AD-A274 696

‖‖‖‖‖‖‖‖‖‖‖

12

SUPER FRACTIONAL BROWNIAN MOTION, FRACTIONAL SUPER BROWNIAN MOTION AND RELATED SELF-SIMILAR (SUPER) PROCESSES

by

Robert J. Adler[1,2] and Gennady Samorodnitsky[1,3]

Faculty of Industrial Engineering & Management
Technion – Israel Institute of Technology

and

School of Operations Research and Industrial Engineering
Cornell Univerity

DTIC
ELECTE
JAN 12 1994
S
A

Abstract

We consider the full weak convergence, in appropriate function spaces, of systems of noninteracting particles undergoing critical branching and following a self similar spatial motion with stationary increments. The limit processes are measure valued, and are of the super and historical process type. In the case in which the underlying motion is that of a fractional Brownian motion, we obtain a characterisation of the limit process as a kind of stochastic integral against the historical process of a Brownian motion defined on the full real line.

94-01196
‖‖‖‖‖‖‖‖‖‖

94 1 10 162

1. INTRODUCTION

In this section we shall start with as non-technical as possible, and therefore occasionally imprecise, a description of the main ideas and results of this paper. All the requisite qualifying details will appear in the following sections.

(a) The usual setting. Superprocesses arise as infinite density limits of systems of branching Markov processes. We start with a parameter $\mu > 0$ that will eventually become large, a finite measure m on $\Re^d$, and a Poisson point process Π^μ on $\Re^d$ with intensity measure μm. The $K = O(\mu)$ random points, $x_0, \ldots, x_K$ of the Poisson process will be the initial positions of a system of particles. Each of these K particles follows the path of independent copies of a Markov process Y, until time $t = 1/\mu$.

At time $1/\mu$ each particle, independently of the others, dies and leaves behind itself a random number of "children", with the family sizes of the different "parent" particles independently distributed. The mean family size is one. The individual particles in the new population then follow independent copies of Y, starting at their place of birth, in the interval $[1/\mu, 2/\mu)$, and the pattern of alternating critical branching and spatial spreading continues until, with probability one, there are no particles left alive. It is clear that this is a purely Markovian system.

The basic process of interest is the measure valued Markov process

$$X_t^\mu(A) \;=\; \frac{\{\text{Number of particles in } A \text{ at time } t\}}{\mu}, \qquad (1.1)$$

where $A \in \mathcal{B}^d =$ Borel sets in $\Re^d$. Note that, for fixed t and μ, X_t^μ is a purely atomic measure.

It is now well known that under very mild conditions on Y the sequence $\{X^\mu\}_{\mu \geq 1}$ converges weakly, as $\mu \to \infty$, and on an appropriate Skorohod space, to a measure valued process which is called the superprocess for Y. If, for example, Y is a Brownian motion with generator Δ, the limit process is the super Brownian motion. Details can be found, for example, in Ethier and Kurtz (1986), Walsh (1986), and the major recent review by Dawson (1993).

(b) A non-Markovian model. The motivation for this paper is to investigate the above problem when the underlying motion of the particles is non-Markovian. This problem can be divided into two distinct subproblems, depending on what happens at the birth/death times k/μ.

The simpler of the two cases arises when a general non-Markov motion is allowed, but at birth/death times each of the new particles begins a motion independent of those of its ancestors, and of those of other particles, other than for the fact that it begins its motion at the deathplace of its immediate parent. Much of the Markov structure is thus preserved. We do not consider this case in the present paper.

In the more interesting case, the one which we do consider, the particles follow the talmudic dictum of "Know from whence you have come", (Talmud, *circa* 400), and at regeneration times retain some memory of what their ancestors did in the past. We now begin with a more precise setup.

2

We append to the state space $\mathfrak{R}^d$ a cemetery state Λ, and adopt the convention that $\phi(\Lambda) \equiv 0$ for all functions $\phi\colon \mathfrak{R}^d \to \mathfrak{R}^k$. We fix $0 < \mu < \infty$, and recall that $K = O(\mu)$ is the number of particles alive at time zero.

In order to label our particles, define the family of multi-indices

$$I := \{\alpha = (\alpha_0, \alpha_1 \ldots, \alpha_N)\colon \alpha_0 \in \mathcal{N}, \ \alpha_i \geq 1, \ i \geq 1, \ N \geq 0\}. \tag{1.2}$$

Define the "length" of α by $|\alpha| = N$, and set $\alpha|_i = (\alpha_0, \ldots, \alpha_i)$ and $\alpha - i = (\alpha_0, \ldots, \alpha_{|\alpha|-i})$.

Induce a partial order on I by setting

$$\beta \prec \alpha \iff \beta = \alpha|_i \ \text{ for some } i \leq |\alpha|, \tag{1.3}$$

and, for any $t > 0$, write $\alpha \sim t$, if, and only if

$$\frac{|\alpha|}{\mu} \leq t < \frac{1 + |\alpha|}{\mu}. \tag{1.4}$$

Furthermore, for two indices α and β, set

$$\tau^{\alpha,\beta} = \max\{k\colon \alpha_i = \beta_i, \text{ for all } i \leq k\}, \tag{1.5}$$

and write $\alpha \wedge \beta = \alpha|_{\tau^{\alpha,\beta}} = \beta|_{\tau^{\alpha,\beta}}$. For each $\alpha \in I$, let $\{Y^\alpha(t), \ t \geq 0\}$, be a copy of the generic process Y with $Y_0 \equiv 0$, chosen so that the family $\{Y^\alpha\}_{\alpha \in I}$ satisfies

$$Y^\alpha(t) \equiv Y^{\alpha-i}(t), \tag{1.6}$$

for all $1 \leq i \leq |\alpha|$. Again, for each $\alpha \in I$, let $\{\mathcal{F}_t^\alpha\}_{t \geq 0}$ be a filtration. We require that $\mathcal{F}_t^\alpha \subseteq \mathcal{F}_t^\beta$, for all $t \geq 0$, whenever $\alpha \prec \beta$, and that $Y^\alpha(t)$ is $\mathcal{F}_t^\alpha$ measurable. In general, we think of $\mathcal{F}_t^\alpha$ as the σ-algebra generated by $\{Y^\alpha(s) \ s \leq t\}$, but, even in the most interesting of our examples to follow, this will not be the case.

Furthermore, for any $\alpha, \beta \in I$, and $t > \tau^{\alpha,\beta}/\mu$, we shall require that $Y^\alpha(t)$ and $Y^\beta(t)$ are conditionally independent given $\mathcal{F}_{\tau^{\alpha,\beta}/\mu}^{\alpha \wedge \beta}$. (When $\mathcal{F}_t^\alpha$ is the σ-algebra generated by Y^α, this is equivalent to requiring conditional independence given $\{Y^{\alpha \wedge \beta}(s)\colon 0 \leq s \leq \tau^{\alpha,\beta}/\mu\}$; i.e. Children recall their parent's behaviour, but, given this, ignore their peers.)

It is not hard, although slightly tedious, to show via a standard Kolmogorov-type construction as in Theorem 2.7.2 of Ash (1972) that such a system is well defined. This almost defines our particle system, but we are not quite finished, for we have yet to introduce random branching. This we do by thinning out the full tree of particles. For each $\alpha \in I$, let N^α be an independent copy of a generic non-negative random variable N, which represents the size of families at branching. We assume throughout that $EN = 1$. Now define the stopping times

$$\tau^\alpha = \begin{cases} 0 & \text{if } \alpha_0 > K \\ \min_{0 \leq i \leq |\alpha|} \left\{ \frac{i+1}{\mu}\colon N^{\alpha|_i} = 0 \right\} & \text{if this set} \neq \emptyset \text{ and } \alpha_0 \leq K \\ \frac{1+|\alpha|}{\mu} & \text{otherwise.} \end{cases}$$

3

We can now finally define the particle paths of interest to us by setting, for each $\alpha \in I$,

$$X^\alpha(t) = \begin{cases} x_{\alpha_0} + Y^\alpha(t) & \text{if } t < \tau^\alpha \\ \Lambda & \text{if } t \geq \tau^\alpha. \end{cases} \tag{1.7}$$

Thus X^α is only "alive" (i.e. $\neq \Lambda$) at time t if it has a continuous stream of ancestors.

The definition (1.1) of our basic measure valued process for the finite system is now more precisely definable as

$$X_t^\mu(A) = \frac{\#\{X_t^\alpha \in A : \alpha \sim t\}}{\mu}. \tag{1.8}$$

We shall also find it useful to construct another measure-valued process associated with the above branching system of moving particles. Assume that the generic process Y has sample paths in $D[0,\infty)$, the space of $\Re^d$-valued cadlag functions (equipped with the Skorohod J_1 topology). Furthermore, for any function $f : [0,\infty) \to \Re^d$ and $t > 0$ define a new function $f^{(t)}$ by $f^{(t)}(s) = f(t \wedge s)$. Then the *historical process* for the above particle system is well defined as

$$H_t^\mu(A) = \frac{\#\{(X^\alpha)^{(t)} \in A : \alpha \sim t\}}{\mu}, \tag{1.9}$$

where now A is a Borel set in $D([0,\infty), \Re^d)$.

