The Annals of Probability
2000, Vol. 28, No. 4, 1852–1869

LINEAR BOUNDS FOR STOCHASTIC DISPERSION

BY MIKE CRANSTON, MICHAEL SCHEUTZOW AND DAVID STEINSALTZ

University of Rochester, Technische Universität Berlin and Technische Universität Berlin

It has been suggested that stochastic flows might be used to model the spread of passive tracers in a turbulent fluid. We define a stochastic flow by the equations

$$\phi_0(x) = x,$$
$$d\phi_t(x) = F(dt, \phi_t(x)),$$

where $F(t, x)$ is a field of semimartingales on $x \in \mathbb{R}^d$ for $d \geq 2$ whose local characteristics are bounded and Lipschitz. The particles are points in a bounded set $\mathscr{X}$, and we ask how far the substance has spread in a time T. That is, we define

$$\Phi_T^* = \sup_{x\in\mathscr{X}} \sup_{0\leq t\leq T} \|\phi_t(x)\|,$$

and seek to bound $\mathrm{P}\{\Phi_T^* > z\}$.

Without drift, when $F(\cdot, x)$ are required to be martingales, although single points move on the order of $\sqrt{T}$, it is easy to construct examples in which the supremum Φ_T^* still grows linearly in time—that is, $\liminf_{T\to\infty} \Phi_T^*/T > 0$ almost surely. We show that this is an upper bound for the growth; that is, we compute a finite constant K_0, depending on the bounds for the local characteristics, such that

$$\limsup_{T\to\infty} \frac{\Phi_T^*}{T} \leq K_0 \text{ almost surely.}$$

A linear bound on growth holds even when the field itself includes a drift term.

1. Introduction. It has been proposed that stochastic flows might form the basis of an instructive model for the spread of "passive tracers" within or on the surface of a fluid. Individually the tracers are supposed to be diffusing, while the motions of adjacent particles are correlated, so that they form a coherent stochastic flow. On an infinitesimal scale this flow is driven by a field $F(t, x)$ of continuous semimartingales:

$$\phi_{st}(x) = x + \int_s^t F(du, \phi_{su}(x)).$$

We let $\phi_t(x) := \phi_{0t}(x)$.

Certain questions arise naturally when we imagine the flow to be carrying a mass of particles—an oil slick, for instance, or some other pollutant which we would like to confine. A particle begins at a point x within a bounded

Received May 1998; revised April 2000.
AMS 1991 *subject classification*. 60H20.
Key words and phrases. Stochastic flows, passive dispersion, chaining methods, empirical processes.

region $\mathscr{X}$, and is transported by time t to a point $\phi_t(x)$. R. Carmona [3] has noted that the boundary seems in general to expand exponentially in time, developing a convoluted fractal structure. C. Zirbel and E. Çinlar [15] have studied mean properties of the spreading substance (for the special case of isotropic Brownian flows), and in particular have shown that the center of mass advances subdiffusively.

In this paper we give a partial answer to one question which was posed, but not addressed, in the work of Zirbel and Çinlar: the time when the substance first leaves a ball of radius R. That is, we want to find bounds on

$$\mathrm{P}\left\{\sup_{0\le t\le T}\sup_{x\in\mathscr{X}}\|\phi_t(x)\| > R\right\}.$$

While the boundary grows exponentially long, this does not imply that any individual points move so far. But neither will the fastest point need to move at the same order of speed as a single diffusion. Even in the absence of drift, when $F(\cdot, x)$ are martingales, although generic points are displaced only as the square root of the time, it is not difficult to contrive a system in which the supremum over all points grows linearly with time. (One example is given in Section 3. In [5], we show that all isotropic Brownian flows with positive Lyapunov exponents have this property.) This linear growth was first noticed several years ago in simulations of a similar model by R. Carmona. He conjectured at the time, but did not prove, that the supremum (or diameter of the image) could grow no faster than linearly. (R. Carmona [personal communication]. The conjecture appears in Section 5.2 of his paper with F. Cerou [4], where it is variously attributed as well to Ya. Sinai, and to M. Isichenko [7].) A version of this conjecture is stated in Section 2, and is proved in Section 6.

2. Defining the problem. We start, as usual, with a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, \mathrm{P})$ satisfying the standard conditions, and $\mathscr{X}$, a bounded subset of $\mathbb{R}^d$, for $d \ge 2$. On this probability space is defined a field $F(t, x, \omega) = M(t, x, \omega) + V(t, x, \omega)$ of continuous $\mathbb{R}^d$-valued semimartingales which are 0 at time $t = 0$. Here $M(\cdot, x)$ is a martingale and $V(\cdot, x)$ is a continuous adapted process of locally bounded variation; t is a time in $[0, \infty)$ and x is the spatial point in $\mathbb{R}^d$. We impose the standard continuity condition

$$\langle M(\cdot, x, \omega), M(\cdot, y, \omega)\rangle_t = \int_0^t a(s, x, y, \omega)\,ds,$$
$$V(t, x, \omega) = \int_0^t b(s, x, \omega)\,ds,$$

where $a(s, x, y, \omega)$ and $b(s, x, \omega)$ (a $d\times d$ matrix and a d-vector, respectively) are continuous in (x, y) for almost every (s, ω), and predictable in (s, ω) for each (x, y). Defining

$$\mathscr{A}(s, x, y, \omega) = a(s, x, x, \omega) - a(s, y, x, \omega) - a(s, x, y, \omega) + a(s, y, y, \omega),$$

we also assume that there are some constants a, A, b, B such that these functions satisfy

$$\|\mathscr{A}(s, x, y, \omega)\| \le a^2\|x - y\|^2, \tag{1}$$

$$\|a(s, x, x, \omega)\| \le A^2, \tag{2}$$

$$\|b(s, x, \omega) - b(s, y, \omega)\| \le b\|x - y\|, \tag{3}$$

$$\|b(s, x, \omega)\| \le B, \tag{4}$$

for all $x, y \in \mathbb{R}^d$ and all $s \in [0, \infty)$ almost surely. We will often elide the ω. By another egregious abuse of notation, we allow $\|x\|$ to represent the Euclidean norm when x is a vector, while $\|\mathscr{A}\|$ represents the corresponding operator norm when $\mathscr{A}$ is a matrix.

