilon - \mu_t^\varepsilon\|_p^q &= \mathbf{E}_{X_0}\Big(\mathbf{E}_{X_0}(\|\mu_{t+\delta}^\varepsilon - \mu_t^\varepsilon\|_p^q \,|\, \mathcal{F}_t)\Big) \\
&\le \mathbf{E}_{X_0}\Big(|N_t|^{q-1}\sum_{j=1}^{N_t} \mathbf{E}_{X_t^{(j)}}\|\mu_\delta^\varepsilon - \delta_{X_t^{(j)}}\|_p^q\Big) \\
&\le C\delta^{pq/2}\mathbf{E}_{X_0}(|N_t|^q) \le C\delta^{pq/2} \,,
\end{aligned}
$$

where we made use of the Markov property, (4.7), and (4.13). $\qquad\square$

Proof of Proposition 4.8. Denote by X_t the location at time t of a single particle starting at $X_0 = (x_0, v_0, 0)$ and evolving under the rescaled random walk stopped when it reaches the boundary of $\mathcal{M}$. Denote by x_t the position in $\mathbf{R}$ corresponding to X_t. From the properties of the random walk and the definition of $\|\cdot\|_p$, for every $q > 1$ there exists a constant C such that the bound

$$\mathbf{E}_{X_0}\|\delta_{X_\delta} - \delta_{X_0}\|_p^q \le C\delta^{qp/2} \,,$$

holds, independently of the initial condition and independently of $\varepsilon \le 1$.

Let us now bound the contribution from the descendants of the initial particle. Ordering the particles alive at time t in such a way that the original particle has label 1

(if it is not alive anymore, we consider it as being located on the boundary, where it was stopped), we have the bound

$$\|\delta_{X_\delta} - \mu_\delta^\varepsilon\|_p \le \sum_{j=2}^{N_\delta} d_p(X_\delta^{(j)}, \partial\mathcal{M}) \,,$$

so that

$$\begin{aligned}
\mathbf{E}_{X_0}\|\delta_{X_\delta} - \mu_\delta^\varepsilon\|_p^q &\le \mathbf{E}_{X_0}\Big(\sum_{j=2}^{N_\delta} d_p(X_\delta^{(j)}, \partial\mathcal{M})\Big)^q \\
&\le \Big(\Big(\mathbf{E}_{X_0}(N_\delta - 1)^{2q-1}\Big)\Big(\mathbf{E}_{X_0}\sum_{j=2}^{N_\delta} d_p^{2q}(X_\delta^{(j)}, \partial\mathcal{M})\Big)\Big)^{1/2} \\
&\le C\Big(\mathbf{E}_{X_0}\sum_{j=2}^{N_\delta} d_p^{2q}(X_\delta^{(j)}, \partial\mathcal{M})\Big)^{1/2} \,,
\end{aligned}$$

where the second inequality follows from Proposition 4.2. Recall now that if $X_\delta^{(j)} = (x, v, n)$, then one has $d_p(X_\delta^{(j)}, \partial\mathcal{M}) = |x - v/a|^p$, where v/a is guaranteed to take values between $\inf_{s\le\delta} x_s$ and x. As a consequence, we have the bound

$$|x - v/a|^p \le |x - \inf_{s\le\delta} x_s|^p \le |x - x_0|^p + \sup_{s\le\delta}|x_s - x_0|^p \,,$$

so that

$$\begin{aligned}
\mathbf{E}_{X_0}\sum_{j=2}^{N_\delta} d_p^{2q}(X_\delta^{(j)}, \partial\mathcal{M}) &\lesssim \mathbf{E}_{X_0}\Big(N_\delta \sup_{s\le\delta}|x_s - x_0|^{2pq}\Big) \\
&\quad + \mathbf{E}_{X_0}\int|x - x_0|^{2pq}\mu_\delta^\varepsilon(dx, dv, dk) \lesssim \delta^{pq} \,,
\end{aligned}$$

where the last bound is a consequence of Propositions 4.2 and 4.6, as well as standard bounds on the supremum of a random walk. The claim now follows at once. $\qquad\square$

5 Convergence of fixed-time distributions

In this section, we show that any limiting process obtained from the tightness result of the previous section necessarily coincides with the Brownian fan constructed in Section 3.2. With the notations of that section at hand, our convergence result can be formulated as follows.

Theorem 5.1 *Consider the setting of Theorem 4.1 with an initial condition $X_0 = (x, v, 0) \in \mathcal{M}$, and consider μ_t as above with the "initial condition" for $\mu^{[\infty]}$ given by a Brownian motion starting at x, killed when it reaches the level $-v/a$.*

Then, for every $p \in (0, 1]$, there exists a version of the process $\{\mu_t\}_{t\ge0}$ which is a continuous Markov process with values in $\ell^p(\mathcal{M})$. Furthermore, denoting the law of its restriction to the time interval $[0, 1]$ by $\mathcal{L}$, the sequence of measures $\mathcal{L}^\varepsilon$ converges weakly to $\mathcal{L}$ in $\mathcal{C}([0, 1], \ell^p(\mathcal{M}))$.

Proof. It suffices to show that, for any fixed collection of times $\{t_1, \ldots, t_k\}$, the law of $\{\mu_{t_i}^\varepsilon\}_{i \le k}$ converges weakly to that of $\{\mu_{t_i}\}_{i \le k}$. Indeed, Corollary 4.10 then implies that the process μ_t satisfies Kolmogorov's continuity criterion and therefore has a continuous version. By Theorem 4.1, we deduce weak convergence in $\mathcal{C}([0,1], \ell^p(\mathcal{M}))$ from the convergence of marginals. Using the Markov property, the superposition property of the process, and Proposition 4.2, we reduce ourselves to the case $k = 1$ with an initial condition consisting of one single particle.

Denote now by $\mu^{\varepsilon, [\infty]}$ the random integer-valued measure on $\mathcal{E}$ obtained by running Algorithm TDMC until time 1. Observe that $\mu^{\varepsilon, [\infty]}$ can be built in the following way. For an excursion $w \in \mathcal{E}$, we build a random measure $\mathscr{Q}^\varepsilon(w)$ by the following procedure. For every $k \in \mathbf{N}$ with $\varepsilon k > \mathfrak{s}(w)$ and $\varepsilon(k+1) < \mathfrak{e}(w) \wedge 1$ we set

$$\Delta w_k = w_{(k+1)\varepsilon} - w_{k\varepsilon} .$$

If $\Delta w_k > 0$, we then draw a random variable N_k^1 with law $\mathcal{I}(\exp(a\Delta w_k) - 1)$ (see (1.5) and Lemma 4.3). For $j = 1, \ldots, N_k^1$, we build i.i.d. excursions $w^{k,j} \in \mathcal{E}$ by the following procedure. First, draw a uniform random variable $u \sim \mathcal{U}(\exp(-a\Delta w_k), 1)$ and set $v = \log u - aw_{k\varepsilon}$. Then, denote by $\{y_{\ell\varepsilon}^{k,j}\}_{\ell=0}^L$ an instance of the random walk (4.1), started at $w_{(k+1)\varepsilon}$ and stopped just before it becomes smaller than $-v/a$ (so that $y_{L\varepsilon}^{k,j} > -v/a$). The excursion $w^{k,j}$ is then given by $\mathfrak{s}(w^{k,j}) = k\varepsilon$, $\mathfrak{e}(w^{k,j}) = (L+k+2)\varepsilon$, $w_{\varepsilon(\ell+k+1)}^{k,j} = y_{\ell\varepsilon}^{k,j}$ for $\ell \in \{0, \ldots, L\}$, $w^{k,j}(\ell\varepsilon) = -v/a$ for the remaining integer values of ℓ, and linear interpolation in between integer values. We then set

$$\mathscr{Q}^\varepsilon(w) = \sum_{k=1}^\infty \sum_{j=1}^{N_k^1} \delta_{w^{k,j}} ,$$

which is the point measure describing the children of the particle with trajectory w. Similarly to before, we build $\mu^{\varepsilon, [\infty]}$ recursively by the following procedure:

- Build an excursion $w^0 \in \mathcal{E}$ as above, starting at x and stopped at $-v/a$, where $(x, v, 0) \in \mathcal{M}$ is the initial condition appearing in the statement. Set $\mu^{\varepsilon, 0} = \delta_{w^0}$.
- Given $\mu^{\varepsilon, \ell}$, define $\mu^{\varepsilon, \ell+1}$ by

$$\mu^{\varepsilon, \ell+1} = \int_\mathcal{E} \mathscr{Q}^\varepsilon(w) \, \mu^{\varepsilon, \ell}(dw) ,$$

  where the $\{\mathscr{Q}^\varepsilon(w)\}_{w \in \mathcal{E}}$ are all independent (and independent of the $\mu^{\varepsilon, \ell'}$ with $\ell' \le \ell$). Note that the integral is actually a finite sum, so the construction makes sense.
- Set $\mu^{\varepsilon, [\ell]} = \sum_{\ell'=0}^\ell \mu^{\varepsilon, \ell'}$ for positive ℓ (including the case $\ell = \infty$).

If we set

$$\mu_t^\varepsilon = E_t^\star \mu^{\varepsilon, [\infty]} ,$$

where E_t is defined as in Section 3.2, then the process μ_t^ε is indeed equal in law to the process considered in Section 4[1].

Denote now by $\mathcal{E}_\gamma$ the set of excursions of height at least γ, namely

$$\mathcal{E}_\gamma = \{w \in \mathcal{E} \, : \, \exists t \in \mathfrak{l}(w) \text{ with } w_t \ge w(\mathfrak{s}(w)) + \gamma\} .$$

---

[1] Strictly speaking, the two processes agree only at times that are multiples of ε since the two interpolation procedures we used may differ when the trajectories of two particles from the same generation cross each other. This is irrelevant.