If $A \in \mathcal{B}^d$, then $A_t := \{f \in D[0,\infty) : f(t) \in A\}$ is a Borel set in $D[0,\infty)$. Thus it is immediate that $X_t^\mu(A) = H_t^\mu(A_t)$, and so, in principle, knowledge of the historical process implies knowledge of the basic process X_t^μ. (The converse is not generally true, although for special processes, in high dimensions, it is. cf. Barlow and Perkins, (1993).)

(c) Self-similar processes. To get some interesting limit theory from the set-up described above, we shall have to make some assumptions about the structure of the generic process Y. In particular, we shall assume that $\{Y(t), t \geq 0\}$ is an H-self similar $\Re^d$-valued stochastic process with stationary increments. That is, for any $n \geq 1$, $0 \leq t_1 < t_2 < \ldots < t_n$, $h \geq 0$ and $c > 0$, there exists a real H such that

$$\big(Y(ct_1), \ldots, Y(ct_n)\big) \overset{\mathcal{L}}{=} c^H \big(Y(t_1), \ldots, Y(t_n)\big) \tag{1.10}$$

(H-self similarity) and

$$\begin{aligned} \big(Y(t_2 + h) - Y(t_1 + h), Y(t_3 + h) - Y(t_2 + h), \ldots, Y(t_n + h) - Y(t_{n-1} + h)\big) \\ \overset{\mathcal{L}}{=} \big(Y(t_2) - Y(t_1), Y(t_3) - Y(t_2), \ldots, Y(t_n) - Y(t_{n-1})\big) \end{aligned} \tag{1.11}$$

(stationarity of the increments). The common alias for "H-self similar with stationary increments" is "H-sssi", and we shall use it in the sequel. We shall also assume that

$$0 < H < 1, \text{ and } E\|Y(1)\|^p < \infty \text{ for some } p > 1/H. \tag{1.12}$$

It is worth noting that mere existence of the first moment of $Y(1)$ forces H to be in $(0,1]$, and $H = 1$ leads to a degenerate process (Vervaat 1985). Note further that the H-sssi property

of Y implies that $E\|Y(t) - Y(s)\|^p = a|t - s|^{pH}$ for every $t, s \geq 0$ with $a = E\|Y(1)\|^p$, and, of course, that $Y(0) \equiv 0$.

An immediate consequence of the moment assumption (1.12) is that Y has a version with continuous sample paths. This can be checked in a variety of ways, one of the simplest being via a version of Kolomogorov-Loéve theorem, (e.g. Billingsley (1968), Theorem 12.4). We shall assume throughout therefore that all copies of Y have continuous sample paths. When a component Y_i of $Y \in \Re^d$ is a one-dimensional Gaussian process, it is the well known *fractional Brownian motion* and so can be represented as

$$Y_i(t) = k_i \int_{-\infty}^{t} \left((t - r)^{H-1/2} - (-r)_+^{H-1/2}\right) B(dr), \quad t \geq 0 \tag{1.13}$$

for $H \neq 1/2$, where k_i is a constant and B is a standard Brownian motion on $(-\infty, \infty)$. For $H = 1/2$, (1.13) is not really very informative, and Y_i is simply a constant multiple of standard Brownian motion. It is easy to check (cf. (1.15) below) that the covariance function of fractional Brownian motion is given by

$$\text{Cov}(Y_i(t), Y_i(s)) = \frac{1}{2}\text{Var}(Y_i(1))\left[t^{2H} + s^{2H} - |t - s|^{2H}\right]. \tag{1.14}$$

It is also straightforward to check that *every* (one dimensional) H-sssi process with finite second moments has the same covariance function (1.14) as the fractional Brownian motion. However, these are not the only examples, and there do exist non-Gaussian one dimensional H-sssi processes with finite second moments. (In fact, with all moments finite; cf. Surgailis (1981).)

Our main weak convergence result concerning measure valued processes (1.8) is stated and proved in Section 2, under the dependence structure outlined in Section 1(b) above, and the assumptions that the generic particle motion Y is H-sssi, satisfying, together with the branching mechanism N, an appropriate moment condition. The result parallels the weak convergence to the superprocess known in the purely Markov case.

(d) More on fractional Brownian motion. Since the general convergence results just described do not require any particular structure for the particle motion Y, it is not possible to say very much about the limiting measure-valued process. To do so, we specialize to the particular case of the particles moving according to the d-dimensional fractional Brownian motion. (i.e. each component of Y satisfies (1.13) for some k_i, and the same H.) Given the general terminology of superprocesses, the measure valued limit process could well be termed a "super fractional Brownian motion". What we shall establish is a representation for this limiting process which is related to the usual super Brownian motion.

To do this, we require a system of branching Brownian motions on $(-\infty, \infty)$, so as to exploit (1.13) to give a compact and useful representation of the particle system.

Retaining the indexing notation developed above, let $\{\{W^\alpha(t)\}_{t \geq 0}\}_{\alpha \in I}$ be a system of branching Brownian motions, all starting at the origin. Recall that at $t = 0$ there are K initial points for this system. Let $\tilde{W}^1, \ldots, \tilde{W}^K$ be K extra Brownian motions in positive time, independent of one another and of the initial system. For $\alpha \in I$ define the process

$$B^\alpha(t) = 1_{[t>0]}(t)W^\alpha(t) + 1_{[t \leq 0]}(t)\tilde{W}^{\alpha_0}(-t), \quad -\infty < t < \infty. \tag{1.15}$$

5

Then each B^α is a Brownian motion on the entire real line, centered by setting $B^\alpha(0) \equiv 0$. Now set, for each $\alpha \in I$

$$Y^\alpha(t) = \int_{-\infty}^t ((t-r)^{H-1/2} - (-r)_+^{H-1/2}) B^\alpha(dr), \qquad t > 0. \tag{1.16}$$

Finally, define the thinned system $\{X^\alpha\}$ as at (1.7). However, to simplify the notation for the moment, assume that all the initial points x_{α_0} are zero. That is, all the paths X^α start from the origin of $\Re^d$. It is obvious that the branching system so constructed fulfills all the requirements of the previous section, with the filtrations $\mathcal{F}_t^\alpha$, $t \geq 0$ taken to be $\sigma\{B^\alpha(s): -\infty < s \leq t\}$.

Now, however, we have far more structure than we had before.

To see this, let H_t^μ be the historical process for the B^α system, so that for any Borel set A of $C(\Re, \Re^d)$ and $t \geq 0$

$$H_t^\mu(A) = \frac{1}{\mu} \sum_{\alpha \sim t} 1_A(B^\alpha). \tag{1.17}$$

Taking the integrated form of the measure (1.8), it is now immediate that, for any nice test function $\phi: \Re^d \to \Re$,

$$
\begin{aligned}
X_t^\mu(\phi) :&= \int_{\Re^d} \phi(x) X_t^\mu(dx) \\
&= \frac{1}{\mu} \sum_{\alpha \sim t} \phi(X_t^\alpha) \\
&= \frac{1}{\mu} \sum_{\alpha \sim t} \phi\left(\int_{-\infty}^t ((t-r)^{H-1/2} - (-r)_+^{H-1/2}) B^\alpha(dr) \right) \\
&= \int_{C(\Re, \Re^d)} \phi\left(\int_{-\infty}^t ((t-r)^{H-1/2} - (-r)_+^{H-1/2}) y(dr) \right) H_t^\mu(dy).
\end{aligned}
\tag{1.18}
$$

(It is easy to pass from line to line here if one keeps in mind that, with the exception of the one stochastic integral, all other integrals are really just finite sums.)

In Theorem 3.1 we shall show that, at least for $H > \frac{1}{2}$, it is possible to pass to the $\mu \to \infty$ limit in (1.18). Since the final line in (1.18) is then a functional of the historical super process for Brownian motion (albeit on all of $\Re$ rather than just $\Re_+$), we obtain super fractional Brownian motion as a function of super (historical) Brownian motion, giving us the main result of Section 3, and title of the paper.

It is clear that the explicit representation (1.18) allows for a far clearer understanding of the limit process than does the more general procedure which we descibed earlier, in which nothing other than the existence of a limit process is established. We should note, however, that the previous subsection covers a much wider range of processes than mere fractional Brownian motion. In fact, a representation like (1.18) would seem to be possible only for fractional Brownian motion, or other processes representable as functionals of systems of branching Markov processes. The approach taken in Section 2 works in far greater generality.

As we shall see in Section 3 below, (1.18) only seems to work, in the limit, for $H > \frac{1}{2}$. While it is not totally clear whether this is an artifact of the proof or not, we believe that the result may not be true for smaller H. We shall explain why in Section 3.

2. WEAK CONVERGENCE FOR GENERAL SELF-SIMILAR PROCESSES

Throughout this section we shall assume the branching structure, with conditional independence of new particles, described in Section 1(b). We shall also assume that the particle paths are those of H-ssii processes, and that the Poisson distribution of initial points is governed by a finite intensity measure m.