We observe here that, for each $x, y \in \mathbb{R}^d$, the matrix $\mathscr{A}(t, x, y, \omega)$ is nonnegative definite for almost all $t \in [0, \infty)$ and ω, since for any vector v,

$$v^{\mathsf{T}}\mathscr{A}(t, x, y)v = \frac{d}{dt}\langle v \cdot (M(\cdot, x) - M(\cdot, y))\rangle_t.$$

Observe as well that, for any $d \times d$ real matrix $\mathscr{A}$,

$$\operatorname{Tr}\mathscr{A} \le (d-1)\|\mathscr{A}\| + \inf_{\|v\|=1} v^{\mathsf{T}}\mathscr{A}v. \tag{5}$$

To see this, simply compute the trace in an orthonormal basis which includes the vector v that minimizes $v^{\mathsf{T}}\mathscr{A}v$.

Kunita [8, Theorem 4.5.1] tells us that, under these conditions, the equation

$$\phi_{st}(x, \omega) = x + \int_s^t F(du, \phi_{su}(x))$$

determines a unique continuous stochastic flow $\phi_{st}(x, \omega)$ on $0 \le s, t < \infty$, of homeomorphisms in $x \in \mathbb{R}^d$. We will generally write $\phi_t(x)$ instead of $\phi_{0t}(x, \omega)$, and introduce in addition the notation

$$\Phi_T(x, y) = \sup_{0\le t\le T} \|\phi_t(x) - \phi_t(y)\|;$$

$$\Phi_T^* = \sup_{0\le t\le T} \sup_{x\in\mathscr{X}} \|\phi_t(x)\|.$$

We will also have use for a "driftless" version of the flow:

$$\psi_t(x) = \phi_t(x) - \int_0^t b(s, \phi_s(x))\, ds.$$

This is not a flow, of course, but it is a martingale for each x. We define $\Psi_T(x, y)$ and Ψ_T^* by analogy to Φ_T and Φ_T^*.

We ignore the case $d = 1$. The outer boundary of $\mathscr{X}$ consists then of only two points, which means that the supremum is simply the maximum of two diffusions. The growth is clearly on the order of $\sqrt{T}$ (without drift) or T (with drift), and may be bounded directly by the results of Section 5. The interesting situation arises when we try to control the maximum over infinitely many boundary points.

THEOREM 2.1. *The growth of Φ_T^* is at most linear in T, in the sense that*

$$\limsup_{T\to\infty} \frac{\Phi_T^*}{T} \le (d^2-d)aA\kappa(\tilde{b}) + B, \tag{6}$$

almost surely, where $\tilde{b} := b/a^2(d-1)$, and κ is a continuous increasing function given by (28), *with $\kappa(0) \approx 18.2$. Furthermore, for all positive γ and $\alpha' > 2$,*

$$\sup_{T\ge 0} \mathrm{E}[\rho_{\gamma,\alpha'}(\Phi_T^*/(T\vee 1))] < \infty, \tag{7}$$

where

$$\rho_{\gamma,\alpha'}(z) = \exp\{\gamma z^2/(1\vee \log z)^{\alpha'}\}.$$

There is a bound on this supremum which depends only on γ and α', on the dimension d, and on the local-characteristic bounds a, A, b and B.

3. Heuristic arguments and an example. We begin by first explaining why the expansion should be faster than merely diffusive—that is, order $\sqrt{T}$ growth; second, explaining why in the absence of drift the expansion should not be faster than linear; and third, giving an example to illustrate how the growth can in fact be linear.

To the first question there is a trivial answer: although each point moves diffusively (if the drift is zero), there are infinitely many points, and there is no reason a priori why the maximum should not be significantly larger than any individual element. But there is more to be said here, because the points are not free agents: adjacent points are strongly correlated. What drives the superdiffusive expansion is simply the rate at which the diffusing points lose their correlation. That is, not only do the points spread out individually, but as they spread they behave as though there were more of them.

The continuity conditions imply that the distance between two points following this motion will grow no faster than a geometric Brownian motion. That is,

$$\|\phi_t(x) - \phi_t(y)\| \sim \exp(\lambda t + \sigma W_t),$$

where W_t is a standard Brownian motion and λ is positive. (A precise statement is given in Lemma 5.1.) This short-range exponential expansion is seen in the aforementioned exponential growth of the boundary.

Suppose for simplicity that $d = 2$ and that $\mathscr{X}$ is the unit disk, and imagine that we knew that every piece of the boundary would stretch by a factor of no more than $e^{\lambda t}$ in time t. Then we could choose $n = \lceil 2\pi e^{\lambda t}\rceil$ points $(x_i)_{i=1}^n$ on the unit circle, at intervals of $e^{-\lambda t}$, and

$$\sup_{x\in B_1(0)} \|\phi_t(x)\| \le 1 + \max_{1\le i\le n} \|\phi_t(x_i)\|,$$

where $B_r(x)$ is the Euclidean ball of radius r around the point x. Each $\phi_t(x_i)$ is a diffusion, so $\|\phi_t(x_i)\|$ has an expectation about $\sigma\sqrt{t}$ and subgaussian tails, for some positive σ. But the maximum of n positive subgaussian variables with expectation bounded by $\sigma\sqrt{t}$ is itself subgaussian with expectation no more than $\sigma\sqrt{t}\sqrt{\log n}$. Thus we should expect $\sup_{x\in B_1(0)}\|\phi_t(x)\|$ to be subgaussian, with variance on the order of $\sigma\sqrt{\lambda}t$.

In fact, this is the essence of the chaining argument, by which we bound the rate of expansion—only, since it is not literally true that points never separate at a speed faster than $e^{\lambda t}$, no one finite set of points will suffice. Instead, we rely upon ever finer meshes of points to trap the supremum.

We now construct an example which does exhibit linear growth. (In our following paper [5], we prove that a wide class of somewhat more natural examples, the isotropic Brownian flows with positive Lyapunov exponents, also grows linearly.) It is a process in $\mathbb{R}^2$ driven by a single standard Brownian motion W_t. It exploits the fact that a tangential Brownian motion—a stochastic rotation—automatically acquires a deterministic radial component.

We take $\mathscr{X}$ to be the unit disk. Begin by choosing any $a > 0$ and $k \in (0, a^2/2)$. Let $\rho(x) = (x \wedge 1) \vee (-1)$, and define the stochastic differential equations

$$
\begin{aligned}
dX_t &= a\rho(Y_t)\,dW_t,\\
dY_t &= -a\rho(X_t - kt)\,dW_t.
\end{aligned}
$$

This may be understood as a random rotation, about a center which drifts to the right with constant rate k. Clearly, these equations satisfy the conditions (1) through (4), with $B = b = 0$ and $A = \sqrt{2}a$. Let ϕ be the stochastic flow associated with these equations. We will show that, for every $t \geq 0$, we have

$$\phi_t(B_1(0)) \supset B_1(kt)$$

almost surely, which implies that

$$t^{-1}\sup_{x\in B_1(0)}\|\phi_t(x)\| \geq k$$

almost surely.