We also write $\mu_\gamma^{\varepsilon,[n]}$ for the measure obtained exactly like $\mu^{\varepsilon,[n]}$, but where we replace $\mathscr{Q}^\varepsilon(w)$ by its restriction $\mathscr{Q}_\gamma^\varepsilon(w)$ to the set $\mathcal{E}_\gamma$ at every step, so that $\mu_\gamma^{\varepsilon,[n]} \leq \mu_\varepsilon^{[\infty]}$ almost surely. In words, $\mathscr{Q}_\gamma^\varepsilon$ is obtained from $\mathscr{Q}^\varepsilon$ by discarding all excursions of height less than γ, as well as the descendants of any such excursions.

Combining Propositions 4.2 and 4.6, we see that, for ε_0 small enough, there exist constants C and $\alpha > 0$ such that one has the bound

$$\sup_{\varepsilon \leq \varepsilon_0} \mathbf{P}(E_t^\star \mu^{\varepsilon,[\infty]} \neq E_t^\star \mu_\gamma^{\varepsilon,[n]}) \leq C(e^{-\alpha n} + \gamma^p) , \tag{5.1}$$

uniformly over ε and $t \in [0, 1]$.

Following an argument along the lines of the proof of Theorem 2.14, a similar bound can be shown to hold for $\mathbf{P}(E_t^\star \mu^{[\infty]} \neq E_t^\star \mu_\gamma^{[n]})$, where $\mu_\gamma^{[n]}$ is the recursive Poisson process of depth n constructed like $\mu^{[\infty]}$, but with $\mathcal{Q}(w, \cdot)$ replaced by its restriction $\mathcal{Q}_\gamma(w, \cdot)$ to $\mathcal{E}_\gamma$. As a consequence, we conclude that it is sufficient to show that, for every fixed $\gamma > 0$ and $n \geq 1$,

$$\lim_{\varepsilon \to 0} \mathcal{D}(\mu_\gamma^{\varepsilon,[n]}) = \mathcal{D}(\mu_\gamma^{[n]}) , \tag{5.2}$$

where $\mathcal{D}(\cdot)$ denotes the law of a random variable and convergence is to be understood in the sense of weak convergence on the space $\mathscr{M}_+(\mathcal{E})$ endowed with the Wasserstein-1 distance (see [Vil03] and Section 5.2 below).

Our aim then is to make use of Theorem 5.7 below which gives a general convergence result for recursive Poisson point processes. The drawback is that $\mu_\gamma^{\varepsilon,[n]}$ is itself not a recursive Poisson point process, due to the fact that $\mathscr{Q}_\gamma^\varepsilon(w)$ is not a realisation of a Poisson point process. However, it would be one if, in the construction of $\mathscr{Q}^\varepsilon(w)$, we had drawn N_k^1 according to a Poisson distribution with mean $\exp(a\Delta w_k) - 1$. Denote by $\tilde{\mathscr{Q}}^\varepsilon(w)$ the Poisson point process obtained in this way, by $\tilde{\mathscr{Q}}_\gamma^\varepsilon$ its restriction to $\mathcal{E}_\gamma$, and by $\tilde{\mu}_\gamma^{\varepsilon,[n]}$ the recursive Poisson point process of depth n obtained by iterating $\tilde{\mathscr{Q}}_\gamma^\varepsilon$. We then claim that it is possible to find a coupling between $\tilde{\mu}_\gamma^{\varepsilon,[n]}$ and $\mu_\gamma^{\varepsilon,[n]}$ such that

$$\mathbf{P}(\tilde{\mu}_\gamma^{\varepsilon,[n]} \neq \mu_\gamma^{\varepsilon,[n]}) \leq C_{n,\gamma}\sqrt{\varepsilon} , \tag{5.3}$$

where the constant $C_{n,\gamma}$ depends on n and γ, but not on ε. Indeed, if we denote by U_λ a random variable with law $\mathcal{I}(\lambda)$ and by $\bar{U}_\lambda$ a Poisson random variable with mean λ, then it is straightforward to check that there exists a constant C and a coupling between U_λ and $\bar{U}_\lambda$ such that

$$\mathbf{P}(|U_\lambda - \bar{U}_\lambda| = 1) \leq C(1 \wedge \lambda^2) , \quad \mathbf{P}(|U_\lambda - \bar{U}_\lambda| > 1) \leq C(1 \wedge \lambda^3) .$$

Furthermore, by Proposition 2.2, the probability that the random walk started at x reaches γ before becoming negative is bounded from above by $C_\gamma(x + \sqrt{\varepsilon})$ for some constant C_γ depending on γ. As a consequence, one can construct a coupling between $\mathscr{Q}_\gamma^\varepsilon(w)$ and $\tilde{\mathscr{Q}}_\gamma^\varepsilon(w)$ such that

$$\mathbf{P}(\mathscr{Q}_\gamma^\varepsilon(w) \neq \tilde{\mathscr{Q}}_\gamma^\varepsilon(w)) \leq C_\gamma \sum_{k\varepsilon \in [0,1]} |\Delta w_k|^2 (|\Delta w_k| + \sqrt{\varepsilon}) . \tag{5.4}$$

Similarly, regarding the total mass of $\mathscr{Q}_\gamma^\varepsilon$, one has the bound

$$\mathbf{E}\mathscr{Q}_\gamma^\varepsilon(w, \mathcal{E}) = \mathbf{E}\mathscr{Q}^\varepsilon(w, \mathcal{E}_\gamma) \leq C_\gamma \sum_{k\varepsilon \in [0,1]} |\Delta w_k|^2 e^{c|\Delta w_k|} , \tag{5.5}$$

for some constant c. If w is an excursion created by the procedure above, the expected values of (5.4) and (5.5) are bounded by $C_\gamma \sqrt{\varepsilon}$ and C_γ respectively, for a possibly different constant depending on γ. Proceeding as in Lemma 2.7, it follows that, if we set

$$F_\varepsilon(w) = 1 + \frac{1}{\sqrt{\varepsilon}} \sum_{k\varepsilon \in [0,1]} |\Delta w_k|^3 + \sum_{k\varepsilon \in [0,1]} |\Delta w_k|^2 e^{c|\Delta w_k|} ,$$

we obtain the bound $\mathbf{E} \int_{\mathcal{E}} F_\varepsilon(w) \mu_\gamma^{\varepsilon,[n]}(dw) < C_{n,\gamma}$, uniformly over ε. (But this constant might potentially grow very fast as $\gamma \to 0$ and $n \to \infty$!) Combining this with (5.4), the bound (5.3) then follows at once.

As a consequence, it is sufficient to show, instead of (5.2), that

$$\lim_{\varepsilon \to 0} \mathcal{D}(\tilde{\mu}_\gamma^{\varepsilon,[n]}) = \mathcal{D}(\mu_\gamma^{[n]}) . \tag{5.6}$$

This is the content of Proposition 5.10 below, which completes the proof. $\qquad\square$

The remainder of this section is devoted to the proof of (5.6). The outline of the proof goes as follows. First, in Section 5.1, we show that the law of a single excursion of the random walk (4.1), conditioned on hitting a prescribed level γ, converges as $\varepsilon \to 0$ to the Brownian excursion, conditioned to reach γ. In Section 5.2 we then provide a general criterion for the convergence of recursive Poisson point processes. Finally, in Section 5.3 we combine these results in order to obtain (5.6).

5.1 Convergence of excursion measures

As before, we denote by y_t the rescaled random walk given by

$$y_{(k+1)\varepsilon} = y_{k\varepsilon} + \sqrt{\varepsilon} \xi_{k+1} ,$$

where the ξ_k are an i.i.d. sequence of random variables with law ν. As before, we extend this to arbitrary times by linear interpolation. We also assume that ν has some exponential moment as before. The aim of this section is to show that if we start y_t at some initial condition $x \sim \sqrt{\varepsilon}$ and condition it on reaching a prescribed height γ before becoming negative, then its law converges to that of an unnormalised Brownian excursion, conditioned to reach height γ (call it w^γ). Throughout this whole section, we will only consider the process on a fixed time interval, which for definiteness we choose to be equal to $[0,1]$.

A more precise description of the law $\mathbf{Q}_\gamma$ of w^γ is given by the identity

$$\mathbf{Q}_\gamma(\,\cdot\,) = \frac{\int_0^\infty s^{-3/2} g_\gamma(s) \mathbf{Q}_{s,\gamma}(\,\cdot\,) \, ds}{\int_0^\infty s^{-3/2} g_\gamma(s) \, ds} ,$$

where $\mathbf{Q}_{s,\gamma}$ denotes the law of a Brownian excursion of length s, conditioned to reach level γ, and

$$g_\gamma(s) = \mathbf{Q}_{s,0}(\{w \,:\, \sup_{t \le s} w_t \ge \gamma\}) = \mathbf{Q}(\{w \,:\, \sup_{t \le 1} w_t \ge \gamma/\sqrt{s}\}) ,$$

where $\mathbf{Q}$ is the standard Itô excursion measure. Since $g_\gamma(s) \to 0$ exponentially fast for small s, this does indeed define a probability measure on $\mathcal{C}(\mathbf{R}_+, \mathbf{R})$. We then turn this into a probability measure on $\mathcal{C}([0,1], \mathbf{R})$ by restriction.