Let $M_F(\mathfrak{R}^d)$, endowed with the topology of weak convergence, be the space of finite Radon measures on $\mathfrak{R}^d$, and $C([0,T), M_F(\mathfrak{R}^d))$ and $D([0,T), M_F(\mathfrak{R}^d))$ the spaces of continuous and cadlag functions from $[0,T]$ to $M_F(\mathfrak{R}^d)$. When $T = \infty$ we denote these more briefly by $C(M_F(\mathfrak{R}^d))$ and $D(M_F(\mathfrak{R}^d))$

The following is the main convergence result of this section.

Theorem 2.1 *Suppose that for some $p > 2$ both $E\|Y(1)\|^p < \infty$ and $EN^p < \infty$. If $H > 1/p$, then the sequence $\{X^\mu\}_{\mu=1}^\infty$ of $M_F(\mathfrak{R}^d)$-valued processes converges weakly in $D(M_F(\mathfrak{R}^d))$, as $\mu \to \infty$, to a $M_F(\mathfrak{R}^d)$-valued process X. Furthermore, X has a version with all sample paths in $C(M_F(\mathfrak{R}^d))$.*

Remark: We shall see below that the finite-dimensional distributions of X^μ converge without assuming existence of any moment higher than second for the branching distribution and for *every $H \in (0,1)$*. The extra assumptions are required to establish tightness. We do not know, however, whether the condition $H > 1/p$ is necessary, or only a function of our method of proof.

At an intuitive level, the qualitative aspects of this requirement are readily understood. Lack of smoothness at the level of X^μ can be due to one of two sources; irregularities in the particle paths, or large spurts of population growth. Recalling that the sample paths of the process Y become rougher as H decreases, the need to balance the moments of Y and N against the smoothness parameter seems more than reasonable.

Proof of Theorem 2.1. The proof is rather long, and so will be broken up into a number of stages.

Initially we shall prove that the sequence $\{X^\mu\}_\mu$ converges weakly in $D(M_F(\overline{\mathfrak{R}^d}))$ to a stochastic process X which has a version with all sample paths in $C(M_F(\overline{\mathfrak{R}^d}))$, where $\overline{\mathfrak{R}^d}$ is the one-point compactification of $\mathfrak{R}^d$. Only at the very end shall we move back to $D(M_F(\mathfrak{R}^d))$.

To establish the weak convergence of $\{X^\mu\}_\mu$ we start with the usual reduction:

Let C_l be the collection of real valued continuous functions on $\mathfrak{R}^d$ which have a continuous extension to the one-point compactification of $\mathfrak{R}^d$ (that is, the functions with a finite limit at ∞). We claim that it is enough to prove that for every $\varphi \in C_l$ the sequence $\{X^\mu(\varphi)\}_\mu$ converges weakly in $D(\mathfrak{R})$ to a process which has a version with all sample paths in $C(\mathfrak{R})$. It will then follow from Theorem 3.7.1 of Dawson (1993) that the sequence $\{X^\mu\}_\mu$ is tight in $D(M_F(\overline{\mathfrak{R}^d}))$. To establish convergence of finite dimensional distributions, note that since the limiting process for each $\{X^\mu(\varphi)\}_\mu$ has continuous sample paths and $\overline{\mathfrak{R}^d}$ is compact, every limit point of the sequence $\{X^\mu\}_\mu$ must have a continuous version. It then follows from the assumed convergence and Theorem 3.7.8(a) of Ethier and Kurtz (1986) that for every

$t_1 \geq 0, \ldots, t_p \geq 0$ and $\varphi \in C_l$ the sequence $\{(X_{t_1}^\mu(\varphi), \ldots, X_{t_p}^\mu(\varphi))\}_\mu$ converges weakly in $\Re^p$, and so the sequence $\{(X_{t_1}^\mu, \ldots, X_{t_p}^\mu)\}_\mu$ converges weakly in $(M_F(\overline{\Re^d}))^p$. A final application of Theorem 3.7.8(b) of Ethier and Kurtz (1986) shows that the sequence $\{X^\mu\}_\mu$ cannot have more than one limiting point. This gives the full weak convergence that we require.

We therefore now fix $\varphi \in C_l$ and consider the sequence $\{X^\mu(\varphi)\}_\mu$ of $D(\Re)$-valued processes. It is obviously enough to consider a function φ with a bounded first derivative (such functions are dense in C_l; see Theorem 3.7.1. of Dawson(1993)), and henceforth we shall consider only such functions. We start by establishing the convergence of finite dimensional distributions of this sequence. The proof is a reasonably straightforward application of Lemma 3.1 of Dynkin (1991).

We shall give enough details here to allow the reader to check that Dynkin's theorem applies to our situation, without repeating that result in full detail. However, it is worth noting that Dynkin's result gives convergence of finite dimensional distributions for branching systems of not-necessarily Markovian systems in an extremely general setting. His result is therefore crucial in the following argument. It does not, however, give full weak convergence, (i.e. in the function space setting which includes, of necessity, tightness of some kind). Furthermore, in the non-Markov setting, it works only at the level of the historical process.

Choose any $t_1, \ldots, t_p$ in $[0, \infty)$. We define p continuous real valued functions on $C(\Re, \Re^d)$ by

$$f_i(y) = \varphi(y(t_i)), \quad i = 1, \ldots, p.$$

Observe that for every $\mu \geq 1$

$$(X_{t_1}^\mu(\varphi), \ldots, X_{t_p}^\mu(\varphi)) = (H_{t_1}^\mu(f_1), \ldots, H_{t_p}^\mu(f_p))$$

where H^μ is the historical particle process (1.9) corresponding to X^μ. Apply now Lemma 3.1 of Dynkin (1991) to the sequence $\{H^\mu\}_\mu$ with, in his notation, $K(dt) = dt$, branching mechanism $\varphi_\mu^t(x, z)$ independent of t, μ and x and, finally, with the additive functional A given by

$$A(B, \omega) = \sum_{i=1}^p \mathbf{1}_{t_i \in B} f_i(\omega)$$

for a Borel set B and $\omega \in C(\Re, \Re^d)$. The application is straightforward: simply treat H^μ as a regular (non-historical) superprocess with respect to the time inhomogeneous Markov process $\hat{Y}$ defined by

$$\hat{Y}(t) = (t, Y(s), s \leq t).$$

(A detailed example of how to do this is given in the following section, where we have more need for rigour.) Note that for each $t \geq 0$, $\hat{Y}(t)$ is $\Re_+ \times C(\Re, \Re^d)$-valued. Since Dynkin's result does not require any regularity assumptions on the Markov process, we conclude immediately that the sequence

$$\{(H_{t_1}^\mu(f_1), \ldots, H_{t_p}^\mu(f_p)), \ \mu = 1, 2, \ldots\}$$

converges weakly in $\Re^p$ as $\mu \to \infty$, and, therefore, so does the sequence

$$\{(X_{t_1}^\mu(\varphi), \ldots, X_{t_p}^\mu(\varphi)), \ \mu = 1, 2, \ldots\}.$$

This concludes our treatment of the finite dimensional distributions. (Note, by the way, that we have not been able to say anything useful about their structure!) We now proceed to establish tightness for the sequence $\{X^\mu(\varphi)\}_\mu$ in $D(\Re)$. This, in itself, is a rather long calculation that needs to be broken up into a number of steps.

A simple extension of Theorem 7.2 of Ethier and Kurtz (1986) (and the fact that φ is bounded) show that it is enough to prove that for every $\epsilon > 0$ and $T > 0$ there is a $\delta > 0$ such that

$$\overline{\lim_{\mu \to \infty}} P\Big(\omega\big(X^\mu(\varphi), \delta, T\big) \geq \epsilon\Big) \leq \epsilon, \tag{2.1}$$

where the uniform oscillation function ω is defined by

$$\omega\big(f, \delta, T\big) = \sup_{\substack{0 \leq s, t \leq T \\ |t-s| \leq \delta}} |f(t) - f(s)|. \tag{2.2}$$

As a first step, we study a discrete version of the oscillation (2.2), which is somewhat easier to deal with. For $\mu = 1, 2, \ldots, \delta > 0$ and $M \geq 1$ consider

$$\omega_\mu\big(f, \delta, M\big) = \max_{\substack{i_1, i_2 \leq M\mu \\ |i_1 - i_2| \leq \delta\mu}} |f(i_1/\mu) - f(i_2/\mu)|. \tag{2.3}$$