Define

$$R_t = \sqrt{Y_t^2 + (X_t - kt)^2}.$$

Itô's formula implies that

$$
\begin{aligned}
dR_t^2 = {}& -2aY_t\rho(X_t - kt)\,dW_t + 2a(X_t - kt)\rho(Y_t)\,dW_t - 2(X_t - kt)k\,dt\\
&+ a^2\Big[\rho(X_t - kt)^2 + \rho(Y_t)^2\Big]\,dt.
\end{aligned}
$$

When $R_t \leq 1$, the right-hand side becomes

$$(-2(X_t - kt)k + a^2R_t^2)\,dt.$$

Since $a^2 - 2k > 0$, there is a positive ε such that $dR_t/dt > 0$ when $1 - \varepsilon \leq R_t \leq 1$. If $R_0 \geq 1$, it follows that

$$\mathrm{P}\{R_t \geq 1 \text{ for all } t \geq 0\} = 1.$$

By the continuity of $\phi_t(x)$ (with respect to x), this implies that

$$\mathrm{P}\{\phi_t(B_1(0)) \supset B_1(kt) \text{ for all } t \geq 0\} = 1.$$

Note that the lower bound for the speed here, $k = a^2/2$, is about a factor of 50 smaller than the upper bound $18.2\sqrt{2}a^2$ given by Theorem 1.

4. Basic results: chaining. The proof relies on two well-known, elementary tools—or, rather, on one well-known and one elementary tool. The well-known tool is Itô's formula, certainly a sophisticated result, but familiar to the point of banality. We use the comparison theorem, a direct consequence of Itô's formula, to derive effective bounds on the separation of two given points under the action of the stochastic flow. These will be presented in Section 5.

The elementary technique which turns these two-point approximations into a bound on the supremum is "chaining": a local hero within its home province of empirical-process theory but somewhat neglected in the wider world, ever since it was first introduced by Lévy to derive the modulus of continuity for Brownian motion. The idea is as simple as it is powerful: Given a totally bounded metric space $(\mathscr{X}, d)$ and a random continuous function ϕ: $\mathscr{X} \to \mathbb{R}^+$, we seek to bound $\sup_{x\in\mathscr{X}} \phi(x)$, on the basis of information only about the action of ϕ on single points and pairs of points.

This project relies upon a sequence of metric skeletons: given some positive real numbers δ_j such that $\sum_{j=0}^{\infty} \delta_j < \infty$, we find finite sets $\mathscr{X}_j$ in $\mathscr{X}$ such that, for every point $x \in \mathscr{X}$, there is a point $x_j \in \mathscr{X}_j$ with $d(x, x_j) \leq \delta_j$. For simplicity we also assume that $\mathscr{X}_0 = \{x_0\}$, with $d(x, x_0) \leq \delta_0$ for all $x \in \mathscr{X}$.

There are many variations on this theme: extensive accounts of the chaining method, theory and practice may be found in D. Pollard's expository paper [10] and in his lecture notes [11], as well as in the texts by M. Ledoux and M. Talagrand [9] and J. Wellner and A. van der Vaart [14]. There is no single most general form, and we will derive here a version which is optimized for our present applications. We make no claims for the originality of this lemma, but we have not found an exactly equivalent result in the literature.

Most renditions of the chaining concept rely upon Orlicz norms, but these are too crude in our context, because it requires that the tail bounds have the same form throughout the range of δ's. To obtain the best results, we will need to split up the δ's into three different ranges, so that we will need to work throughout directly with probability bounds. The increased power is paid for with a significant burden of extra bookkeeping, which we have attempted to hold to the sunny side of opaque. One prior work which takes a similar approach is K. Alexander [2].

LEMMA 4.1. *Let ϕ: $\mathscr{X} \to \mathbb{R}^+$ be an almost-surely continuous random function, with $(\mathscr{X}_j)$ and (δ_j) defined as above. For any positive z and ε and any* *sequence of positive* ε_j *with* $\varepsilon + \sum_{j=0}^{\infty} \varepsilon_j \le 1$,

$$
\begin{aligned}
\mathrm{P}\Bigl\{\sup_{x\in\mathscr{X}} \phi(x) > z\Bigr\} &\le \mathrm{P}\{\phi(x_0) > \varepsilon z\} \\
&\qquad + \sum_{j=0}^{\infty} |\mathscr{X}_{j+1}| \sup_{d(x,y)\le\delta_j} \mathrm{P}\{|\phi(x)-\phi(y)| > \varepsilon_j z\}.
\end{aligned}
\tag{8}
$$

PROOF. First, for each j and each $x \in \mathscr{X}_{j+1}$, define $g_j(x)$ to be a point in $\mathscr{X}_j$ such that $d(x, g_j(x)) \le \delta_j$. We claim that, for each $x \in \mathscr{X}$, there is an infinite sequence $x_0, x_1, \dots$ with $x_j \in \mathscr{X}_j$, such that $x = \lim_{j\to\infty} x_j$ and $x_j = g_j(x_{j+1})$.

Define $\delta_j^* = \sum_{i=j}^{\infty} \delta_i$, and let $\widetilde{\mathscr{X}_j}$ be the set of points in $\mathscr{X}_j$ within a distance δ_j^* of x. Also let $\widetilde{\mathscr{X}}$ be the disjoint union over j of all the $\widetilde{\mathscr{X}_j}$—that is, $\widetilde{\mathscr{X}} = \{(j,x): x \in \widetilde{\mathscr{X}_j}\}$. These sets are all nonempty, and if $\tilde{x} \in \widetilde{\mathscr{X}}_{j+1}$, then $g_j(\tilde{x}) \in \widetilde{\mathscr{X}_j}$. The tree whose vertices are $\widetilde{\mathscr{X}}$ and whose edges connect $\tilde{x} \in \widetilde{\mathscr{X}}_{j+1}$ to $g_j(\tilde{x})$ is infinite and connected, and every vertex has finite degree, so it must have at least one infinite path (by König's Infinity Lemma [13], Theorem 2.1). Since $\lim_{j\to\infty} \delta_j^* = 0$, this infinite path fulfills the claim. For every $x \in \mathscr{X}$, we choose one such sequence, and denote it by $(x_j(x))$.