We view $\mathcal{C}([0,1], \mathbf{R})$ as a subset of $\mathcal{E}$ via the injection $\iota \colon \mathcal{C}([0,1], \mathbf{R}) \hookrightarrow \mathcal{E}$ given by

$$\mathfrak{s}(\iota w) = 0 , \qquad \mathfrak{e}(\iota w) = 1 \wedge \inf\{t > 0 \,:\, w(t) = 0\} ,$$

and by setting the path component of ιw equal to w, stopped when it reaches the time $\mathfrak{e}(\iota w)$. We endow $\mathcal{E}$ as before with the distance d given in (2.24). Since we only deal with excursions starting at 0 and stopped before time 1, the distance d is equivalent on this set to the (pseudo-)distance $\bar{d}$ given by

$$\bar{d}(w, w') = 1 \wedge \left(|\mathfrak{e}(w) - \mathfrak{e}(w')| + \sup_{t \in [0,1]} |w_t - w'_t| \right) .$$

Regarding the space $\mathcal{C}([0, 1], \mathbf{R})$, we endow it throughout this section with the metric

$$d(w, w') = 1 \wedge \sup_{t \in [0,1]} |w_t - w'_t| .$$

We furthermore denote by $\| \cdot \|_d$ the Wasserstein-1 metric associated with any distance function d, which is just the dual of the corresponding Lipschitz (semi-)norm. The main theorem of this section is the following:

Theorem 5.2 *Let $x > 0$ and denote by $\mathbf{Q}^\varepsilon_{z,\gamma}$ the law of the random walk (4.1), starting at $z = \sqrt{\varepsilon} x$ and conditioned to reach level γ before becoming negative. Then, for every $\bar{\beta} < \frac{1}{16}$ there exists a constant C such that*

$$\|\mathbf{Q}^\varepsilon_{z,\gamma} - \mathbf{Q}_\gamma\|_d \le C\varepsilon^{\bar{\beta}} , \qquad \|\iota^\star \mathbf{Q}^\varepsilon_{z,\gamma} - \iota^\star \mathbf{Q}_\gamma\|_{\bar{d}} \le C\varepsilon^{\frac{2\bar{\beta}}{5}} ,$$

uniformly over all $\varepsilon \in (0, 1]$, all $x \in [0, \varepsilon^{-1/3}]$, and all $\gamma \in [\varepsilon^{1/16}, 1]$.

Our main abstract ingredient in the proof is the following criterion for the convergence of conditional probabilities when the probability of the set on which the measures are conditioned converges to 0:

Lemma 5.3 *Let μ and π be two probability measures on some Polish space $\mathcal{Y}$ with metric d and diameter 1 and let $\mathcal{D}_\mu \colon \mathcal{Y} \to [0, 1]$ and $\mathcal{D}_\pi \colon \mathcal{Y} \to [0, 1]$. For $\varrho > 0$, set*

$$\begin{aligned} A_\varrho &= \{ x \in \mathcal{Y} : \exists y \in \mathcal{Y} \text{ with } \mathcal{D}_\pi(y) > \mathcal{D}_\pi(x) + d(x, y)/\varrho \} , \\ \bar{A}_\varrho &= \{ x \in \mathcal{Y} : d(x, A_\varrho) \le \varrho \} . \end{aligned}$$

Assume that δ, ε_1 and ε_2 are such that

$$\int_{\mathcal{Y}} \mathcal{D}_\pi(x)\pi(dx) \ge \delta , \quad \|\mu - \pi\|_d \le \varepsilon_1 , \quad \sup_x |\mathcal{D}_\pi(x) - \mathcal{D}_\mu(x)| \le \varepsilon_2 ,$$

where $\| \cdot \|_d$ denotes the Wasserstein-1 distance with respect to d.
Define measures $\tilde{\mu}$ and $\tilde{\pi}$ by

$$\tilde{\mu}(A) = c_\mu \int_A \mathcal{D}_\mu(x)\mu(dx) , \qquad \tilde{\pi}(A) = c_\pi \int_A \mathcal{D}_\pi(x)\pi(dx) ,$$

where c_μ and c_π are such that these are probability measures.
Then, the bound

$$\|\tilde{\mu} - \tilde{\pi}\|_d \le \frac{1}{\delta}\left(\frac{3\varepsilon_1}{\varrho} + \varepsilon_2 + 2\pi(\bar{A}_\varrho) \right) , \tag{5.7}$$

holds for every $\varrho \le 1$. In particular, one has $\int_{\mathcal{Y}} \mathcal{D}_\mu(x)\mu(dx) > 0$ whenever the right hand side in (5.7) is strictly smaller than 1, so that the bound is non-trivial.

Proof. Let $f \colon \mathcal{Y} \to \mathbf{R}$ be a test function such that $\mathrm{Lip}_d(f) \leq 1$. Since the diameter of $\mathcal{Y}$ is assumed to be 1, we can assume without loss of generality (by possibly adding a constant to f) that $\sup_x |f(x)| \leq \frac{1}{2}$. Since one has the identity

$$\|\tilde{\mu} - \tilde{\pi}\|_d = \sup_{\mathrm{Lip}_d(f) \leq 1} \left(c_\mu \int f(x) \mathcal{D}_\mu(x) \, \mu(dx) - c_\pi \int f(x) \mathcal{D}_\pi(x) \, \pi(dx) \right) = \sup_{\mathrm{Lip}_d(f) \leq 1} \mathcal{I}_f \,,$$

our aim is to bound $\mathcal{I}_f$, uniformly over f. Note first that, by the bound on f,

$$\mathcal{I}_f \leq \frac{c_\pi}{2} \left| \frac{1}{c_\pi} - \frac{1}{c_\mu} \right| + \frac{c_\pi}{2} \int |\mathcal{D}_\mu(x) - \mathcal{D}_\pi(x)| \, \mu(dx)$$

$$+ c_\pi \left| \int f(x) \mathcal{D}_\pi(x) \, \mu(dx) - \int f(x) \mathcal{D}_\pi(x) \, \pi(dx) \right| \,.$$

Note that the first term is nothing but a particular instance of the last term with $f = \frac{1}{2}$. Since the second term is furthermore trivially bounded by $\varepsilon_2 / (2\delta)$, it suffices to bound the last term.

The problem in bounding this term arises of course from the fact that $\mathcal{D}_\pi$ is not Lipschitz continuous. For any $\varrho > 0$, we can however "mollify" this function by setting

$$\mathcal{D}_\pi^\varrho(x) = \sup_{y \in \mathcal{Y}} \left(\mathcal{D}_\pi(y) - \frac{d(x,y)}{\varrho} \right) \,.$$

It is then straightforward to check that $\mathrm{Lip}_d(\mathcal{D}_\pi^\varrho) \leq \varrho^{-1}$, that $\mathcal{D}_\pi(x) \leq \mathcal{D}_\pi^\varrho(x) \leq \sup_y \mathcal{D}_\pi(y)$, and that furthermore $\mathcal{D}_\pi^\varrho(x) = \mathcal{D}_\pi(x)$ for all $x \notin A_\varrho$. It then follows from the definition of ε_1 that

$$\left| \int f(x) \mathcal{D}_\pi^\varrho(x) \, \mu(dx) - \int f(x) \mathcal{D}_\pi^\varrho(x) \, \pi(dx) \right| \leq \varepsilon_1 (1 + (2\varrho)^{-1}) \,.$$

Furthermore,

$$\left| \int f(x) \mathcal{D}_\pi^\varrho(x) \, \pi(dx) - \int f(x) \mathcal{D}_\pi(x) \, \pi(dx) \right| \leq \frac{\pi(A_\varrho)}{2} \,,$$

and similarly for the term with π replaced by μ. In order to bound $\mu(A_\varrho)$, we set as above

$$F^\varrho(x) = \sup_{y \in \mathcal{Y}} \left(\mathbf{1}_{A_\varrho}(y) - \frac{d(x,y)}{\varrho} \right) \,,$$

so that

$$\mu(A_\varrho) \leq \int F^\varrho(x) \, \mu(dx) \leq \frac{\varepsilon_1}{\varrho} + \int F^\varrho(x) \, \pi(dx) \leq \frac{\varepsilon_1}{\varrho} + \pi(\bar{A}_\varrho) \,,$$

where we used the fact that F^ϱ vanishes outside of $\bar{A}_\varrho$. Collecting all of these bounds, the claim follows. □

An alternative description of w^γ is given by the following. Let Y be a Bessel-3 process starting at the origin and let τ_γ be its first hitting time of γ, i.e. $\tau_\gamma = \inf\{t \geq 0 : Y_t \geq \gamma\}$. Let furthermore B be a Brownian motion independent of Y, which is stopped when it reaches the level γ. Then, one has the decomposition

$$w_t^\gamma = \begin{cases} Y_t & \text{for } t \leq \tau_\gamma, \\ \gamma - B_{t - \tau_\gamma} & \text{for } t \geq \tau_\gamma. \end{cases} \tag{5.8}$$

This is a consequence of the symmetry of the Brownian excursion under time reversal, combined with [RW94, Theorem 49.1] for example.

We can use the above decomposition to obtain the following bound:

Lemma 5.4 *For any $\beta < \frac{1}{4}$, there exists a constant C such that, for every $\gamma, \gamma' \in (0, 1]$, one has the bound*

$$\|\mathbf{Q}_\gamma - \mathbf{Q}_{\gamma'}\|_d \le C|\gamma - \gamma'|^\beta .$$

Proof. The decomposition (5.8) suggests a natural coupling between w^γ and $w^{\gamma'}$ by building them from the same basic building blocks Y and B. The characterisation of the Bessel-3 process as the norm of a 3-dimensional Brownian motion, together with standard hitting estimates for Brownian motion, imply that

$$\mathbf{P}(|\tau_\gamma - \tau_{\gamma'}| > \zeta) \le 1 \wedge \frac{C|\gamma - \gamma'|}{\zeta^{3/2}} ,$$

so that in particular $\mathbf{P}(|\tau_\gamma - \tau_{\gamma'}| \ge \sqrt{|\gamma - \gamma'|}) \le C|\gamma - \gamma'|^{1/4}$. The result now follows from the fact that both B and Y are almost surely α-Hölder continuous for every $\alpha < \frac{1}{2}$. $\square$

Lemma 5.5 *Let B_z^γ be a Brownian motion starting at z, conditioned to hit γ before 0, and stopped upon its return to 0. Then, for every $\beta < 1$, there exists a constant C depending on β such that the bound*

$$\|\mathcal{D}(B_\varepsilon^\gamma) - \mathbf{Q}_\gamma\|_d \le C_\gamma \varepsilon^\beta ,$$

holds uniformly over $\varepsilon \in (0, \gamma \wedge 1]$ and $\gamma > 0$.