We shall handle this oscillation function via metric entropy methods, for which we need to start with some moment calculations. For $0 \leq i_1 < i_2 \leq M\mu$, write

$$X^\mu_{i_1/\mu}(\varphi) - X^\mu_{i_2/\mu}(\varphi) = \mu^{-1} \sum_{\alpha \sim i_1/\mu} \varphi\Big(Y^\alpha\big(i_1/\mu\big) + x_{\alpha_0}\Big) \mathbf{1}\Big(i_1/\mu < \tau\big(\alpha\big)\Big)$$

$$- \mu^{-1} \sum_{\alpha \sim i_2/\mu} \varphi\Big(Y^\alpha\big(i_2/\mu\big) + x_{\alpha_0}\Big) \mathbf{1}\Big(i_2/\mu < \tau\big(\alpha\big)\Big)$$

$$= \mu^{-1} \sum_{\alpha \sim i_1/\mu} \varphi\Big(Y^\alpha\big(i_1/\mu\big) + x_{\alpha_0}\Big)\Big(1 - N_\alpha(i_1, i_2)\Big) \mathbf{1}\Big(i_1/\mu < \tau\big(\alpha\big)\Big)$$

$$+ \Bigg[\mu^{-1} \sum_{\alpha \sim i_1/\mu} \varphi\Big(Y^\alpha\big(i_1/\mu\big) + x_{\alpha_0}\Big) N_\alpha(i_1, i_2) \mathbf{1}\Big(i_1/\mu < \tau\big(\alpha\big)\Big)$$

$$- \mu^{-1} \sum_{\alpha \sim i_2/\mu} \varphi\Big(Y^\alpha\big(i_2/\mu\big) + x_{\alpha_0}\Big) \mathbf{1}\Big(i_2/\mu < \tau\big(\alpha\big)\Big)\Bigg]$$

$$:= \Delta_1(i_1, i_2) + \Delta_2(i_1, i_2), \tag{2.4}$$

where for $\alpha \sim i_1/\mu$, $N_\alpha(i_1, i_2)$ is the number of "descendants" of Y^α alive at the time i_2/μ. Observe that

$$\Delta_2(i_1, i_2) = \mu^{-1} \sum_{\alpha \sim i_2/\mu} \Big(\varphi\Big(Y^\alpha\big(i_1/\mu\big) + x_{\alpha_0}\Big) - \varphi\Big(Y^\alpha\big(i_2/\mu\big) + x_{\alpha_0}\Big)\Big) \mathbf{1}\Big(i_2/\mu < \tau\big(\alpha\big)\Big)$$

$$:= \sum_{j=1}^{N(i_2/\mu)} D_j, \tag{2.5}$$

where $N(i_2/\mu)$ is the total number of particles alive at time i_2/μ.

We shall use two simple inequalities which can be easily checked from the first principles. The first is a version of Burkholder's inequality. Let $X_1, X_2, \ldots$ be independent zero mean random variables such that $E|X_i|^p < \infty$ for every $i \geq 1$. Then, letting c be a finite positive constant, in general dependent on p, that is allowed to change from line to line,

$$E|\sum_{j=1}^{m} X_j|^p \leq c\left(\left(\sum_{j=1}^{m} \sigma_j^2\right)^{p/2} + \sum_{j=1}^{m} E|X_j|^p\right), \tag{2.6}$$

for every $m \geq 1$, where $\sigma_j^2 := EX_j^2$. Secondly, let N_n be a critical branching process starting with a single individual, with a progeny distribution having a finite p^{th} moment for some $2 \leq p \leq 4$ (say). Then there is a $c \in (0, \infty)$ such that for any $n \geq 1$

$$EN_n^p \leq cn^{p-1}. \tag{2.7}$$

Observe that

$$E|\Delta_2(i_1, i_2)|^p \leq c\left(E|\sum_{j=1}^{N(i_2/\mu)} \left(D_j - ED_j\right)|^p + |ED_1|^p EN(i_2/\mu)^p\right). \tag{2.8}$$

Write

$$E|\sum_{j=1}^{N(i_2/\mu)} \left(D_j - ED_j\right)|^p := E|\sum_{k=1}^{N(0)} S_k|^p,$$

where S_k is the sum of all the terms under the summation in the left hand side corresponding to a single progenitor existing at time 0 and $N(0)$ is the initial number of particles. Then, by (2.6), we have that

$$E\left|\sum_{j=1}^{N(i_2/\mu)} \left(D_j - ED_j\right)\right|^p \leq c\left(E\left(\sum_{k=1}^{N(0)} \mathrm{Var}(S_k)\right)^{p/2} + E\sum_{k=1}^{N(0)} E|S_k|^p\right)$$
$$= c\left(EN(0)^{p/2}\left(\mathrm{Var}(S_1)\right)^{p/2} + EN(0)E|S_1|^p\right). \tag{2.9}$$

We now evaluate the moments of S_1. Let N_i^j be the size of the ith generation of the jth particle from the first generation. For any $2 \leq a \leq 4$ we have

10

$$E|S_1|^a := E|\sum_{j=1}^{N_{i_2}^1}(D_j - ED_j)|^a$$

$$\leq E\Big(E_D|\sum_{j=1}^{N_{i_2}^1}(D_j - ED_j)|^4\Big)^{a/4}$$

$$\leq E\Big(\sum_{j_1=1}^{N_{i_2}^1}\cdots\sum_{j_4=1}^{N_{i_2}^1}E_D\big((D_{j_1}-ED_{j_1})\ldots(D_{j_4}-ED_{j_4})\big)\Big)^{a/4} \qquad (2.10)$$

$$\leq cE\Big(\sum_{j_1=1}^{N_{i_2}^1}\cdots\sum_{j_4=1}^{N_{i_2}^1}|i_1/\mu - i_2/\mu|^{4H}\Big)^{a/4}$$

$$= c|i_1/\mu - i_2/\mu|^{aH}E[N_{i_2}^1]^a$$

$$\leq c|i_1/\mu - i_2/\mu|^{aH}\mu^{a-1},$$

where E_D stands for expectation taken with respect to the D_j's. Applying (2.10) first with $a = 2$ and then with $a = p$ we conclude from (2.9) that

$$E|\sum_{j=1}^{N(i_2/\mu)}\big(D_j - ED_j\big)|^p \leq c|i_1/\mu - i_2/\mu|^{pH}\mu^p. \qquad (2.11)$$

To complete a bound for (2.8) we need to evaluate the moments of $N(i_2/\mu)$ appearing there. Retaining the same notation, a repeated application of (2.6) and (2.7) (in the notation of (2.7)) yields that

$$E\big(N(i_2/\mu)^p\big) \leq cE\Big(N(0)^p + \Big(\sum_{j=1}^{N(0)}E\big(N_{i_2}^j - 1\big)^2\Big)^{p/2} + \sum_{j=1}^{N(0)}E|N_{i_2}^j - 1|^p\Big) \leq c\mu^p. \qquad (2.12)$$

It follows now from (2.8), (2.11) and (2.12) that

$$E|\Delta_2(i_1, i_2)|^p \leq c|i_1/\mu - i_2/\mu|^{pH}. \qquad (2.13)$$

We estimate the pth moment of $\Delta_1(i_1, i_2)$ in a similar manner:

$$E|\Delta_1(i_1, i_2)|^p \leq c\mu^{-p}E\Big(\Big(\sum_{j=1}^{N(i_1/\mu)}E\big(N_{i_2-i_1}^j - 1\big)^2\Big)^{p/2} + \sum_{j=1}^{N(i_1/\mu)}E|N_{i_2-i_1}^j - 1|^p\Big)$$

$$\leq c\mu^{-p}\Big[(i_2 - i_1)^{p/2}EN(i_1/\mu)^{p/2} + (i_2 - i_1)^{p-1}EN(i_1/\mu)\Big] \qquad (2.14)$$

$$\leq c\Big(|i_1/\mu - i_2/\mu|^{p/2} + |i_1/\mu - i_2/\mu|^{p-1}\Big),$$

where at the last step we used (2.12). It follows from (2.4), (2.13) and (2.14) that for any $0 \leq i_1 < i_2 \leq M\mu$

$$E|X_{i_1/\mu}^\mu(\varphi) - X_{i_2/\mu}^\mu(\varphi)|^p \leq c\Big(|i_1/\mu - i_2/\mu|^{p/2} + |i_1/\mu - i_2/\mu|^{pH}\Big). \qquad (2.15)$$

We can now use our moment estimates and a metric entropy argument to bound the oscillation function. On the set $S^\mu := \{0, 1/\mu, \ldots, M - 1/\mu, M\}$ define a metric d^μ by setting

$$d^\mu(i_1/\mu, i_2/\mu) = \left(E|X^\mu_{i_1/\mu}(\varphi) - X^\mu_{i_2/\mu}(\varphi)|^p\right)^{1/p}.$$

For an $\epsilon > 0$ let $N(\epsilon)$ be the smallest number of d^μ-balls of radii not exceeding ϵ needed to cover the whole of S^μ. (For ease of notation, we shall not explicitly display the dependence of N on μ.) It follows from (2.15) that

$$N(\epsilon) \leq c\epsilon^{-\max(2, 1/H)}.$$

In particular, this implies that there is a constant $K < \infty$ such that for every $\mu \geq 1$

$$\int_0^M \left(N(\epsilon)\right)^{1/p} d\epsilon \leq K.$$

It now follows from Theorem 11.6 of Ledoux and Talagrand (1991) that there is a function $h: (0, 1) \to (0, 1)$ such that $h(\delta) \to 0$ as $\delta \to 0$ and such that for every $\mu \geq 1$ and $\delta \in (0, 1)$

$$E \max_{\substack{i_1, i_2 \leq M\mu \\ |i_1 - i_2| \leq \delta\mu}} |X^\mu_{i_1/\mu}(\varphi) - X^\mu_{i_2/\mu}(\varphi)| \leq h(\delta). \tag{2.16}$$