By continuity, for almost every ϕ and every point x, we have

$$
\phi(x) = \phi(x_0) + \sum_{j=0}^{\infty} (\phi(x_{j+1}(x)) - \phi(x_j(x))).
$$

Thus,

$$
\begin{aligned}
\sup_{x\in\mathscr{X}} \phi(x) &\le \phi(x_0) + \sup_{x\in\mathscr{X}} \sum_{j=0}^{\infty} |\phi(x_{j+1}(x)) - \phi(x_j(x))| \\
&\le \phi(x_0) + \sum_{j=0}^{\infty} \max_{x_{j+1}\in\mathscr{X}_{j+1}} |\phi(x_{j+1}) - \phi(g_j(x_{j+1}))|.
\end{aligned}
$$

It follows then that

$$
\begin{aligned}
&\mathrm{P}\Bigl\{\sup_{x\in\mathscr{X}} \phi(x) > z\Bigr\} \\
&\quad \le \mathrm{P}\Biggl\{\phi(x_0) + \sum_{j=0}^{\infty} \max_{x\in\mathscr{X}_{j+1}} |\phi(g_j(x)) - \phi(x)| > z\Biggl(\varepsilon + \sum_{j=0}^{\infty} \varepsilon_j\Biggr)\Biggr\} \\
&\quad \le \mathrm{P}\Biggl(\{\phi(x_0) > \varepsilon z\} \cup \bigcup_{j=0}^{\infty} \bigcup_{x\in\mathscr{X}_{j+1}} \{|\phi(g_j(x)) - \phi(x)| > \varepsilon_j z\}\Biggr) \\
&\quad \le \mathrm{P}\{\phi(x_0) > \varepsilon z\} + \sum_{j=0}^{\infty} \sum_{x\in\mathscr{X}_{j+1}} \mathrm{P}\{|\phi(x) - \phi(g_j(x))| > \varepsilon_j z\} \\
&\quad \le \mathrm{P}\{\phi(x_0) > \varepsilon z\} + \sum_{j=0}^{\infty} |\mathscr{X}_{j+1}| \sup_{d(x,y)\le\delta_j} \mathrm{P}\{|\phi(x) - \phi(y)| > \varepsilon_j z\}. \qquad \square
\end{aligned}
$$

If the skeletons are not too big (this particular topic of concern is referred to as "metric entropy" in the chaining business), and if the probabilities in this last expression fall sufficiently quickly as $d(x, y) \to 0$, then we may hope to obtain a reasonably well-behaved sum.

5. Basic results: the two-point motion. The first part of the solution is an estimate on the rate of separation of any two starting points x and y, under the action of the stochastic flow. The difference $\phi_t(x) - \phi_t(y)$ is a semimartingale. Its local diffusion and drift rate are bounded according to the conditions (1) through (4) on the one hand by $4A^2$ and $2B$, on the other hand by $a^2\|\phi_t(x)-\phi_t(y)\|^2$ and $b\|\phi_t(x)-\phi_t(y)\|$. The former provides an irreproachable global bound on the distance between the two points, to wit $2A|W_t|+2Bt$, where W_t is a Brownian motion. Such a bound is adequate when x and y are far apart, but is much too coarse when they are adjacent, when the Lipschitz properties of the flow are essential to chain together the motions. There we bound the spread instead by a multiplicative Brownian motion—much worse in the long run, but with the advantage that the separation remains a multiple of $\|x - y\|$.

The goal is to merge these two bounds. When the initial interval $\|x - y\|$ is manageably small, we run with (1) and (3), until the exponential growth rate destroys the initial correlation. At that point we switch to the global diffusion bounds (2) and (4).

LEMMA 5.1. *If the conditions* (1) *through* (4) *are satisfied, we may define, for each pair of points x, y in $\mathscr{X}$, a standard Brownian motion W_s such that*

$$\Phi_T(x, y) \le \sup_{0\le s\le T} \Big\{\|x - y\| \exp\{aW_s + \lambda T\}\Big\}$$

for all $T \ge 0$, where $\lambda = b + (d-1)a^2/2$.

PROOF. If $x = y$, the statement is trivial; so assume $x \ne y$. Define the stochastic process $\Delta_t = \|\phi_t(x) - \phi_t(y)\|$. By Itô's formula, this satisfies the differential equation

$$\begin{aligned} d\Delta_t = {} & \frac{\phi_t(x) - \phi_t(y)}{\Delta_t} \cdot (M(dt, \phi_t(x)) - M(dt, \phi_t(y))) \\ & + \frac{\phi_t(x) - \phi_t(y)}{\Delta_t} \cdot (V(dt, \phi_t(x)) - V(dt, \phi_t(y))) \\ & + \frac{1}{2} \sum_{i,\,j=1}^{d} \Delta_t^{-1}\Big(\delta_{i,j} - \Delta_t^{-2}(\phi_t^i(x) - \phi_t^i(y))(\phi_t^j(x) - \phi_t^j(y))\Big) \\ & \qquad \times \mathscr{A}_{ij}(t, \phi_t(x), \phi_t(y))\, dt. \end{aligned} \tag{9}$$

Writing $\mathscr{A}$ for $\mathscr{A}(t, \phi_t(x), \phi_t(y))$, the drift component is bounded by

$$b\Delta_t + \frac{\operatorname{Tr} \mathscr{A}}{2\Delta_t} - \frac{1}{2\Delta_t} \inf_{\|v\|=1} v^{\mathsf{T}} \mathscr{A} v \le \lambda \Delta_t. \tag{10}$$

Here we have used the bound (5), together with the assumption that $\|\mathscr{A}\| \le a^2\Delta_t^2$.

Let

$$N_t = \int_0^t \frac{\phi_s(x) - \phi_s(y)}{\Delta_s^2} \cdot (M(ds, \phi_s(x)) - M(ds, \phi_s(y))).$$

$N_t(t \ge 0)$ is a continuous local martingale, with

$$\frac{d\langle N\rangle_t}{dt} = \Delta_t^{-4}(\phi_t(x) - \phi_t(y))^{\mathsf{T}}\mathscr{A}(t, \phi_t(x), \phi_t(y))(\phi_t(x) - \phi_t(y)) \le a^2; \tag{11}$$

and let $\tilde{\Delta}_t(t \ge 0)$ be the solution of the stochastic differential equation

$$\begin{aligned} d\tilde{\Delta}_t &= \tilde{\Delta}_t\, dN_t + \lambda\tilde{\Delta}_t\, dt, \\ \tilde{\Delta}_0 &= \|x - y\|. \end{aligned} \tag{12}$$

The diffusive components and the initial conditions of the two equation systems are identical, while the drift component of (9) is smaller than that of (12). By the Comparison Theorem [6, Theorem VI.1.1], this implies that $\tilde{\Delta}_t \ge \Delta_t$ almost surely for all $t \ge 0$. (N.B.: Ikeda and Watanabe state their theorem for N_t a Brownian motion, but in fact their proof does not call on any special properties, and is just as valid if N is a general continuous local martingale.)