Proof. Let w^γ be as above and let τ_ε be the first passage time of w^γ through ε. Then, it follows from the decomposition (5.8) that one has the exact identity

$$B_\varepsilon^\gamma(\cdot) \stackrel{\text{law}}{=} w^\gamma(\cdot - \tau_\varepsilon) . \tag{5.9}$$

It then follows from the small ball estimates of Brownian motion that, for every $\bar{\zeta} < 2$, one has the bound

$$\mathbf{P}(\tau_\varepsilon \ge \varepsilon^{\bar{\zeta}}) \lesssim \varepsilon .$$

The desired estimate then follows at once from the Hölder regularity of w^γ. $\square$

Proposition 5.6 *Let $\alpha \in (0, \frac{1}{8})$. Suppose further that $\gamma > 0$ is fixed and denote by y_t^γ the random walk y_t conditioned to hit $[\gamma, \infty)$ before hitting $\mathbf{R}_-$ and stopped when it then hits $\mathbf{R}_-$. Assume that $y_0^\gamma = \varepsilon^\alpha$. Then the law of y^γ converges weakly to $\mathbf{Q}_\gamma$ as $\varepsilon \to 0$. Furthermore, for every $\beta > 0$, there exists a constant C such that the bound*

$$\|\mathcal{D}(y^\gamma) - \mathbf{Q}_\gamma\|_d \le C(\sqrt{\gamma}\varepsilon^{\frac{1}{8}-\alpha} + \varepsilon^{\alpha-\beta}) ,$$

holds uniformly over all $\varepsilon \le 1$ and $\gamma \in [\varepsilon^\alpha, 1]$.

Proof. By Lemma 5.5, it suffices to compare the law of y^γ with that of $B_{\varepsilon^\alpha}^\gamma$. The result will then be a consequence of Lemma 5.3.

To see this, we partition the state space $\mathcal{Y} = \{w \in \mathcal{C}([0,1], \mathbf{R}) \ : \ w_0 = \varepsilon^\alpha\}$ into three sets in the following way. Given a continuous function w with $w_0 \in (0, \gamma)$, we set $\tau = 1 \wedge \inf\{t > 0 \ : \ w_t \notin [0, \gamma]\}$, and we define sets $A^{(i)}$ with $i \in \{1, 2, 3\}$ by

$$A^{(1)} = \{w \ : \ w_\tau = 0\}, \qquad A^{(2)} = \{w \ : \ w_\tau = \gamma\}, \qquad A^{(3)} = \{w \ : \ \tau = 1\}.$$
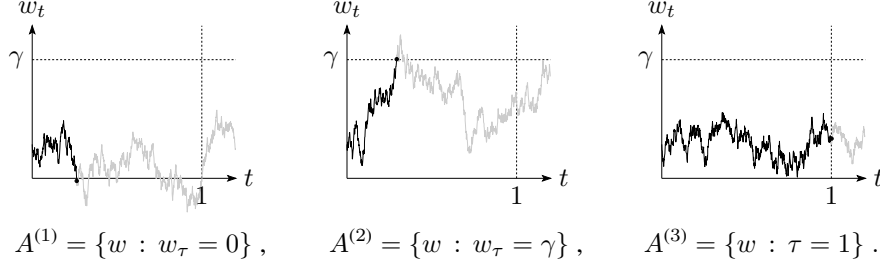
Define furthermore functions F_γ and F_γ^ε on $\mathcal{Y}$ by

$$F_\gamma(w) = \frac{w_\tau}{\gamma}, \qquad F_\gamma^\varepsilon(w) = \begin{cases} 0 & \text{if } w \in A^{(1)}, \\ 1 & \text{if } w \in A^{(2)}, \\ \bar{P}_{\frac{\gamma}{\sqrt{\varepsilon}}, \frac{w_1}{\sqrt{\varepsilon}}} & \text{if } w \in A^{(3)}, \end{cases}$$

where $\bar{P}_{z,\gamma}$ is defined as in the discussion before Proposition 2.2. With these definitions at hand, if we set $\mu = \mathcal{D}(y)$ with $y_0 = \varepsilon^\alpha$, $\pi = \mathcal{D}(B_{\varepsilon^\alpha})$, $\mathcal{D}_\mu = F_\gamma^\varepsilon$, and $\mathcal{D}_\pi = F_\gamma$, then we are precisely in the setting of Lemma 5.3 with $\tilde{\mu} = \mathcal{D}(y^\gamma)$ and $\tilde{\pi} = \mathcal{D}(B_{\varepsilon^\alpha}^\gamma)$.

Note first that, since F_γ is precisely the probability that a Brownian motion started from w_1 hits γ before 0, we have

$$\delta = \int \mathcal{D}_\pi(w)\pi(dw) = \frac{\varepsilon^\alpha}{\gamma}.$$

Furthermore, it follows from Corollary 2.3 (and from the fact that F_γ and F_γ^ε coincide outside of $A^{(3)}$) that

$$\varepsilon_1 = \sup_w |F_\gamma(w) - F_\gamma^\varepsilon(w)| \lesssim \frac{\varepsilon^{\frac{1}{4}}}{\gamma}.$$

(We could have replaced $\frac{1}{4}$ by any exponent less than $\frac{1}{2}$ here.) Regarding the distance between the unconditioned measures, we obtain from [Fra73] the bound

$$\varepsilon_2 = \|\mu - \pi\|_d \lesssim \varepsilon^{\frac{1}{8}}.$$

It thus remains to obtain a bound on $\bar{A}_\varrho$. By the definition of A_ϱ and of F_γ, $w \in A_\varrho$ implies that either $w \in A^{(1)} \cup A^{(3)}$ and $d(w, A^{(2)}) \le \varrho$ or $w \in A^{(1)}$ and $d(w, A^{(3)}) \le \varrho$. This implies that

$$A_\varrho \subset \left\{w \ : \ \sup_{t \in [0,1]} w_t \in [\gamma - \varrho, \gamma]\right\} \cup \left\{w \ : \ \inf_{t \in [0,1]} w_t \in [-\varrho, 0]\right\},$$

so that

$$\bar{A}_\varrho \subset \left\{w \ : \ \sup_{t \in [0,1]} w_t \in [\gamma - 2\varrho, \gamma + \varrho]\right\} \cup \left\{w \ : \ \inf_{t \in [0,1]} w_t \in [-2\varrho, \varrho]\right\}.$$

Since (by the reflection principle) the law of the extremum of Brownian motion over a finite time interval has a smooth density with respect to Lebesgue measure, there exists a constant C independent of ε and δ such that $\pi(\bar{A}_\varrho) \le C\varrho$.

Inserting these bounds into Lemma 5.3, we thus obtain the bound

$$\|\mathcal{D}(y^\gamma) - \mathcal{D}(B^\gamma_{\varepsilon^\alpha})\|_d \lesssim \frac{\gamma}{\varepsilon^\alpha} \left(\frac{\varepsilon^{\frac{1}{4}}}{\varrho \gamma} + \varepsilon^{\frac{1}{8}} + \varrho \right) .$$

Setting $\varrho = \varepsilon^{\frac{1}{8}} \gamma^{-\frac{1}{2}}$ completes the proof. $\qquad\square$

We now have all the ingredients required for the proof of Theorem 5.2.

Proof of Theorem 5.2. Assume as in the previous proof that y^γ_t is the random walk y_t conditioned to hit $[\gamma, \infty)$ before hitting $\mathbf{R}_-$ and stopped when it then hits $\mathbf{R}_-$. In contrast to the above setup we now assume that $y^\gamma_0 = z = x\sqrt{\varepsilon}$ for some $x \geq 0$. Let k_0 be given by

$$k_0 = \inf \left\{ k > 0 : y^\gamma_{k\varepsilon} \geq \varepsilon^{\frac{1}{16}} \right\} .$$

It then follows from [Fra73], combined with standard small ball estimates for Brownian motion that, for every $\beta < \frac{1}{8}$ and every $p > 0$ there exists a constant C such that

$$\mathbf{P}(\varepsilon k_0 > \varepsilon^\beta) \lesssim \varepsilon^p , \qquad (5.10)$$

uniformly over $\varepsilon \leq 1$, for all x such that $x\sqrt{\varepsilon} \leq \varepsilon^{\frac{1}{16}}$. Furthermore, the probability that $y^\gamma_{k_0\varepsilon} > 2\varepsilon^{\frac{1}{16}}$ (say) is exponentially small in ε, again uniformly over x. It then follows from Proposition 5.6 (choosing $\alpha = \frac{1}{16}$) that, for every $\bar\beta < \frac{1}{16}$ one can construct a joint realisation of y^γ and w^γ such that

$$\mathbf{E}d(y^\gamma, w^\gamma(\cdot - \varepsilon k_0)) \lesssim \varepsilon^{\bar\beta} .$$

(Here we extend w^γ by setting it to 0 for negative times.) On the other hand, the Hölder regularity of the sample paths of w^γ together with (5.10) implies that

$$\mathbf{E}d(w^\gamma, w^\gamma(\cdot - \varepsilon k_0)) \lesssim \varepsilon^{\bar\beta} ,$$

so that the bound on $\|\mathbf{Q}^\varepsilon_{z,\gamma} - \mathbf{Q}_\gamma\|_d$ follows.