An application of Markov's inequality now establishes a discrete version of (2.1): for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$\varlimsup_{\mu \to \infty} P\left(\omega_\mu\left(X^\mu(\varphi), \delta, T\right) \geq \epsilon\right) \leq \epsilon. \tag{2.17}$$

It is easy now to bridge the gap between (2.17) and (2.1): observe that for any $t, s \in [0, T]$ we can place t in some interval $[i_1/\mu, i_1 + 1/\mu]$ and s in some interval $[i_2/\mu, i_2 + 1/\mu]$. Then an argument very similar to that leading to (2.15) shows that there is a $c \in (0, \infty)$ such that for every $t, s \in [0, T]$

$$E|X^\mu_t(\varphi) - X^\mu_s(\varphi)|^p \leq c\left(|t - s|^{p/2} + |t - s|^{pH}\right), \tag{2.18}$$

and then (2.1) follows as before by an application of Markov inequality. This completes the proof of the fact that $\{X^\mu\}_\mu$ converges weakly in $D(M_F(\overline{\Re^d}))$ to a stochastic process X which has a version with all sample paths in $C(M_F(\overline{\Re^d}))$.

It remains to prove that the limiting process X is supported by $\Re^d$, in the sense that there is an event Ω_+ of probability 1 such that for every $t \geq 0$ and $\omega \in \Omega_+$ we have $X_t(\{\infty\}) = 0$.

We may and shall restrict t to be in the interval $[0, T]$ for some $T > 0$. For a fixed $r > 0$ consider a bounded Hölder function φ_r on $\Re^d$ such that $0 \leq \varphi_r(x) \leq 1$ for any $x \in \Re^d$, with $\varphi_r(x) = 1$ if $\|x\| > r$, $\varphi_r(x) = 0$ if $\|x\| < r/2$, and with a Hölder constant δ, say, fixed. ($\delta > 0$.) (Such a function always exists for r large enough, once δ has been fixed.)

Of course, (2.16) and (2.18) apply fully to φ_r, and since the constant c appearing there depends only on $\max_{x \in \Re^d} |\varphi_r(x)|$ and on the Hölder constant of φ_r, we conclude that for

12

every $t, s \in [0, T]$, and any r large enough, (2.18) holds, and the constant c does not depend on r. Writing

$$d_r^\mu(t, s) = \left(E|X_t^\mu(\varphi_r) - X_s^\mu(\varphi_r)|^p \right)^{1/p},$$

$t, s \in [0, T]$, we estimate next the d_r^μ-diameter D_r of $[0, T]$. We therefore have

$$\left(d_r^\mu(t, s) \right)^p \leq c \max_{0 \leq t \leq T} E|X_t^\mu(\varphi_r)|^p$$

$$\leq c\mu^{-p} \max_{0 \leq t \leq T} E| \sum_{\alpha \sim t} \mathbf{1}\Big(\|Y^\alpha(t) + x_{\alpha_0}\| > r/2 \Big) \mathbf{1}\big(t < \tau(\alpha) \big)|^p$$

$$\leq c\mu^{-p} \max_{0 \leq t \leq T} E| \sum_{\alpha \sim t} \mathbf{1}\Big(\|Y^\alpha(t)\| > r/4 \Big) \mathbf{1}\big(t < \tau(\alpha) \big)|^p$$

$$+ c\mu^{-p} \max_{0 \leq t \leq T} E| \sum_{\alpha \sim t} \mathbf{1}\Big(\|x_{\alpha_0}\| > r/4 \Big) \mathbf{1}\big(t < \tau(\alpha) \big)|^p.$$

We now repeat the kind of arguments that led to (2.13) and (2.14) to see that for any $0 \leq t \leq T$

$$E| \sum_{\alpha \sim t} \mathbf{1}\Big(\|Y^\alpha(t)\| > r/4 \Big) \mathbf{1}\big(t < \tau(\alpha) \big)|^p$$

$$\leq c\Big[E| \sum_{\alpha \sim t} \big[\mathbf{1}\Big(\|Y^\alpha(t)\| > r/4 \Big) - P\Big(\|Y^\alpha(t)\| > r/4 \Big) \big] \mathbf{1}\big(t < \tau(\alpha) \big)|^p$$

$$+ P\Big(\|Y^\alpha(t)\| > r/4 \Big)^p E N(t)^p \Big]$$

$$\leq c\Big[\mu^{p/2}(EW^2)^{p/2} + \mu E|W|^p + P\Big(\|Y^\alpha(T)\| > r/4 \Big)^p E N(t)^p \Big],$$

where

$$W = \sum_{j=1}^{N_{[t\mu]}} \Big[\mathbf{1}\Big(\|Y^{\alpha_j}(t)\| > r/4 \Big) - P\Big(\|Y^{\alpha_j}(t)\| > r/4 \Big) \Big]. \tag{2.19}$$

Repeating the computation in (2.10) we easily conclude that for any $2 \leq a \leq 4$

$$E|W|^a \leq c\mu^{a-1} P\Big(\|Y^\alpha(t)\| > r/4 \Big)^{a/4} \leq c\mu^{a-1} P\Big(\|Y^\alpha(T)\| > r/4 \Big)^{a/4}.$$

Substituting the last bound into (2.19) and using (2.12) we conclude that

$$E| \sum_{\alpha \sim t} \mathbf{1}\Big(\|Y^\alpha(t)\| > r/4 \Big) \mathbf{1}\big(t < \tau(\alpha) \big)|^p \leq b\mu^p P\Big(\|Y^\alpha(T)\| > r/4 \Big)^{p/4}. \tag{2.20}$$

Furthermore, let $N_r(0)$ be the number of initial particles located at a distance of more than r from the origin. Since $N_r(0)$ has Poisson distribution with mean $\mu m(B_r^c)$, where B_r is the

closed ball of radius r centered at the origin, we obtain as in (2.12) that for any $0 \leq t \leq T$

$$E| \sum_{\alpha \sim t} 1\Big(\|x_{\alpha_0}\| > r/4 \Big) 1\big(t < \tau(\alpha) \big)|^p$$

$$\leq c E \Big(N_{r/4}(0)^p + \Big(\sum_{j=1}^{N_{r/4}(0)} E\big(N_{i_2} - 1\big)^2 \Big)^{p/2} + \sum_{j=1}^{N_{r/4}(0)} E|N_{i_2} - 1|^p \Big) \qquad (2.21)$$

$$\leq c\mu^p \big(m(B_{r/4}^c) \big)^{p/4}.$$

A consequence of (2.20) and (2.21) is that

$$\big(d_r^\mu(t,s) \big)^p \leq c\Big[P\Big(\|Y^\alpha(T)\| > r/4 \Big)^{p/4} + \big(m(B_{r/4}^c) \big)^{p/4} \Big] := D(r).$$

Observe that $D(r) \downarrow 0$ as $r \to \infty$. We can get now an upper bound on the size of $\sup_{0 \leq t \leq T} |X_t^\mu(\varphi_r)|$ by using (2.18) with Theorem 11.1 of Ledoux and Talagrand (1992) . We have

$$E \sup_{0 \leq t,s \leq T} |X_t^\mu(\varphi_r) - |X_s^\mu(\varphi_r)| \leq c \int_0^{D(r)} \Big(\epsilon^{-\max(2,1/H)} \Big)^{1/p} d\epsilon$$

$$= c D(r)^\theta$$

for some $\theta > 0$. Therefore,

$$E \sup_{0 \leq t \leq T} |X_t^\mu(\varphi_r)| \leq E|X_0^\mu(\varphi_r)| + E \sup_{0 \leq t,s \leq T} |X_t^\mu(\varphi_r) - |X_s^\mu(\varphi_r)|$$

$$\leq c(\mu E N_{r/2}(0) + D(r)^\theta) \qquad (2.22)$$

$$= c\Big(m(B_{r/2}^c) + D(r)^\theta \Big).$$

Note that the constants in the right hand side of (2.22) do not depend on μ.

Since $X^\mu(\varphi_r) \Rightarrow X(\varphi_r)$ weakly in $D(\Re)$ as $\mu \to \infty$, we conclude that

$$\sup_{0 \leq t \leq T} X_t^\mu(\varphi_r) \Rightarrow \sup_{0 \leq t \leq T} X_t(\varphi_r) \quad \text{in distribution}$$

as $\mu \to \infty$, and so by Fatou's lemma we get

$$E \sup_{0 \leq t \leq T} |X_t(\varphi_r)| \leq c\Big(m(B_{r/2}^c) + D(r)^\theta \Big).$$

By the definition of φ_r we immediately obtain

$$E \sup_{0 \leq t \leq T} X_t(\{\infty\}) \leq c\Big(m(B_{r/2}^c) + D(r)^\theta \Big),$$

and since the above holds for any r large enough, we let $r \to \infty$ to get

$$E \sup_{0 \leq t \leq T} X_t(\{\infty\}) = 0,$$

14

implying that there is an event Ω_+ of probability 1 such that for every $t \geq 0$ and $\omega \in \Omega_+$ we have $X_t(\{\infty\}) = 0$.