By an application of Itô's formula, the solution of (12) is simply

$$\tilde{\Delta}_t = \|x - y\| \exp\left\{N_t - \frac{1}{2}\langle N\rangle_t + \lambda t\right\}.$$

Since N_t is a continuous local martingale, with $N_0 = 0$ almost surely and quadratic variation bounded by (11), we may represent N (possibly on an enriched probability space) in the form $N_t = aW_{\tau(t)}$, where W_t is a standard Brownian motion, and $\tau(t)$ is a family of stopping times with $\tau(t) \le t$ almost surely (cf. [12, Theorem V.1.4]). In particular, on this probability space,

$$\Phi_T(x, y) \le \sup_{0\le s\le T} \tilde{\Delta}_s \le \|x - y\| \sup_{0\le s\le T} \exp\{aW_s + \lambda T\}. \qquad \square$$

We may now use simple facts about Brownian motion to derive useful bounds on the two-point separation. We assume that A and a are nonzero.

PROPOSITION 5.2. *For all x, y in $\mathscr{X}$, all positive times T, and any positive z, the two-point separation $\Phi_T(x, y)$ satisfies*

$$\mathrm{P}\{\Phi_T(x, y) > z\} \le \frac{2}{\sqrt{\pi}} \exp\left\{-\frac{1}{2a^2T}\left((\log z - \log\|x - y\| - \lambda T)^+\right)^2\right\}, \tag{13}$$

$$\mathrm{P}\{\Phi_T(x, y) > z\} \le \frac{4d}{\sqrt{\pi}} \exp\left\{-\frac{1}{8A^2T}\left(\left(\frac{z}{d} - \|x - y\| - 2BT\right)^+\right)^2\right\}, \tag{14}$$

where $\lambda = b + (d-1)a^2/2$. We also have the one-point bound

$$\mathrm{P}\Big\{\sup_{0\le t\le T}\|\phi_t(x)\| > z\Big\} \le \frac{4d}{\sqrt{\pi}}\exp\bigg\{-\frac{1}{2A^2T}\bigg(\Big(\frac{z}{d} - \|x\| - BT\Big)^+\bigg)^2\bigg\}. \tag{15}$$

PROOF. By Lemma 5.1, if $a > 0$,

$$\mathrm{P}\{\Phi_T(x, y) > z\} \le \mathrm{P}\Big\{\sup_{0\le t\le T} W_t \ge \frac{1}{a}(\log z - \log\|x-y\| - \lambda T)\Big\}.$$

Applying the reflection principle and the approximation

$$\int_z^\infty \exp(-x^2/2)\,dx \le \sqrt{2}\exp(-z^2/2)$$

for z positive (see, e.g., [1], 7.1.13) proves (13).

The bound (14) is obtained by ignoring the continuity of the stochastic flux, and pretending that the points move diffusively, with diffusion constant no more than A and drift no more than B. Formally, for each coordinate $1 \le i \le d$,

$$\begin{aligned}\big|\phi_t^i(x) - \phi_t^i(y)\big| &= \bigg|x^i - y^i + \int_0^t \big[b^i(s, \phi_s(x)) - b^i(s, \phi_s(y))\big]\,ds \\ &\qquad + \int_0^t M^i(ds, \phi_s(x)) - \int_0^t M^i(ds, \phi_s(y))\bigg| \\ &\le |x^i - y^i| + 2Bt + |\widetilde{M}_t^i|,\end{aligned}$$

where $\widetilde{M}_s^i$ is a local martingale with quadratic variation

$$\frac{d\langle\widetilde{M}^i\rangle_t}{dt} = \mathscr{A}_{ii}(t, \phi_t(x), \phi_t(y)) \le 4A^2.$$

By the same argument as in Lemma 5.1, this implies that we may define d standard Brownian motions (not necessarily independent) with

$$\sup_{0\le t\le T}|\phi_t^i(x) - \phi_t^i(y)| \le |x^i - y^i| + 2BT + 2A\sup_{0\le t\le T}|W_t^i|.$$

Thus,

$$\begin{aligned}\mathrm{P}\{\Phi_T(x, y) > z\} &\le d\max_{1\le i\le d}\mathrm{P}\Big\{\sup_{0\le t\le T}|W_t^i| \ge \frac{1}{2A}\Big(\frac{z}{d} - |x^i - y^i| - 2BT\Big)\Big\} \\ &\le \frac{4d}{\sqrt{\pi}}\exp\bigg\{-\frac{1}{8A^2T}\bigg(\Big(\frac{z}{d} - \|x-y\| - 2BT\Big)^+\bigg)^2\bigg\}.\end{aligned}$$

The one-point bound (15) is proved identically. □

The corresponding bounds for the driftless version follow directly.

COROLLARY 5.3. *For all x, y in $\mathscr{X}$, all positive times T, and any positive z, the two-point separation $\Psi_T(x, y)$ satisfies*

$$\mathrm{P}\{\Psi_T(x,y)>z\}\le\frac{2}{\sqrt{\pi}}\exp\biggl\{-\frac{1}{2a^2T}\biggl(\biggl(\log\frac{z}{bT+1}-\log\|x-y\|-\lambda T\biggr)^+\biggr)^2\biggr\} \tag{16}$$

and

$$\mathrm{P}\{\Psi_T(x, y) > z\} \le \frac{4d}{\sqrt{\pi}}\exp\biggl\{-\frac{1}{8A^2T}\biggl(\biggl(\frac{z}{d} - \|x-y\|\biggr)^+\biggr)^2\biggr\}, \tag{17}$$

where $\lambda = b + (d-1)a^2/2$. We also have the one-point bound

$$\mathrm{P}\Bigl\{\sup_{0\le t\le T}\|\psi_t(x)\| > z\Bigr\} \le \frac{4d}{\sqrt{\pi}}\exp\biggl\{-\frac{1}{2A^2T}\biggl(\biggl(\frac{z}{d} - \|x\|\biggr)^+\biggr)^2\biggr\}. \tag{18}$$

PROOF. We note first that by (3),

$$\begin{aligned}
\Psi_T(x, y) &= \sup_{0\le t\le T}\|\psi_t(x)-\psi_t(y)\| \\
&= \sup_{0\le t\le T}\biggl\|\phi_t(x)-\phi_t(y)-\int_0^t (b(s,\phi_s(x))-b(s,\phi_s(y)))\,ds\biggr\| \\
&\le \sup_{0\le t\le T}\|\phi_t(x)-\phi_t(y)\| + \int_0^T \|b(s,\phi_s(x))-b(s,\phi_s(y))\|\,ds \\
&\le \Phi_T(x, y) + \int_0^T b\|\phi_s(x)-\phi_s(y)\|\,ds \\
&\le (1+bT)\Phi_T(x, y).
\end{aligned}$$

The bound (16) is thus an immediate consequence of (13).