In order to obtain the bound on $\|\iota^\star \mathbf{Q}^\varepsilon_{z,\gamma} - \iota^\star \mathbf{Q}_\gamma\|_{\bar d}$, we make use of the same coupling between y^γ and w^γ as above, so that $\mathbf{E}d(y^\gamma, w^\gamma) \lesssim \varepsilon^{\bar\beta}$. It then remains to obtain a bound on

$$|\mathfrak{e}(\iota y^\gamma) - \mathfrak{e}(\iota w^\gamma)| .$$

For this, note first that, by Chebychev's inequality, one has the bound

$$\mathbf{P}(d(y^\gamma, w^\gamma) \geq \varepsilon^\beta) \lesssim \varepsilon^{\bar\beta - \beta} , \qquad (5.11)$$

valid for every $\beta \leq \bar\beta$. Consider now any two paths y^γ and w^γ at distance less than ε^β and define

$$\tau_1 = 1 \wedge \inf\{t > 0 : w^\gamma(t) \leq \varepsilon^\beta\} , \qquad \tau_2 = 1 \wedge \inf\{t > \tau_1 : w^\gamma(t) < -\varepsilon^\beta\} .$$

In this way, one has both $\mathfrak{e}(\iota w^\gamma) \in [\tau_1, \tau_2]$ and $\mathfrak{e}(\iota y^\gamma) \in [\tau_1, \tau_2]$, so that it remains to obtain a bound on $\tau_2 - \tau_1$. The explicit expression for the hitting time of a line for a Brownian motion yields

$$\mathbf{P}(\tau_2 - \tau_1 \geq \varepsilon^\alpha) \leq \varepsilon^{\beta - \frac{\alpha}{2}} ,$$

for any $\alpha < 2\beta$. Choosing $\beta = \frac{3\bar\beta}{5}$ and $\alpha = \frac{2\bar\beta}{5}$ and combining this with (5.11), we then obtain

$$\mathbf{P}(|\mathfrak{e}(\iota y^\gamma) - \mathfrak{e}(\iota w^\gamma)| \geq \varepsilon^{\frac{2\bar\beta}{5}}) \lesssim \varepsilon^{\frac{2\bar\beta}{5}} ,$$

from which the bound follows. $\qquad\square$

5.2 Convergence of recursive Poisson point processes

The aim of this section is to provide a general result allowing us to bound the distance between two recursive Poisson point processes of the same depth n in terms of their respective kernels. This result is the main abstract result on which the proof of the convergence result (5.6) will be based. One difficulty that we have to overcome is that there is very little uniformity in the proximity of the kernel describing $\mu_\gamma^{\varepsilon,[n]}$ to the one describing $\mu_\gamma^{[n]}$.

Throughout this section, given a Polish space $\mathcal{X}$ with a distance function d bounded by 1, we define the Wasserstein-1 distance between any two positive measures with finite mass (and not just probability measures!) by

$$\|\mu - \pi\|_1 = \sup_{\substack{\mathrm{Lip} f \le 1 \\ \|f\|_\infty \le 1}} \left(\int_{\mathcal{X}} f(x)\mu(dx) - \int_{\mathcal{X}} f(x)\pi(dx) \right) .$$

If μ and π happen to have the same mass, then the additional constraint on the supremum norm $\|f\|_\infty$ of f is redundant in the above expression, and we simply recover the usual Wasserstein-1 distance. In the case where the masses of μ and π are different however, our choice of definitions ensures that

$$\|\mu - \pi\|_1 \approx |\mu(\mathcal{X}) - \pi(\mathcal{X})| + |\mu(\mathcal{X})| \left\| \frac{\mu}{\mu(\mathcal{X})} - \frac{\pi}{\pi(\mathcal{X})} \right\|_1 , \qquad (5.12)$$

where $\approx$ denotes that both quantities are bounded by multiples of each other, with proportionality constants that are independent of μ and π.

The main result of this section is the following:

Theorem 5.7 *Let $\mathcal{Q}$ and $\bar{\mathcal{Q}} \colon \mathcal{X} \to \mathscr{M}_+(\mathcal{X})$ be two measurable maps and assume that the Polish space $\mathcal{X}$ is endowed with a metric d bounded by 1. Let $A \subset \mathcal{X}$, $\varepsilon \in (0,1]$ and $K \ge 1$ be such that the bounds*

$$\sup_{x \in \mathcal{X}} \mathcal{Q}(x, \mathcal{X}) \le K , \qquad\qquad \|\mathcal{Q}(x) - \mathcal{Q}(y)\|_1 \le K d(x,y) , \qquad (5.14)$$

$$\sup_{x \in A} \bar{\mathcal{Q}}(x, A^c) \le \varepsilon , \qquad\qquad \sup_{x \in A} \|\mathcal{Q}(x) - \bar{\mathcal{Q}}(x)\|_1 \le \varepsilon , \qquad (5.15)$$

hold, where we use the notation $A^c = \mathcal{X} \setminus A$.

Fix $n > 0$, $\bar{x} \in A$ and $x \in \mathcal{X}$, and denote by $\mu_x^{[n]}$ and $\bar{\mu}_{\bar{x}}^{[n]}$ the recursive Poisson point processes with respective kernels $\mathcal{Q}$ and $\bar{\mathcal{Q}}$. Then, there exists a coupling between $\mu_x^{[n]}$ and $\bar{\mu}_{\bar{x}}^{[n]}$ such that

$$\mathbf{E}(1 \wedge \|\mu_x^{[n]} - \bar{\mu}_{\bar{x}}^{[n]}\|_1) \lesssim C(\sqrt{\varepsilon} + d(x, \bar{x})) ,$$

where the proportionality constant C depends only on K and n.

Remark 5.8 One useful feature of the way that we set up the bounds in the statement is that we only require information about the kernel $\bar{\mathcal{Q}}$ on the set A. In the application we have in mind, the kernel $\bar{\mathcal{Q}}$ will be the one describing $\tilde{\mu}_\gamma^{\varepsilon,[n]}$, while the set A will consist of trajectories exhibiting "typical" behaviour at small scales.

Before we turn to the proof of this theorem, we show that if $\pi_n \to \pi$ in the Wasserstein-1 sense, then the (usual) Poisson point processes with these intensity measures also converge to each other weakly in the Wasserstein-1 distance:

Proposition 5.9 *Let π and $\bar{\pi}$ be two finite positive measures on a Polish space $\mathcal{X}$ endowed with a metric $d \le 1$ and let μ and $\bar{\mu}$ be the corresponding Poisson point processes on $\mathcal{X}$. Then, there exists a constant C and a coupling between μ and $\bar{\mu}$ such that*

$$\mathbf{E}(1 \wedge \|\mu - \bar{\mu}\|_1) \le C(\|\pi - \bar{\pi}\|_1 \wedge 1) \ .$$

Proof. The proof relies on the fact that, if $\mathcal{P}(\lambda)$ denotes the Poisson distribution with parameter λ, one has the total variation bound

$$d_{\mathrm{TV}}(\mathcal{P}(\lambda), \mathcal{P}(\bar{\lambda})) \le |\lambda - \bar{\lambda}| \wedge 1 \ , \tag{5.16}$$

see for example [AJ06].

We can construct μ (and similarly for $\bar{\mu}$) in the following way. First, draw a Poisson random variable N with parameter $\pi(\mathcal{X})$. Then, draw N independent random variables $\{X_1, \ldots, X_N\}$ with law $\pi/\pi(\mathcal{X})$ and set

$$\mu = \sum_{k=1}^{N} \delta_{X_k} \ . \tag{5.17}$$

By (5.16), we can now construct a Poisson random variable $\bar{N}$ with parameter $\bar{\pi}(\mathcal{X})$ such that

$$\mathbf{P}(\bar{N} \ne N) \le |\pi(\mathcal{X}) - \bar{\pi}(\mathcal{X})| \ . \tag{5.18}$$

Assuming that $\bar{N} = N$, we can then draw random variables $\{\bar{X}_1, \ldots, \bar{X}_N\}$ in such a way that the pairs $(\bar{X}_k, X_k)$ are distributed according to a coupling between $\pi/\pi(\mathcal{X})$ and $\bar{\pi}/\bar{\pi}(\mathcal{X})$ that minimises their expected distance. If $\bar{N} \ne N$, then we simply draw $\{\bar{X}_1, \ldots, \bar{X}_N\}$ according to $\bar{\pi}/\bar{\pi}(\mathcal{X})$, independently of the X_k.

If we then define $\bar{\mu}$ similarly to (5.17), it follows that

$$\|\mu - \bar{\mu}\|_1 \le \begin{cases} N \|\frac{\pi}{\pi(\mathcal{X})} - \frac{\bar{\pi}}{\bar{\pi}(\mathcal{X})}\|_1 & \text{if } N = \bar{N}, \\ N + \bar{N} & \text{otherwise.} \end{cases}$$

As a consequence, we obtain the bound

$$\mathbf{E}(1 \wedge \|\mu - \bar{\mu}\|_1) \le \mathbf{P}(N \ne \bar{N}) + \left\| \frac{\pi}{\pi(\mathcal{X})} - \frac{\bar{\pi}}{\bar{\pi}(\mathcal{X})} \right\|_1 \mathbf{E}N \ ,$$

so that the claim follows from (5.18), the definition of N, and (5.12). $\qquad\square$

Proof of Theorem 5.7. Note first that, by combining the first bound in (5.14) with the second bound in (5.15), we obtain the bound

$$\sup_{x \in A} \bar{\mathcal{Q}}(x, \mathcal{X}) \le K + \varepsilon \ .$$

It follows that we have the recursive bound

$$\mathbf{E}(\bar{\mu}_{\bar{x}}^n(A) \,|\, \bar{\mu}_{\bar{x}}^{[n-1]}(A^{\mathrm{c}}) = 0) \le 2K \mathbf{E}(\bar{\mu}_{\bar{x}}^{n-1}(A) \,|\, \bar{\mu}_{\bar{x}}^{[n-2]}(A^{\mathrm{c}}) = 0) \ ,$$

so that, since $\bar{\mu}_{\bar{x}}^0(A) = 1$ by assumption,

$$\mathbf{E}(\bar{\mu}_{\bar{x}}^n(A) \,|\, \bar{\mu}_{\bar{x}}^{[n-1]}(A^{\mathrm{c}}) = 0) \le (2K)^n \ . \tag{5.19}$$