This completes the proof of Theorem 2.1 . ∎

3. THE FRACTIONAL BROWNIAN MOTION CASE

In this section we have two tasks. The first is to *carefully* define the process that we plan to use in obtaining a nice representation for the limit of branching fractional Brownian motions. The second is to establish the weak convergence result.

(a) Historical Brownian motion with tails. In this subsection we shall closely follow Perkin's (1992) construction of historical Brownian motion, and of certain stochastic integrals associated with it. Much of what we have to say could well be swept under the carpet with a totally justifiable "it is easy to see as in Perkin's paper", and we shall actually often do so. However, we nevertheless require enough notation to properly define our process, and to establish some technical results, in the following subsections, which cannot be "easily seen", and deserve checking.

Since the dimension d of our processes is fixed throughout, write, respectively, C and C^+ for the space of continuous functions from $(-\infty, \infty)$ and $[0, \infty)$ to $\Re^d$, endowed with topology of uniform convergence on compact intervals, and let $\mathcal{C}$ and $\mathcal{C}^+$ denoting the corresponding σ-algebras. Let $\mathcal{C}_t$ denote the canonical filtration on C (uncompleted and not made right continuous). If $y \in C$, $w \in C^+$, and $s \in \Re$, let

$$(y/s/\omega) = \begin{cases} y(t) & \text{if } t < s, \\ w(t-s) & \text{if } t \geq s. \end{cases} \tag{3.1}$$

Denote Wiener measure starting at x on $(C^+, \mathcal{C}^+)$ by P_x, and let $M_F(C)$ denote the space of finite measures on C. For $y \in C$, set

$$y^s(t) = \begin{cases} y(t) & \text{if } t < s, \\ y(s) & \text{if } t \geq s, \end{cases}$$

and for $s \geq 0$ let

$$M_F(C)^s = \{\nu \in M_F(C) : y = y^{(s)} \text{ for } \nu \text{ a.a. } y\}.$$

Choose $\nu \in M_F(C)^{(s)} - \{0\}$. Define $P_{s,\nu} \in M_F(C)$ by

$$P_{s,\nu}(A) = \int_C P_{y(s)}(w : (y/s/w) \in A)\,\nu(dy).$$

Let $\hat{\Re}^d = \{(s, y) \in \Re_+ \times C : y = y^{(s)}\}$, and for $(s, y) \in \hat{\Re}^d$ let $P_{s,y} = P_{s,\delta_y}$. Finally, define probability measures $\hat{P}_{s,y}$ and $\hat{P}_{s,\nu}$ on $\hat{C} = C([0, \infty], \hat{\Re}^d)$ by

$$\hat{P}_{s,y}(A) = P_{s,y}(\{w : (s + \cdot, (y/s/w)^{(s+\cdot)}) \in A\}),$$
$$\hat{P}_{s,\nu}(A) = \int \hat{P}_{s,y}(A)\,\nu(dy). \tag{3.2}$$

15

where $\nu \in M_F(C)^{(s)} - \{0\}$ and $A \in \hat{\mathcal{C}} =$ Borel sets in $\hat{C}$. If $\hat{B}_t(\omega) = \omega_t$ denotes the coordinate map on $\hat{C}$ and $\{\hat{\mathcal{C}}_{t+}\}$ is the canonical right-continuous filtration then $\hat{B} = (C, \hat{\mathcal{C}}, \hat{\mathcal{C}}_{t+}, \hat{B}_t, \hat{P}_{s,y})$ is a continuous strong Markov process with a Borel semigroup.

So far, we have done nothing other than to repeat Perkins' (1992) construction of the usual historical Brownian motion, with a redefinition of C as functions on the whole real line rather than just its positive half. Further details can therefore be found in Perkins' paper, and the references cited therein.

To see how all this relates to the constructions of the previous section, we require some notation. For any function y on $\Re$, define $\bar{y}$ by $\bar{y}(t) = y(-t)$. Similarly, if A is a Borel set in $C(-\infty, \infty)$, let $\bar{A}$ be the set defined by $\bar{A} = \{\bar{y} : y \in A\}$. Finally, if m is a probability measure on $C(-\infty, \infty)$, define $\bar{m}$ by $\bar{m}(A) = m(\bar{A})$. We leave it to the reader to check the following easy lemma.

Lemma 3.1 *In the construction above, take $\nu = \bar{P}_0$ at (3.2); i.e. Wiener measure in "negative time". If $\hat{B}_t(\omega) = (t, B_t(\omega))$, then, for $t \geq 0$, B_t is the full path, from $-\infty$ to t, of the "Brownian motion with tails" described in the previous section, and constructed explicitly at (1.15), but stopped at time t. We let $\hat{P}^B_{0,\bar{P}_0}$ denote the law of B_t on C.*

Returning to the general setting above, it now follows from the results of Fitzsimmons (1988) or Dawson and Perkins (1991) that B_t has an associated $M_F(\hat{\Re}^d)$ valued superprocess $\hat{H}$, itself a continuous, strong Markov process with a Borel semigroup. If $\hat{Q}_{0,\nu}$ denotes the law of $\hat{H}$ with $\hat{H}_0 = \delta_0 \times \nu$, then $\hat{H}_t = \delta_t \times H_t$, for all $t \geq 0$ a.s. We let Q^0_ν denote the law of H on $(\Omega_H, \mathcal{H})$, where $\Omega_H = C([0, \infty), M_F(C))$ and $\mathcal{H} = \bigvee_n \sigma(\omega_s : s \leq n)$. The process H is a time-inhomogeneous Borel strong Markov process with $M_F(C)$ valued paths. This is the historical process that we shall need to work with. In the setup of Lemma 3.1, it is the "historical Brownian motion with tails" of the previous section.

(b) Some properties of historical Brownian motion with tails. Let us start this section by agreeing, henceforth, to denote "Brownian motion with tails" by BMT, and its historical version by HBMT.

We now need to collect a number of properties of HBMT. The first is a straightforward application of Theorem 7.13 of Dawson and Perkins (1991), and describes the weak convergence of a particle system of BMT's to HBMT.

Retaining the notation of Section 1(a), let $\{\hat{B}^\alpha\}_\alpha$ be a system of branching Markov processes on $\hat{C}$, with transition probabilities as at (3.2). Since the initial values $\hat{B}^\alpha(0)$ are all of the form $(0, B^0)$, where B is a BMT, and the intensity m underlying the initial Poisson measure Π^μ is of the form $m = \delta_0 \times P^B_{0,\bar{P}_0}$.

Denote the associated, particle system, historical process by $\hat{H}^\mu$; i.e. for $A \in \hat{\mathcal{C}}$,

$$\hat{H}^\mu_t(A) := \mu^{-1} \sum_{\alpha \sim t} 1_A(\hat{B}^\alpha(t)). \tag{3.3}$$

Note that $\hat{H}^\mu_t = \delta_t \times H^\mu_t$, where

$$H^\mu_t(A) := \mu^{-1} \sum_{\alpha \sim t} 1_A(B^\alpha(t)) \tag{3.4}$$

16

and now $A \in \mathcal{C}$.

By Theorem 7.13 of Dawson and Perkins (1991) it follows immediately that $\hat{H}^\mu$ converges weakly in $D([0,\infty), M_F(\hat{\mathbb{R}}^d))$ to the sample continuous process $\hat{H}$ defined above. An immediate consequence of this result is

Theorem 3.2 *Under the above conditions, the sequence $\{H^\mu\}_{\mu \geq 1}$ converges, in $D([0,\infty), M_F(C))$ to the sample continuous process H defined above.*

Our next step is to show that, under the random measures H_t, the paths y behave like the branching BMT's used in our stochastic integral representation of fractional Brownian motion. This is crucial to all that follows.

We start with a slight extension (because of tails) of a somewhat simplified (because of non-random T in what follows) version of Perkins' (1992) Campbell measures:

Definition 3.3 *Let $0 < T < \infty$. The probability measures $\mathbf{P}_T$ on $(\Omega_H \times C, \mathcal{H} \times \mathcal{C})$ defined by*

$$\mathbf{P}_T(A \times B) := \frac{E\{1_A H_T(B)\}}{E\{H_T(1)\}} = \frac{E\{1_A H_T(B)\}}{M}, \tag{3.5}$$

where $M = E\{H_T(1)\} = E\{H_0(1)\}$, are called the Campbell measures for the process H.

The next theorem follows directly from the definition of Campbell measures and the properties of H as in the proof of Theorem 2.6 of Perkins (1992).

Theorem 3.4 *Denote a generic point of $\Omega_H \times C$ by (ω, y). Let y_t be the co-ordinate mapping on C. Then, for each $T > 0$, $B_t(y) := y_t$ is a BMT stopped at T on $(\Omega_H \times C, \mathcal{H} \times \mathcal{C}, \mathbf{P}_T)$.*

The importance of this result for us lies in the following theorem, whose proof is a simpler version (bar the extension to negative time) of the construction of the Itô integral in Section 3 of Perkins (1992).