The bounds (17) and (18) follow exactly the same scheme as their analogues in Proposition 5.2, since

$$\begin{aligned}
|\psi_t^i(x)-\psi_t^i(y)| &= \biggl|x^i - y^i + \int_0^t M^i(ds,\phi_s(x)) - \int_0^t M^i(ds,\phi_s(y))\biggr| \\
&\le |x^i-y^i| + |\widetilde{M}_t^i|. \qquad \square
\end{aligned}$$

6. Proof of the theorem.

Observe first that

$$\Phi_T^* = \sup_{0\le t\le T}\sup_{x\in\mathscr{X}}\biggl\|\psi_t(x)+\int_0^t b(s,\psi_s(x))\,ds\biggr\| \le \Psi_T^* + BT.$$

Thus, it will suffice to prove that

$$\limsup\frac{\Psi_T^*}{T} \le d(d-1)aA\kappa(\tilde{b}),$$

and that the exponential moment bounds (7) hold for Ψ_T^* in place of Φ_T^*.

By [8], Theorem 4.5.1, there is a modification of the system (ϕ_t) in which the maps are homeomorphisms depending continuously on t. It is a simple consequence of the bounds (3) and (4) on $b(s, x)$ that, for any $t, u \in \mathbb{R}^+$ and $x, y \in \mathscr{X}$,

$$\left\| \int_0^t b(s, \phi_s(x))\, ds - \int_0^u b(s, \phi_s(y))\, ds \right\| \le bt\, \Phi_t(x, y) + |t-u|B$$

almost surely, so the functions $\psi_t(x)$ are also almost-surely continuous in t and x, allowing us to apply chaining bounds.

Without loss of generality, we may assume that $\mathscr{X}$ is the unit ball in $\mathbb{R}^d$. The maximum must occur on the boundary, so we may ignore the interior of the ball: $\Psi_T^* = \sup_{t \le T} \sup_{\|x\|=1} \|\psi_t(x)\|$. For simplicity we throw in the origin, and take $\mathscr{X} = \{0\} \cup \{x \in \mathbb{R}^d : \|x\| = 1\}$.

Choose $\varepsilon \in (0, \frac{1}{4})$, and $\alpha \in (2, 2e)$. Let k, T and θ be positive, so that

$$kT > 16 \frac{d^2}{\varepsilon^2}, \tag{19}$$

$$k > \frac{1}{\theta e^p}\left(a^2\left(e + \frac{1}{2}\right)(d-1) + b + \frac{a^2}{\theta e^p kT} + \frac{5 + \log(bT+1)}{T}\right), \tag{20}$$

where $p = \lceil \log kT \rceil - \log kT \in [0, 1)$. Define $\delta_0 = 1$, and $\delta_j = \exp(-\theta e^j)$ for $j \ge 1$. We may find a sequence of δ_j-skeletons $(\mathscr{X}_j)$ for $\mathscr{X}$, taking $\mathscr{X}_0 = \{0\}$, such that the number of points in $\mathscr{X}_j$ is $|\mathscr{X}_j| \le \gamma_d \delta_j^{-d+1}$, where a crude calculation shows that $\gamma_d \le 2^d d^{(d+1)/2}$. We also define

$$\begin{aligned} j_1 &= \lfloor \log kT - \alpha \log\log kT \rfloor, \\ j_2 &= \lceil \log\, kT \rceil, \\ \beta &= (\sqrt{e} - 1)(1 - 2\varepsilon). \end{aligned}$$

We use these to define a sequence (ε_j) with $\sum \varepsilon_j \le 1$:

$$\varepsilon_j = \begin{cases} \varepsilon/(j_1+1), & \text{if } 0 \le j \le j_1, \\ \beta \exp((j - j_2)/2), & \text{if } j_1 + 1 \le j \le j_2 - 1, \\ (1-2\varepsilon)e^{-j-1}, & \text{if } j \ge j_2. \end{cases}$$

The sum of the ε_j is

$$\begin{aligned} \sum_{j=0}^{\infty} \varepsilon_j &= (j_1+1)\frac{\varepsilon}{j_1+1} + (\sqrt{e}-1)(1-2\varepsilon)\exp(-j_2/2) \sum_{j=j_1+1}^{j_2-1} e^{j/2} \\ &\quad + (1-2\varepsilon) \sum_{j=j_2}^{\infty} e^{-j-1} \\ &= \varepsilon + (1-2\varepsilon)\exp(-j_2/2)(\exp(j_2/2) - \exp(j_1+1)/2) \\ &\quad + \frac{\exp(-j_2)(1-2\varepsilon)}{e-1} \end{aligned}$$

$$
\begin{aligned}
&\le \varepsilon + (1-2\varepsilon)\left(1 - \frac{1}{\sqrt{e}(\log kT)^{\alpha/2}}\right) + \frac{1-2\varepsilon}{(e-1)kT}\\
&\le 1-\varepsilon.
\end{aligned}
$$

(Note that $2 < \alpha < 2e$ implies that $(\log x)^{\alpha/2} < x$ for all $x > 1$.)

Since the supremum of a sum is smaller than the sum of the suprema, we may apply the chaining inequality (8) to the function $\Psi(x) := \sup_{0\le t\le T} \|\psi_t(x)\|$:

$$
\begin{aligned}
\mathrm{P}\{\Psi_T^* > kT\} \le{}& \mathrm{P}\left\{\sup_{0\le t\le T} \|\psi_t(0)\| > \varepsilon kT\right\}\\
&+ \sum_{j=0}^{\infty} |\mathscr{X}_{j+1}| \sup_{\|x-y\|\le\delta_j} \mathrm{P}\{\Psi_T(x,y) > \varepsilon_j kT\}.
\end{aligned}
\tag{21}
$$

By inequality (18), the first term is bounded by

$$
\mathrm{P}\left\{\sup_{0\le t\le T} \|\psi_t(0)\| > \varepsilon kT\right\} \le \frac{4d}{\sqrt{\pi}} \exp\left\{-\frac{(\varepsilon kT)^2}{2A^2d^2T}\right\}.
\tag{22}
$$