On the other hand, defining the positive measures π^n and $\bar\pi^n$ by

$$\pi^{n+1} = \int_{\mathcal{X}} \mathcal{Q}(y)\mu_x^n(dy) \qquad \text{and} \qquad \bar\pi^{n+1} = \int_{\mathcal{X}} \bar{\mathcal{Q}}(y)\bar\mu_{\bar x}^n(dy)$$

we have the inequality

$$\begin{aligned}
\mathbf{P}(\bar\mu_{\bar x}^{[n]}(A^c) > 0) &\le \mathbf{P}(\bar\mu_{\bar x}^{[n-1]}(A^c) > 0) + \mathbf{P}(\bar\mu_{\bar x}^n(A^c) > 0 \,|\, \bar\mu_{\bar x}^{[n-1]}(A^c) = 0) \\
&\le \mathbf{P}(\bar\mu_{\bar x}^{[n-1]}(A^c) > 0) + \mathbf{E}(\bar\pi^n(A^c) \,|\, \bar\mu_{\bar x}^{[n-1]}(A^c) = 0) \\
&\le \mathbf{P}(\bar\mu_{\bar x}^{[n-1]}(A^c) > 0) + \varepsilon D + \mathbf{P}(\bar\mu_{\bar x}^{n-1}(A) > D \,|\, \bar\mu_{\bar x}^{[n-1]}(A^c) = 0) \\
&\le \mathbf{P}(\bar\mu_{\bar x}^{[n-1]}(A^c) > 0) + \varepsilon D + \frac{1}{D}\mathbf{E}(\bar\mu_{\bar x}^{n-1}(A) \,|\, \bar\mu_{\bar x}^{[n-1]}(A^c) = 0) \\
&\le \mathbf{P}(\bar\mu_{\bar x}^{[n-1]}(A^c) > 0) + \varepsilon D + \frac{(2K)^{n-1}}{D} \,,
\end{aligned}$$

which is valid uniformly over all $D > 0$. Choosing $D \sim 1/\sqrt{\varepsilon}$, we thus obtain the recursion relation

$$\mathbf{P}(\bar\mu_{\bar x}^{[n]}(A^c) > 0) \le \mathbf{P}(\bar\mu_{\bar x}^{[n-1]}(A^c) > 0) + C\sqrt{\varepsilon} \,,$$

from which it follows that

$$\mathbf{P}(\bar\mu_{\bar x}^{[n]}(A^c) > 0) \le C\sqrt{\varepsilon} \,, \tag{5.20}$$

where in both cases the constant C depends on K and n, but not on ε.

Note now that we have the bound

$$\begin{aligned}
\|\pi^{n+1} - \bar\pi^{n+1}\|_1 &\le \Big\| \int_{\mathcal{X}} \mathcal{Q}(y)\,(\mu_x^n - \bar\mu_{\bar x}^n)(dy) \Big\|_1 + \int_{\mathcal{X}} \|\mathcal{Q}(y) - \bar{\mathcal{Q}}(y)\|_1 \bar\mu_{\bar x}^n(dy) \\
&\le \|\mu_x^n - \bar\mu_{\bar x}^n\|_1 (\mathrm{Lip}(\mathcal{Q}) + \|\mathcal{Q}\|_\infty) + \varepsilon\bar\mu_{\bar x}^n(A) \\
&\quad + \int_{A^c} \|\mathcal{Q}(y) - \bar{\mathcal{Q}}(y)\|_1 \bar\mu_{\bar x}^n(dx) \,,
\end{aligned}$$

so that the bound

$$1 \wedge \|\pi^{n+1} - \bar\pi^{n+1}\|_1 \le 2K(\|\mu_x^n - \bar\mu_{\bar x}^n\|_1 \wedge 1) + \varepsilon D + \mathbf{1}_{\bar\mu_{\bar x}^n(A) > D} + \mathbf{1}_{\bar\mu_{\bar x}^n(A^c) > 0} \,,$$

is valid for every $D > 0$. Furthermore, by Chebyshev's inequality and (5.19)–(5.20), one has

$$\begin{aligned}
\mathbf{P}(\bar\mu_{\bar x}^n(A) > D) &\le \mathbf{P}(\bar\mu_{\bar x}^n(A) > D \,|\, \bar\mu_{\bar x}^{n-1}(A^c) = 0) + \mathbf{P}(\bar\mu_{\bar x}^{n-1}(A^c) > 0) \\
&\le \frac{K^n}{D} + C\sqrt{\varepsilon} \,.
\end{aligned}$$

Choosing again $D \sim 1/\sqrt{\varepsilon}$, we thus obtain the bound

$$\mathbf{E}(1 \wedge \|\pi^{n+1} - \bar\pi^{n+1}\|_1) \le 2K\mathbf{E}(\|\mu_x^n - \bar\mu_{\bar x}^n\|_1 \wedge 1) + C\sqrt{\varepsilon} \,.$$

Applying Proposition 5.9, we conclude that, given μ_x^n and $\bar\mu_{\bar x}^n$, it is possible to construct a coupling between μ_x^{n+1} and $\bar\mu_{\bar x}^{n+1}$ such that

$$\mathbf{E}(1 \wedge \|\mu_x^{n+1} - \bar\mu_{\bar x}^{n+1}\|_1) \lesssim \mathbf{E}(1 \wedge \|\mu_x^n - \bar\mu_{\bar x}^n\|_1) + \sqrt{\varepsilon} \,,$$

from whence the claim now follows at once. $\qquad\square$

5.3 Convergence of the truncated distributions

We are now in a position to provide the proof of (5.6). Again, throughout this section, we make the standing assumption that the one-step distribution ν for the random walk (4.1) has some exponential moment. We also use the notations $\tilde{\mu}_\gamma^{\varepsilon,[n]}$ and $\mu_\gamma^{[n]}$ from (5.6). Let $\mathcal{E}$ be the space of real-valued excursions as before. We then introduce the map $\mathscr{J}^\varepsilon \colon \mathcal{E} \to \mathscr{M}_+(\mathbf{R}^2)$ given by

$$\mathscr{J}^\varepsilon(w)(dz, dt) = a\varepsilon \sum_{\varepsilon k \in [0,1]} e^{a\sqrt{\varepsilon} z} \, \delta_{\varepsilon k}(dt) \, \mathbf{1}_{[0, \Delta_k^\varepsilon w]}(z) \, dz \;, \qquad (5.21)$$

where δ_z denotes the Dirac measure located at z, $\Delta^\varepsilon w_k$ is defined by

$$\Delta_k^\varepsilon w = \frac{w_{(k+1)\varepsilon} - w_{k\varepsilon}}{\sqrt{\varepsilon}} \;,$$

and we used the convention that $\mathbf{1}_{[0,z]} = 0$ if $z < 0$.

As before, denote by $\mathbf{Q}_{z,\gamma}^\varepsilon$ the law of the random walk (4.1), starting at $z\sqrt{\varepsilon}$, and conditioned to hit level γ before becoming negative. We stop it as soon as it hits $\mathbf{R}_-$, so that we interpret $\mathbf{Q}_{z,\gamma}^\varepsilon$ as a measure on $\mathcal{E}_0$. Recall also that $\bar{P}_{z,\gamma/\sqrt{\varepsilon}}$, with $\bar{P}_{z,\gamma}$ defined as in Section 2.1.2 denotes the probability that this event actually happens.

With this notation, the measure $\mathcal{Q}_\gamma^\varepsilon$ describing $\tilde{\mu}_\gamma^{\varepsilon,[n]}$ (i.e. $\mathcal{Q}_\gamma^\varepsilon(w, \cdot)$ is the intensity measure of $\tilde{\mathscr{D}}_\gamma^\varepsilon(w)$) is given by

$$\mathcal{Q}_\gamma^\varepsilon(w, \cdot) = \int_{\mathbf{R}^2} \Theta_{w,t}^\star \mathbf{Q}_{z,\gamma}^\varepsilon \, \bar{P}_{z,\frac{\gamma}{\sqrt{\varepsilon}}} \, \frac{\mathscr{J}^\varepsilon(w)(dz, dt)}{\sqrt{\varepsilon}}$$
$$= \int_{\mathbf{R}^2} \left(\frac{1}{\gamma} \Theta_{w,t}^\star \mathbf{Q}_{z,\gamma}^\varepsilon \right) \left(\frac{\gamma}{\sqrt{\varepsilon}} \bar{P}_{z,\frac{\gamma}{\sqrt{\varepsilon}}} \right) \mathscr{J}^\varepsilon(w)(dz, dt) \;.$$

Note now that if w is a typical realisation of $\mathbf{Q}_{z,\gamma}^\varepsilon$, then $\mathscr{J}^\varepsilon(w)$ is expected to be close to the measure

$$\mathscr{J}(w) = \mathbf{1}_{I(w) \cap [0,1]}(t) \, dt \otimes \hat{\nu}(dz) \;,$$

where $\hat{\nu}$ is the measure on $\mathbf{R}_+$ given by

$$\int G(z)\hat{\nu}(dz) = \int_0^\infty \int_0^z G(y) \, dy \, \nu(dz) \;,$$

for any test function G. This is because $e^{a\sqrt{\varepsilon}\Delta_k^\varepsilon w} \sim 1$ and the law of $\Delta_k^\varepsilon w$ would be given by ν, were it not for the conditioning.