Theorem 3.5 *Fix $0 < T < \infty$, and $0 < H < 1$. Then the integrals*

$$Y_t := \int_{-\infty}^{t} \left((t-r)^{H-1/2} - (-r)_+^{H-1/2} \right) y(dr), \tag{3.6}$$

can be defined, $\mathbf{P}_T$ a.s., in the same fashion as an Itô integral. Furthermore, under $\mathbf{P}_T$, $\{X_t\}_{-\infty < t \leq T}$ is a fractional Brownian motion of index H.

We require one more technical construction before we can turn to the main result of this section.

Definition 3.6 *Let $0 < T < \infty$. The probability measures $\mathbf{P}_T^{(2)}$ on $(\Omega_H \times C \times C, \mathcal{H} \times \mathcal{C} \times \mathcal{C})$ defined by*

$$\mathbf{P}_T^{(2)}(A \times B \times C) := \frac{E\{1_A H_T(B) H_T(C)\}}{E\{H_T^2(1)\}} = \frac{E\{1_A H_T(B) H_T(C)\}}{MT + M^2} \tag{3.7}$$

with M as above, are called the product Campbell measures for the process H.

Product measures of this form, for regular super Brownian motion, were studied in Tribe and Adler (1993). The same techniques used there suffice to establish

Theorem 3.7 *Let $M = E\{H_T(1)\} = E\{H_0(1)\}$. Denote a generic point of $\Omega_H \times C \times C$ by (ω, y_1, y_2). Let $y_i(t)$ represent co-ordinate mappings, and fix $0 < T < \infty$. Then, under $\mathbf{P}_T^{(2)}$, the pair $(B_1(t), B_2(t)) := (y_1(t), y_2(t))$ has the following distribution:*

$$(B_1(t), B_2(t)) \stackrel{\mathcal{L}}{=} \begin{cases} (W_0(t), W_0(t)) & \text{if } -\infty < t \leq \tau, \\ (W_0(\tau) + W_1((t \wedge T) - \tau), W_0(\tau) + W_2((t \wedge T) - \tau)) & \text{if } t > \tau, \end{cases}$$

where W_0 is a BMT, W_1 and W_2 are standard Brownian motions, and τ is a non-negative random variable taking the value 0 with probability $M/(T + M)$ and whose conditional distribution, given $\{\tau > 0\}$, is uniform on $[0, T]$. All four variables, W_0, W_1, W_2 and τ are independent of one another.

(c) Back to branching. With everything properly set up, we can now return to the setting of (1.16)–(1.18), where the branching fractional Brownian motions were set up as a function of a system of branching BMT's.

To set up our main result, let $\{B^\alpha\}_{\alpha \in I}$ be a branching system of BMT's as described above, and define a system of index-H fractional Brownian motions $\{Y^\alpha\}_{\alpha \in I}$ as in (1.16). As pointed out in Section 1(d), this system satisfies the conditional independence structure we require, with the filtrations generated by the B^α.

Furthermore, let Π^μ be a Poisson point process of intensity μm, where m is a finite measure on $\Re^d$, and let $x_1, x_2, \ldots$ be some ordering of the points of Π^μ. Note that the a.s. limit of $\mu^{-1}\Pi^\mu$ is simply m. Let $\{X^\alpha\}_{\alpha \in I}$ be the system of thinned processes based in the Y^α, as in (1.7). Finally, let X^μ be the particle superprocess (1.8), and let H^μ be the historical process (3.4) corresponding to the branching system of BMT's described above, with weak limit H.

Theorem 3.8 *Let X^μ be the particle superprocess described above. Assume that the progeny distribution has p-th moment, for some $p > 2$, and that the measure m has all its mass concentrated at the origin. If $H > 1/p$, then, as $\mu \to \infty$, the sequence $\{X^\mu\}_{\mu=1}^\infty$ converges weakly, in $D(M_F(\Re^d))$, to a $M_F(\Re^d)$-valued process X with a version with all sample paths in $C(M_F(\Re^d))$. Furthermore, X has the following representation, in law, as a function of the historical process H based on a BMT:*

$$\begin{aligned} X_t(\varphi) &= \int_{\Re^d} \varphi(x) \, X_t(dx) \\ &= \int_C \varphi\Big(\int_{-\infty}^t ((t - r)^{H-1/2} - (-r)_+^{H-1/2}) \, y(dr) \Big) \, H_t(dy), \end{aligned} \tag{3.8}$$

where $\varphi: \Re^d \to \Re$ ranges over all bounded, Lipshitz functions. The inner "stochastic integral" is that defined in Theorem 3.5.

Remarks: Note, firstly, that if $H = 1/2$ the fractional Brownian motions are regular Brownian motions, the limit process is super Brownina motion, (3.8) is essentially vacuous, and so there is nothing new to study.

Secondly, Theorem 3.9 does not cover the case when the initial measure is not concentrated at the origin. In this setting the historical process H has to be changed somewhat. In particular, rather than basing the historical process on the BMT process y, we need the process $\tilde{y}(t) := (z, y(t))$, where z is a random variable distributed according to m and independent of y. In that case the representation in (3.8) becomes

$$\int_{\Re^d \times C} \varphi\Big(z + \int_{-\infty}^t ((t-r)^{H-1/2} - (-r)_+^{H-1/2}) y(dr)\Big) H_t(dz, dy). \tag{3.9}$$

Note that for all Borel $A \in \Re^d$, $H_t(C, A) = m(A)$. We will comment on where the proof needs to be changed to accomodate this setting when appropriate.

Finally, given the structure of (3.8), and its version (3.10) below for the historical process H^μ of the finite system, it rather looks like one could prove Theorem 3.8 from the fact that $H^\mu \Rightarrow H^\mu$, along with a simple application of the continuous mapping theorem. However, it follows from the arguments appearing in Nualart and Zakai (1990) that the mapping from the Wiener space of BMT realizations to $\Re^d$ defined by $y \to \int_{-\infty}^t ((t-r)^{H-1/2} - (-r)_+^{H-1/2}) y(dr)$ can be extended to a continuous mapping on C if, and only if, $H \in (1/2, 1]$. Consequently, even if one worked to get this approach to generate a proof, it would work only for $H > 1/2$.

Proof: The first point to note is that the weak convergence and continuity of the limit follow immediately from Theorem 2.1. Thus, all we really need establish is the representation (3.8). For the moment, we fix $t > 0$ and bounded, Lipshitz φ.

To establish (3.8), we commence by noting that (1.18) gives the following version of (3.8) for the finite particle system.

$$\begin{aligned} X_t^\mu(\varphi) :&= \int_{\Re^d} \varphi(x) X_t^\mu(dx) \\ &= \int_C \varphi\Big(\int_{-\infty}^t ((t-r)^{H-1/2} - (-r)_+^{H-1/2}) y(dr)\Big) H_t^\mu(dy). \end{aligned} \tag{3.10}$$

(Note that if $X^\alpha(0) \not\equiv 0$, as assumed here, the last line will change to the spirit of (3.9).)

To make the notation somewhat lighter for remainder of the proof, write the inner integral in (3.10) as $F(y)$, so that (3.10) becomes

$$X_t^\mu(\varphi) = \int_C \varphi\big(F(y)\big) H_t^\vdash(dy). \tag{3.11}$$

We now claim that there exists an approximation, say F_ϵ, to F, so that, for each fixed $\epsilon > 0$, and each φ and $t > 0$,

$$\int_C \varphi(F_\epsilon(y)) H_t^\mu(dy) \Rightarrow \int_C \varphi(F_\epsilon(y)) H_t(dy) \qquad \text{as } \mu \to \infty. \tag{3.12}$$

Furthermore,

$$\int_C \varphi(F_\epsilon(y)) H_t(dy) \xrightarrow{L^2_?} \int_C \varphi(F(y)) H_t(dy) \qquad \text{as } \epsilon \to 0, \tag{3.13}$$

19

and, for each fixed $\mu > 0$,

$$\int_C \varphi(F_\epsilon(y))\, H_t^\mu(dy) \stackrel{L_3}{\rightarrow} \int_C \varphi(F(y))\, H_t^\mu(dy) \qquad \text{as } \epsilon \to 0. \tag{3.14}$$

The statement of the Theorem then follows via a standard argument. (e.g. Billingsley, 1968, Theorem 4.2.)

We start with (3.12), for which purpose we replace the innermost integrand in (3.10) by a simple function from $(-\infty, t] \to R$, of the form

$$f(s) = f_j, \qquad \text{if } s_{j-1} \le s < s_j, \; j = 1, \dots, k,$$

so that $F_\epsilon(y)$ becomes

$$F_\epsilon(y) \;=\; \sum_{j=1}^{k} f_j \cdot (y(s_j) - y(s_{j-1})).$$

Define a map $M : M_F(C) \to \Re$ by

$$M(K) = \int_C \varphi\Big(\sum_{j=1}^{k} f_j (y(s_j) - y(s_{j-1})) \Big)\, K(dy).$$

Note that this map is continuous. (Recall that C has the topology of uniform convergence on compact intervals). Indeed, the function $\Phi : C \to \Re$ defined by

$$\Phi(y) = \varphi(\sum_{j=1}^{k} f_j (y(s_j) - y(s_{j-1})))$$

is, obviously, continuous in our topology and bounded. The continuity of M therefore follows from the definition of the vague convergence. Theorem 3.2 and the continuous mapping theorem now suffice to establish (3.12).