We consider now the sum. Using Corollary 5.3, this may be bounded by

$$
\begin{aligned}
&\frac{4d\gamma_d}{\sqrt{\pi}} \sum_{j=0}^{j_1} \exp\left\{(d-1)\theta e^{j+1} - \frac{1}{8d^2A^2T}\left(\left(\frac{\varepsilon kT}{j_1+1} - d\right)^+\right)^2\right\}\\
&\quad + \frac{4d\gamma_d}{\sqrt{\pi}} \sum_{j=j_1+1}^{j_2-1} \exp\left\{(d-1)\theta e^{j+1} - \frac{1}{8d^2A^2T}\left(\left(\beta\sqrt{kT}e^{(j-p)/2} - d\right)^+\right)^2\right\}\\
&\quad + \frac{2\gamma_d}{\sqrt{\pi}} \sum_{j=j_2}^{\infty} \exp\left\{(d-1)\theta e^{j+1} - \frac{1}{2a^2T}\right.\\
&\qquad\qquad \left.\times\left(\left(-j + \log\frac{kT}{bT+1} + \theta e^j - \lambda T - 3\right)^+\right)^2\right\}.
\end{aligned}
\tag{23}
$$

In the first sum the terms are increasing, so it is bounded by

$$
\begin{aligned}
&(j_1+1)\exp\left\{(d-1)\theta \exp(j_1+1) - \frac{1}{8d^2A^2T}\left(\left(\frac{\varepsilon kT}{j_1+1} - d\right)^+\right)^2\right\}\\
&\quad \le \log(kT)\exp\left\{\frac{e(d-1)\theta kT}{(\log kT)^\alpha} - \frac{1}{8d^2A^2T}\left(\left(\frac{\varepsilon kT}{\log kT} - d\right)^+\right)^2\right\}\\
&\quad \le \log(kT)\exp\left\{\frac{e(d-1)\theta kT}{(\log kT)^\alpha} - \frac{\varepsilon^2k^2T - 2d\varepsilon k\log kT}{8d^2A^2(\log kT)^2}\right\}.
\end{aligned}
$$

Since $\varepsilon^2 kT > 4d\varepsilon\sqrt{kT} > 4d\varepsilon \log kT$ (by (19)), this bound may be simplified to

$$
\log(kT)\exp\left\{\frac{kT}{(\log kT)^2}\left(e(d-1)\theta(\log kT)^{-\alpha+2} - \frac{\varepsilon^2 k}{16d^2A^2}\right)\right\}.
\tag{24}
$$

For the middle sum we have

$$
\begin{aligned}
&\sum_{j=j_1+1}^{j_2-1} \exp\left\{\theta e^{j+1}(d-1) - \frac{1}{8d^2A^2T}\left(\left(\beta\sqrt{kT}e^{(j-p)/2} - d\right)^+\right)^2\right\} \\
&\qquad \le \sum_{j=j_1+1}^{j_2-1} \exp\left\{\theta e^{j+1}(d-1) - \frac{e^j(\beta^2 e^{-p}kT - 2\beta\sqrt{kT}e^{-j/2}d)}{8d^2A^2T}\right\} \\
&\qquad \le (1+\alpha\log\log kT) \\
&\qquad\quad \times \exp\left\{\frac{kT[8ed^2(d-1)\theta A^2 - \beta^2 e^{-p}k + 2d\beta(\log kT)^{\alpha/2}/T]}{8d^2A^2(\log kT)^\alpha}\right\},
\end{aligned}
\tag{25}
$$

as long as the quantity in the braces is negative. (If positive, it is in any case tautologous as a probability bound.)

In the last (infinite) sum, the ratio of successive terms with indices j and $j+1$ is

$$
\begin{aligned}
\exp\Bigg\{&\theta e^{j+1}(d-1) - \frac{1}{2a^2T}\left(\left(-j-3+\log\frac{kT}{bT+1} + \theta e^j - \lambda T\right)^+\right)^2 \\
&- \theta e^{j+2}(d-1) + \frac{1}{2a^2T}\left(\left(-j-4+\log\frac{kT}{bT+1} + \theta e^{j+1} - \lambda T\right)^+\right)^2\Bigg\}.
\end{aligned}
$$

We know from (20) that $\theta e^j(e-1)-1$ is nonnegative, so we may apply the elementary inequality

$$((y+z)^+)^2 - (y^+)^2 \ge 2yz \quad \text{for } z \ge 0$$

with $z = \theta e^j(e-1)-1$, to conclude that the ratio is bounded below by

$$
\begin{aligned}
&\exp\Bigg\{-\theta e^{j+1}(e-1)(d-1) \\
&\qquad + \frac{(\theta e^j(e-1)-1)(-j-3+\log[kT/(bT+1)]+\theta e^j - \lambda T)}{a^2T}\Bigg\} \\
&\ge \exp\Bigg\{-\theta e^{j+1}(e-1)(d-1) \\
&\qquad + \frac{(\theta e^j(e-1)-1)(-\log kT - 4 + \log[kT/(bT+1)] + \theta e^p kT - \lambda T)}{a^2T}\Bigg\} \\
&\ge \exp\left\{\theta e^j(e-1)\left[-e(d-1) + \frac{\theta e^p kT - 4 - \lambda T - \log(bT+1)}{a^2T}\right] - \frac{\theta e^p kT}{a^2T}\right\} \\
&\ge \exp\left\{\theta e^p kT(e-1)\left(-e(d-1) - \frac{4+(e-1)^{-1}+\log(bT+1)}{a^2T} + \frac{\theta e^p k - \lambda}{a^2}\right)\right\}.
\end{aligned}
$$

The second step uses (20), which implies that the map $j \mapsto -j + \theta e^j$ is nondecreasing on $[j_2, \infty)$. The last step uses the fact that the quantity in square brackets is positive, again by (20). Remember that $\lambda = b + (d-1)a^2/2$, so that a final application of (20) shows the ratio in question to be at least 2. It follows that the infinite sum is bounded by

$$
\begin{aligned}
&2\,\exp\biggl\{\theta\exp(j_2+1)(d-1)\\
&\qquad\qquad - \frac{1}{2a^2T}\biggl(-j_2-3+\log\frac{kT}{bT+1}+\exp(j_2)\theta-\lambda T\biggr)^2\biggr\}\\
&\quad = 2\,\exp\biggl\{e^{1+p}(d-1)\theta kT\\
&\qquad\qquad - \frac{1}{2a^2T}((e^p\theta k-\lambda)T-3-p-\log(bT+1))^2\biggr\}\\
&\quad \le 2\,\exp\biggl\{T\biggl(e^{1+p}(d-1)\theta k-\frac{(e^p\theta k-\lambda)^2}{2a^2}\biggr)\\
&\qquad\qquad + \frac{(4+\log(bT+1))(e^p\theta k-\lambda)}{a^2}\biggr\}.
\end{aligned}
\tag{26}
$$