On the other hand, the kernel $\mathcal{Q}_\gamma$ describing the truncated Brownian fan $\mu_\gamma^{[n]}$ with parameter a is given by

$$\mathcal{Q}_\gamma(w, \cdot) = \frac{a}{2\delta} \int_{I(w) \cap [0,1]} \Theta_{w,t}^\star \mathbf{Q}_\gamma \, dt = \frac{1}{\delta} \int_{\mathbf{R}^2} \Theta_{w,t}^\star \mathbf{Q}_\gamma \, G(z) \, \mathscr{J}(w)(dz, dt) \;, \quad (5.22)$$

where $\mathbf{Q}_\gamma$ is the law of a Brownian excursion conditioned to reach level γ and G was defined in (2.13). This is the case because of the well-known fact that $\frac{1}{\gamma}\mathbf{Q}_\gamma$ is the law of the unnormalised Brownian excursion *restricted* to the set of excursions reaching level γ. The second identity in (5.22) is a consequence of the definition of $\hat{\nu}$, combined with Proposition 2.4. With these notations at hand, we have:

Proposition 5.10 *Consider the setting and assumptions of Theorem 5.1. For every $\delta < \frac{1}{32}$, there exists a constant C depending on γ, n and δ such that*

$$\mathbf{E}(1 \wedge \|\tilde{\mu}_\gamma^{\varepsilon,[n]} - \mu_\gamma^{[n]}\|_1) \le C\varepsilon^\delta \ ,$$

uniformly over $\varepsilon \le \varepsilon_0$ for some ε_0 small enough.

Proof. We apply Theorem 5.7 with $\mathcal{Q} = \mathcal{Q}_\gamma$, $\bar{\mathcal{Q}} = \mathcal{Q}_\gamma^\varepsilon$, and A to be determined. Using the results obtained earlier in this section, it turns out that the assumptions are then relatively straightforward to check.

For convenience, we introduce the notation

$$\tilde{\mathscr{J}}^\varepsilon(w)(dz, dt) = G(x)\,\mathscr{J}^\varepsilon(w)(dz, dt) \ ,$$

and similarly for $\tilde{\mathscr{J}}$. We also denote by $\Pi_2 \colon \mathbf{R}^2 \to \mathbf{R}$ the projection onto the second component, so that $\Pi_2^\star \tilde{\mathscr{J}}^\varepsilon(w)$ (and similarly for $\tilde{\mathscr{J}}$) is the projection of $\tilde{\mathscr{J}}^\varepsilon(w)$ onto the t-component. We also denote by $\Omega_\varepsilon \subset \mathcal{E}$ the subset of excursions given by

$$\Omega_\varepsilon = \Omega_\varepsilon^{(1)} \cap \Omega_\varepsilon^{(2)} \ ,$$

where we set

$$\Omega_\varepsilon^{(1)} = \{w \in \mathcal{E} \ : \ \tilde{\mathscr{J}}^\varepsilon(w)(\{x > \varepsilon^{-1/3}\}) = 0\} \ ,$$
$$\Omega_\varepsilon^{(2)} = \{w \in \mathcal{E} \ : \ \|\Pi_2^\star \tilde{\mathscr{J}}(w) - \Pi_2^\star \tilde{\mathscr{J}}^\varepsilon(w)\|_1 \le \varepsilon^{1/10}\} \ .$$

Also, in the definition of $\Omega_\varepsilon^{(2)}$, the Wasserstein-1 distance is taken with respect to the somewhat unusual distance on $\mathbf{R}$ given by

$$d(t, t') = 1 \wedge |t - t'|^{1/3} \ . \tag{5.23}$$

The reason for this seemingly strange choice will become clear later. The set Ω_ε defined in this way will play the role of A when applying Theorem 5.7.

Note first that $\mathcal{Q}_\gamma(w, \mathcal{X}) \le \frac{a}{2\gamma}$, which is bounded independently of w. Furthermore, one has the bound

$$\|\mathcal{Q}_\gamma(w, \mathcal{X}) - \mathcal{Q}_\gamma(\bar{w}, \mathcal{X})\|_1 \le \frac{a}{2\gamma} \int_{\mathfrak{l}(w) \cap \mathfrak{l}(\bar{w}) \cap [0,1]} \|\Theta_{w,t}^\star \mathbf{Q}_\gamma - \Theta_{\bar{w},t}^\star \mathbf{Q}_\gamma\|_1 \, dt$$
$$+ \frac{a}{2\gamma}(|\mathfrak{s}(w) - \mathfrak{s}(\bar{w})| + |\mathfrak{e}(w) - \mathfrak{e}(\bar{w})|) \ .$$

Since furthermore $\|\Theta_{w,t}^\star \mathbf{Q}_\gamma - \Theta_{\bar{w},t}^\star \mathbf{Q}_\gamma\|_1 \le |w_t - \bar{w}_t|$, it does indeed follow that $\mathcal{Q}$ is globally Lipschitz continuous as required, so that (5.14) holds with some constant K depending on γ.

It remains to check that (5.15) holds, with ε replaced by some power of ε. By Theorem 5.2, we already know that there exists a constant C, possibly depending on γ, such that

$$\|\mathbf{Q}_{z,\gamma}^\varepsilon - \mathbf{Q}_\gamma\|_1 \le C\varepsilon^\delta \ ,$$

for any exponent $\delta \in (0, \frac{1}{16})$, and uniformly over all $z \le \varepsilon^{-1/3}$. As a consequence, for every $w \in \Omega_\varepsilon$, we have the bound

$$\|\mathcal{Q}_\gamma(w, \mathcal{X}) - \mathcal{Q}_\gamma^\varepsilon(w, \mathcal{X})\|_1 \le \frac{1}{\gamma} \int_{\mathbf{R}^2} \left(\frac{\gamma}{\sqrt{\varepsilon}} \bar{P}_{z, \frac{\gamma}{\sqrt{\varepsilon}}} - G(z)\right) \mathscr{J}^\varepsilon(w)(dz, dt)$$

$$+ \frac{1}{\gamma} \int_{\mathbf{R}^2} \|\Theta^\star_{w,t} \mathbf{Q}^\varepsilon_{z,\gamma} - \Theta^\star_{w,t} \mathbf{Q}_\gamma\| \, \tilde{\mathscr{J}}^\varepsilon(w)(dz, dt)$$

$$+ \frac{1}{\gamma} \left\| \int_{\mathbf{R}^2} \Theta^\star_{w,t} \mathbf{Q}_\gamma (\Pi^\star_2 \tilde{\mathscr{J}}^\varepsilon(w)(dt) - \Pi^\star_2 \tilde{\mathscr{J}}^\varepsilon(w)(dt)) \right\|_1$$

$$\leq C_\gamma \varepsilon^{1/5} \mathscr{J}^\varepsilon(w)(\mathbf{R}^2) + C_\gamma \varepsilon^\delta \tilde{\mathscr{J}}^\varepsilon(w)(\mathbf{R}^2)$$

$$+ C_\gamma \varepsilon^{1/10} (1 + \mathrm{Lip}\, \Theta^\star_{w,t} \mathbf{Q}_\gamma) \,,$$

where $\mathrm{Lip}\, \Theta^\star_{w,t} \mathbf{Q}_\gamma$ denotes the quantity

$$\mathrm{Lip}\, \Theta^\star_{w,t} \mathbf{Q}_\gamma = \sup_{t \neq t'} \frac{\|\Theta^\star_{w,t} \mathbf{Q}_\gamma - \Theta^\star_{w,t} \mathbf{Q}_\gamma\|_1}{1 \wedge |t - t'|^{1/3}} \,.$$

It is at this stage that the choice (5.23) of distance function becomes clear: with respect to the usual Euclidean distance, the map $t \mapsto \Theta^\star_{w,t} \mathbf{Q}_\gamma$ would not be Lipschitz continuous. In this way however, it follows immediately from the Hölder continuity of Brownian motion that $\mathrm{Lip}\, \Theta^\star_{w,t} \mathbf{Q}_\gamma < \infty$.

Furthermore, it follows from the definition of $\Omega^{(2)}_\varepsilon$ that $\tilde{\mathscr{J}}^\varepsilon(w)(\mathbf{R}^2)$ is bounded by a constant independent of w and ε. Since $G(x)$ is bounded from below by a constant, this implies that the same is true of $\mathscr{J}^\varepsilon(w)(\mathbf{R}^2)$. Combining these bounds, we then obtain

$$\|\mathcal{Q}_\gamma(w, \mathcal{X}) - \mathcal{Q}^\varepsilon_\gamma(w, \mathcal{X})\|_1 \leq C_\gamma \varepsilon^\delta \,,$$

for some constant C and any $\delta < \frac{1}{16}$.

In order to complete the proof, it thus remains to show that $\inf_{w \in \Omega_\varepsilon} \mathcal{Q}^\varepsilon_\gamma(w, \Omega^c_\varepsilon) \to 0$ sufficiently fast as $\varepsilon \to 0$, where Ω^c_ε denotes the complement of Ω_ε. As a consequence of the definition of Ω_ε, this follows if we can show that $\mathbf{Q}^\varepsilon_{z,\gamma}(\Omega^c_\varepsilon) \to 0$ as $\varepsilon \to 0$, uniformly over all $z \leq \varepsilon^{-1/3}$. Instead of considering $\mathbf{Q}^\varepsilon_{z,\gamma}$, it is much easier to consider $\mathbf{Q}^\varepsilon_z$, the law of the rescaled random walk (4.1) over the time interval $[0, 1]$.

Observe that $\mathbf{Q}^\varepsilon_{z,\gamma}$ is obtained from $\mathbf{Q}^\varepsilon_z$ by first conditioning on the event that γ is reached before the walk becomes negative and then stopping the walk. By Proposition 2.2, the probability of this event is bounded from below by $C\sqrt{\varepsilon}$ for some constant C depending on γ but independent of $z \in [0, \varepsilon^{-1/3}]$. Therefore it is sufficient to find a set Ξ_ε with the following two properties:

- There exists an exponent $\zeta > \frac{1}{2}$ and a constant C such that $\mathbf{Q}^\varepsilon_z(\mathcal{X} \setminus \Xi_\varepsilon) \leq C\varepsilon^\zeta$ for every $\alpha \leq 1$ and every $z \in [0, \varepsilon^{-1/3}]$.
- For every $w \in \Xi_\varepsilon$ and every $t \in [0, 1]$, the path $\tilde{w}$ obtained from w by stopping it at time t belongs to Ω_ε.