For the proofs of (3.13) and (3.14), we require some notation and the following brief lemma.

Let

$$L_2^t \;=\; \{f : \int_{-\infty}^{t} f^2(s)\, ds < \infty\},$$

and set $\|f\|_{(t)}^2 = \int_{-\infty}^{t} f^2(s)\, ds$. Note, for later reference, that

$$\varphi_t(r) := (t - r)^{H-1/2} - (r)_+^{H-1/2} \in L_2^t, \tag{3.15}$$

for all $t < \infty$, if $H \in (0,1)$.

Lemma 3.9 *Let φ be bounded Lipshitz and $f, g \in L_2^t$. Then there exists a universal constant $C < \infty$, dependent only on t, such that for every $\mu > 0$*

$$E\Big| \int_C \big[\varphi\big(\int_{-\infty}^{t} f(s)\, y(ds)\big) - \varphi\big(\int_{-\infty}^{t} g(s)\, y(ds)\big)\big]\, H^\mu(dy) \Big|^2 \le C\|\varphi\|_{Lip}^2\, \|f - g\|_{(t)}^2, \tag{3.16}$$

where

$$\|\varphi\|_{Lip} := \sup_{x \neq y} \left\{ \frac{|\varphi(x) - \varphi(y)|}{\|x - y\|} \right\}$$

is the usual Lipshitz norm. Furthermore, (3.16) also holds when H^μ is replaced by the limiting historical process H.

Proof. Note that the left hand side of (3.16), when written as a sum rather than an integral over the particle historical process, is equivalent to

$$\mu^{-2} E \sum_{\alpha \sim t} \sum_{\beta \sim t} \left[\varphi\left(\int_{-\infty}^{t} f(s) B^\alpha(ds) \right) - \varphi\left(\int_{-\infty}^{t} g(s) B^\alpha(ds) \right) \right]$$

$$\times \left[\varphi\left(\int_{-\infty}^{t} f(s) B^\beta(ds) \right) - \varphi\left(\int_{-\infty}^{t} g(s) B^\beta(ds) \right) \right]$$

$$\leq \mu^{-2} \cdot E[\# \text{ of pairs } (\alpha, \beta) \text{ alive at time } t]^2 \cdot E\left[\varphi\left(\int_{-\infty}^{t} (f(s) - g(s)) B(ds) \right) \right]^2$$

$$\leq \mu^{-2} \cdot E[\# \text{ of particles } \alpha \text{ alive at time } t]^2 \cdot \|\varphi\|_{Lip}^2 \cdot E\left\| \int_{-\infty}^{t} [f(s) - g(s)] B(ds) \right\|^2$$

$$\leq C \|\varphi\|_{Lip}^2 \|f - g\|_{(t)}^2,$$

where the last line follows from (2.7).

This proves (3.16) for the case of finite μ. In the case when H^μ is replaced by the limiting historical process H, the left hand side of (3.16) is equivalent to

$$E \int_C \int_C \left[\varphi\left(\int_{-\infty}^{t} f(s) y(ds) \right) - \varphi\left(\int_{-\infty}^{t} g(s) y(ds) \right) \right]$$

$$\times \left[\varphi\left(\int_{-\infty}^{t} f(s) w(ds) \right) - \varphi\left(\int_{-\infty}^{t} g(s) w(ds) \right) \right] H_t(dy) H_t(dw)$$

$$\leq \|\varphi\|_{Lip}^2 \cdot E \int_C \int_C \left(\left[\int_{-\infty}^{t} [f(s) - g(s)] y(ds) \right] \cdot \left[\int_{-\infty}^{t} [f(s) - g(s)] w(ds) \right] \right) H_t(dy) H_t(dw)$$

$$= \|\varphi\|_{Lip}^2 \cdot (Mt + M^2) \cdot \mathbf{P}_t^{(2)} \left(\left[\int_{-\infty}^{t} [f(s) - g(s)] y(ds) \right] \cdot \left[\int_{-\infty}^{t} [f(s) - g(s)] w(ds) \right] \right),$$

where $\mathbf{P}_t^{(2)}$ is the product Campbell measure of Definition 3.6, and $M = EH_0(1) = m(\Re^d)$.

Now use the representation of the pair (y, w) given in Theorem 3.7, to evaluate the $\mathbf{P}_t^{(2)}$ expectation above. Incorporating the factor $(Mt + M^2)$ into the constant completes the proof of the lemma. ∎

We can now return to the proofs of (3.13) and (3.14). However, these are now quite easy. Recall that in order to establish (3.12) we defined F_t by approximating the function φ_t of (3.15) with a simple function on $\Re$. Call this approximation φ_t^ϵ. We now add the additional restraint that $\|\varphi_t - \varphi_t^\epsilon\|_{(t)}^2 \to 0$ as $\epsilon \to 0$. In view of (3.15), this is easy to do. Applying Lemma 3.9, (3.13) and (3.14) now follow immediately.

This completes the proof for fixed t and φ. To complete the proof in general, note that we can, obviously, repeat the above argument for a finite number of times and test functions,

and so show that the joint distribution of $(X_{t_1}^\mu(\varphi_1), \ldots, X_{t_n}^\mu(\varphi_n))$ converges, as $\mu \to \infty$, to the joint distribution of the right hand-side of (3.8) with the same $t_1, \ldots, t_n$ and $\varphi_1, \ldots, \varphi_n$. Therefore, $\langle X_{t_1}^\mu, \ldots, X_{t_n}^\mu \rangle$ converges weakly in $M_F(C^n)$ as $\mu \to \infty$, and, because of the tightness guaranteed by Theorem 2.1, Theorem 7.8 of Ethier and Kurtz (1986) establishes the full weak convergence of X^μ the right hand side of (3.8).

This completes the proof. ∎

4. Acknowledgement We are indebted to Professor E.B. Dynkin for a useful discussion on how to best apply his results to obtain the weak convergence of finite dimensional distributions in the proof of Theorem 2.1.

5. References

[1] Ash, R.B. (1972) *Real Analysis and Probability*, Academic, New York.

[2] Barlow and Perkins (1993), On the filtration of historical Brownian motion, Preprint.

[3] Billingsley (1968), *Convergence of Probability Measures*, Wiley, New York.

[4] Dawson, D. (1993), Measure-valued Markov processes, *Ecole d'Eté de Probabilités, XXI – 1991 Springer Lecture Notes in Math.*, **1541** 1-260.

[5] Dawson, D. and Perkins, E.A. (1991), Historical processes, *Amer Math Soc Memoirs*, **93** 1-179.

[6] Dynkin, E.B. (1991a), Path processes and historical superprocesses, *Prob. Theory and Related Fields*, **90** 1-36.

[7] Dynkin, E.B. (1991b), Branching particle systems and superprocesses, *Ann. Probab.* **19** 1157-1194.

[8] Ethier, S. and Kurtz, T. (1986), *Markov Processes*, Wiley, New York.

[9] Fitzsimmons, P. (1988) Construction and regularity of measure-valued branching processes, *Israel J. Math.* **64**, 337-361, Correction note, (1991), **73** 127.

[10] Ledoux, M. and Talagrand M. (1991) *Probability in Banach Spaces*, Springer, New York.

[11] Nualart, D. and Zakai, M. (1990) Multiple Wiener-Itô integrals possessing a continuous extension, *Probab. Theory and Related Fields* **85** 119-129.

[12] Perkins, E.A. (1992), Measure valued branching diffusions with spatial interactions, *Probab. Theory and Related Fields* **94** 189-245.

[13] Surgailis, D. (1981), On L^2 and non-L^2 multiple stochastic integration, in *Stochastic Differential Systems*, 212-226, *Lecture Notes in Control and Inform. Sci.* **36**, Springer, New York.

[14] Talmud Bavli, (*circa* 400), *Ethics of the Fathers*, III(a).

[15] Tribe, R. and Adler, R.J. (1993), Measure valued processes with singular interactions, In preparation.

[16] Vervaat, W. (1985), Sample paths of self-similar processes with stationary increments, *Ann. Probab.* **13**, 1-27.

[17] Walsh, J.B. (1986), An introduction to stochastic partial differential equations, *Springer Lecture Notes in Math.*, **1180**, 265-439.

Robert Adler
Faculty of Industrial Engineering and Management
Technion – Israel Institute of Technology
Haifa ISRAEL 32000
IERHE01@TECHNION.BITNET

Gennady Samoroditsky
School of Operations Research and Industrial Engineering
Cornell University
Ithaca NY 14583
GENNADY@ORIE.CORNELL.EDU