In order to determine whether a path w belongs to Ξ_ε, we "coarse-grain" it into pieces of length $\varepsilon^{1/3}$ (the precise exponent is not very important as we do not endeavour to obtain optimal convergence rates) and we impose that the contribution of $\tilde{\mathscr{J}}^\varepsilon(w)$ on each piece is very close to what it should be. In other words, setting $I_k \subset [0, 1]$ by $I_k = [\varepsilon^{1/3} k, \varepsilon^{1/3}(k + 1))$, we define Ξ_ε as

$$\Xi_\varepsilon = \Omega^{(1)}_\varepsilon \cap \left\{ w \, : \, \left| \Pi^\star_2 \tilde{\mathscr{J}}^\varepsilon(w)(I_k) - \frac{a}{2} |I_k| \right| \leq \sqrt{\varepsilon} \right\} \,,$$

where $|I_k|$ denotes the length of the interval I_k.

We first note that if $w \in \Xi_\varepsilon$ and $t \in [0, 1]$, then the path $\tilde{w}$ obtained by stopping w at time t does belong to Ω_ε. Since $\tilde{\mathscr{J}}^\varepsilon(\tilde{w}) \leq \tilde{\mathscr{J}}^\varepsilon(w)$, $w \in \Omega^{(1)}_\varepsilon$ implies that $\tilde{w} \in \Omega^{(1)}_\varepsilon$. Denote now by η^ε the measure $\eta^\varepsilon = \Pi^\star_2 \tilde{\mathscr{J}}^\varepsilon(\tilde{w})$ and by η the target measure, namely

$\eta = \frac{a}{2}\lambda|_{[0,t]}$, where λ denotes the Lebesgue measure. By the assumption on $\tilde{\mathscr{J}}^{\varepsilon}(w)$, it follows that one has $|\eta^{\varepsilon}(I_k) - \eta(I_k)| \le \sqrt{\varepsilon}$ for each of the intervals I_k, except possibly for the interval containing t. As a consequence, denoting by η_k^{ε} and η_k the restrictions of η^{ε} and η to I_k, one obtains for all k such that $t \notin I_k$ the bound

$$\|\eta_k^{\varepsilon} - \eta_k\|_1 \le \varepsilon^{\frac{1}{9} + \frac{1}{3}} + \varepsilon^{\frac{1}{2}} \ .$$

(Recall that we use the distance function (5.23), this is the reason for the exponent $\frac{1}{9}$.) Summing over k and using the fact that $\|\eta_k^{\varepsilon} - \eta_k\|_1 \le C\varepsilon^{1/3}$ for the value k such that $t \in I_k$, it follows that one has $\|\eta^{\varepsilon} - \eta\|_1 \le C\varepsilon^{1/9}$, which indeed implies that $\tilde{w} \in \Omega_{\varepsilon}$, at least for ε small enough.

It remains to show that $\mathbf{Q}_z^{\varepsilon}(\mathcal{E} \setminus \Xi_{\varepsilon}) \le C\varepsilon^{\zeta}$ for sufficiently large ζ. Observe first that $\mathbf{Q}_z^{\varepsilon}(\mathcal{E} \setminus \Omega_{\varepsilon}^{(1)}) \le C\varepsilon^p$ for every $p > 0$ thanks to the fact that the distribution ν of the steps of our random walk has exponential tails.

To obtain the required bound on $\mathbf{Q}_z^{\varepsilon}(\mathcal{E} \setminus \Xi_{\varepsilon})$, we use the fact that $\Pi_2^{\star}\tilde{\mathscr{J}}^{\varepsilon}(w)(I_k)$ consists of the sum of $\varepsilon^{-2/3}$ i.i.d. copies of a random variable Y with law given by

$$Y \stackrel{\text{law}}{=} a\varepsilon \int_0^Z e^{a\sqrt{\varepsilon}z} G(z)\,dz \ , \qquad \mathcal{D}(Z) = \nu \ ,$$

where ν is the one-step distribution of the underlying random walk. Note now that, as a consequence of Proposition 2.4 and the fact that ν admits some exponential moment, one has

$$\mathbf{E}Y = \frac{a}{2}\varepsilon + \mathcal{O}(\varepsilon^{3/2}) \ ,$$

for all ε sufficiently small. Similarly, it is straightforward to check that $\mathbf{E}|Y|^p = \mathcal{O}(\varepsilon^p)$ for every $p > 0$. As a consequence, one has for every $p > 0$ the bound

$$\mathbf{E}\left|\Pi_2^{\star}\tilde{\mathscr{J}}^{\varepsilon}(w)(I_k) - \frac{a}{2}|I_k|\right|^p \le C\varepsilon^{2p/3} \ ,$$

for a constant possibly depending on p. It immediately follows that

$$\mathbf{P}\left(\left|\Pi_2^{\star}\tilde{\mathscr{J}}^{\varepsilon}(w)(I_k) - \frac{a}{2}|I_k|\right| > \sqrt{\varepsilon}\right) \le C\varepsilon^p \ ,$$

for every $p > 0$. Summing over k and combining this with the previously obtained bound on $\mathbf{Q}_z^{\varepsilon}(\mathcal{E} \setminus \Omega_{\varepsilon}^{(1)})$, we conclude that $\mathbf{Q}_z^{\varepsilon}(\mathcal{E} \setminus \Xi_{\varepsilon}) \le C\varepsilon^p$ for every $p \ge 0$, which concludes our proof. □

References

[AFRtW06] R. J. ALLEN, D. FRENKEL, and P. REIN TEN WOLDE. Simulating rare events in equilibrium or non equilibrium stochastic systems. *J. Chem. Phys.* **124**, (2006), 024102.

[AJ06] J. A. ADELL and P. JODRÁ. Exact Kolmogorov and total variation distances between some familiar discrete distributions. *J. Inequal. Appl.* Art. ID 64307, 8.

[Ald91] D. ALDOUS. The continuum random tree. I. *Ann. Probab.* **19**, no. 1, (1991), 1–28.

[Bil99] P. BILLINGSLEY. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, 1999.

[Daw93] D. A. DAWSON. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI—1991*, vol. 1541 of *Lecture Notes in Math.*, 1–260. Springer, Berlin, 1993.

[dDG05] N. DE FREITAS, A. DOUCET, and N. E. GORDON. *Sequential Monte Carlo Methods in Practice*. Springer, 2005.

[DM04] P. DEL MORAL. *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications.

[EK86] S. N. ETHIER and T. G. KURTZ. *Markov Processes*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[FG10] A. FIGALLI and N. GIGLI. A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions. *J. Math. Pures Appl. (9)* **94**, no. 2, (2010), 107–130.

[FINR04] L. R. G. FONTES, M. ISOPI, C. M. NEWMAN, and K. RAVISHANKAR. The Brownian web: characterization and convergence. *Ann. Probab.* **32**, no. 4, (2004), 2857–2883.

[Fra73] D. F. FRASER. The rate of convergence of a random walk to Brownian motion. *Ann. Probability* **1**, (1973), 699–701.

[HM54] J. HAMMERSLEY and K. MORTON. Poor man's Monte Carlo. *J. R. Stat. Soc. B* **16**, no. 1, (1954), 23–38.

[Hut09] M. HUTZENTHALER. The virgin island model. *Electron. J. Probab.* **14**, (2009), no. 39, 1117–1161.

[HW14] M. HAIRER and J. WEARE. Improved diffusion Monte Carlo. *Commun. Pure Appl. Math.* **?**, no. ?, (2014), ??–??

[Itô72] K. ITÔ. Poisson point processes attached to Markov processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, 225–239. Univ. California Press, Berkeley, Calif., 1972.

[Jag75] P. JAGERS. *Branching processes with biological applications*. Wiley-Interscience [John Wiley & Sons], London, 1975. Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics.

[JDMD06] A. JOHANSEN, P. DEL MORAL, and A. DOUCET. Sequential Monte Carlo samplers for rare events. In *Proceedings of the 6th International Workshop on Rare Event Simulation*. Bramberg, 2006.

[LG07] J.-F. LE GALL. The topological structure of scaling limits of large planar maps. *Invent. Math.* **169**, no. 3, (2007), 621–670.

[LG10] J.-F. LE GALL. Geodesics in large planar maps and in the Brownian map. *Acta Math.* **205**, no. 2, (2010), 287–360.

[LL10] G. F. LAWLER and V. LIMIC. *Random walk: a modern introduction*, vol. 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[NR02] D. NUALART and A. RASCANU. Differential equations driven by fractional Brownian motion. *Collect. Math.* **53**, no. 1, (2002), 55–81.

[Pra08] A. PRATELLI. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Math. Z.* **258**, no. 3, (2008), 677–690.

[PY82] J. PITMAN and M. YOR. A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete* **59**, no. 4, (1982), 425–457.

[PZ07]   S. PESZAT and J. ZABCZYK. *Stochastic partial differential equations with Lévy noise*, vol. 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007. An evolution equation approach.

[Rou06]   M. ROUSSET. *Méthods Population Monte-Carlo en Temps Continu pour la Physique Numérique*. Ph.D. thesis, L'Université Paul Sabatier Toulouse III, 2006.

[RW94]   L. C. G. ROGERS and D. WILLIAMS. *Diffusions, Markov processes, and martingales. Vol. 1*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, second ed., 1994. Foundations.

[RY91]   D. REVUZ and M. YOR. *Continuous martingales and Brownian motion*, vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1991.

[SS08]   R. SUN and J. M. SWART. The Brownian net. *Ann. Probab.* **36**, no. 3, (2008), 1153–1208.

[Vil03]   C. VILLANI. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.