Putting all of this together, we see that the sum in the chaining inequality (21) is bounded by

$$
\begin{aligned}
&\frac{4d\gamma_d}{\sqrt{\pi}}\biggl(\log(kT)\exp\biggl\{\frac{kT}{(\log kT)^2}\biggl(e(d-1)\theta(\log kT)^{2-\alpha}-\frac{\varepsilon^2k}{16d^2A^2}\biggr)\biggr\}\\
&\qquad + (1+\alpha\log\log kT)\\
&\qquad \times\exp\biggl\{\frac{kT(8e\theta d^2(d-1)A^2-\beta^2e^{-p}k+2d\beta(\log kT)^{\alpha/2}/T)}{8d^2A^2(\log kT)^\alpha}\biggr\}\\
&\qquad + \exp\biggl\{T\biggl(e^{1+p}k\theta(d-1)-\frac{(e^pk\theta-\lambda)^2}{2a^2}\biggr)\\
&\qquad\qquad + \frac{(4+\log(bT+1))(e^pk\theta-\lambda)}{a^2}\biggr\}\biggr).
\end{aligned}
$$

As T gets large, the first terms will go to 0 nearly exponentially fast, no matter what k is (provided, of course, that the conditions (19) and (20) are satisfied). The other two terms, on the other hand, will go to 0 only if k is chosen to make the coefficients of T in the exponents negative. There will be a critical value of k above which this will be true, and since this will be our bound on the speed, we want to make it as small as possible. This is accomplished when the same k is the critical value for both, and we may manipulate θ to make this happen. That is, we fix θ to make

$$
\frac{8e\theta d^2(d-1)A^2}{\beta^2e^{-p}} = \frac{(d-1)a^2}{e^p\theta}\biggl(e+\frac{1}{2}+\tilde{b}+\sqrt{e^2+e+2e\tilde{b}}\biggr),
\tag{27}
$$

the latter being the larger root of

$$
e^{1+p}\theta(d-1)k-\frac{(e^pk\theta-\lambda)^2}{2a^2}=0,
$$

seen as an equation in k. (Here we have substituted the definitions $\lambda = (d-1)a^2/2 + b$, and $\tilde{b} = b/(d-1)a^2$.) This definition involves k and T, through p, which might in principle threaten to confound the relation (20) already assumed. But in fact, θ appears there only as $e^p\theta$, which the definition (27) fixes as a constant. Now choose k to be greater than the quantity in (27), which is (once we have eliminated θ)

$$\begin{aligned} k &> \sqrt{\frac{8e^{1+p}A^2d^2(d-1)}{\beta^2} \cdot (d-1)a^2e^{-p}\left(e + \frac{1}{2} + \tilde{b} + \sqrt{e^2+e+2e\tilde{b}}\right)} \\ &= \sqrt{8e\left(e + \frac{1}{2} + \tilde{b} + \sqrt{e^2+e+2e\tilde{b}}\right)}(\sqrt{e}-1)^{-1}(1-2\varepsilon)^{-1}d(d-1)aA. \end{aligned}$$

The condition that k be greater than the left-hand side of (27) will make the middle sum go to zero with increasing T, while k being greater than the right-hand side will do the same for the infinite tail sum. Note that the latter bound implies (20) for T sufficiently large.

Define $\kappa(\tilde{b})$ to be

$$\kappa(\tilde{b}) := \sqrt{8e\left(e + \tfrac{1}{2} + \tilde{b} + \sqrt{e^2+e+2e\tilde{b}}\right)}(\sqrt{e}-1)^{-1}. \tag{28}$$

Since we may take ε as small as we like (but still positive), this yields an upper bound on the speed

$$K_0 = \kappa(\tilde{b})(d^2-d)aA. \tag{29}$$

Along with (22), this means that if we fix $K_0^* > K_0$, there are positive constants c and $T_0 \geq 1$ (depending on $d,\ a,\ A,\ b,\ \alpha$, but not k) such that, for any $k \geq K_0^*$ and $T \geq T_0$,

$$\mathrm{P}\{\Psi_T^* > kT\} \leq \exp\left\{-c\frac{k^2T}{(\log kT)^\alpha}\right\}. \tag{30}$$

The intermediate bound K_0^* may seem superfluous, but we need it to make this exponential bound uniform in k. As k approaches K_0, we would need to let ε go to 0, sending c to 0 and T_0 to ∞. Above K_0^* we may leave ε fixed.

Since $T \mapsto \Psi_T^*$ is nondecreasing,

$$\begin{aligned} \mathrm{P}\left\{\limsup_{T\to\infty}\frac{\Psi_T^*}{T} > K_0^*\right\} &\leq \mathrm{P}\left\{\limsup_{N\to\infty}\frac{\Psi_{N+1}^*}{N} > K_0^*\right\} \\ &= \mathrm{P}\left\{\limsup_{N\to\infty}\frac{\Psi_N^*}{N} > K_0^*\right\} \\ &\leq \mathrm{P}\{\Psi_N^* > K_0^*N \text{ infinitely often}\}, \end{aligned}$$

where N runs over positive integers. This last probability is 0, by Borel–Cantelli and (30). Since this holds for all $K_0^* > K_0$, it follows that

$$\limsup_{T\to\infty} \frac{\Psi_T^*}{T} \le K_0$$

almost surely. By the comment which began this section,

$$\limsup_{T\to\infty} \frac{\Phi_T^*}{T} \le K_0 + B,$$

completing the proof of (6). Observe that in the driftless setting, with $B = b = 0$, this speed K_0 goes to 0 as a goes to 0. This is to be expected, since $a = 0$ implies that the flow is pushing all points exactly the same way, so the whole region $\mathscr{X}$ moves like a single diffusion. This means that the growth of Φ_T^* is on the order of $\sqrt{T}$, so the linear speed should be 0.

Now consider the nearly Gaussian moments given by the functions $\rho_{\gamma,\alpha}$. Let K_0^* be as above, but required to be at least e^e, so that the function $\rho_{\gamma,\alpha'}$ is increasing in $x \ge K_0^*$ for every $\alpha' \in (\alpha, 2e)$. For $T \ge T_0$,

$$\begin{aligned}
\mathrm{E}\left[\rho_{\gamma,\alpha'}\left(\frac{\Psi_T^*}{T}\right)\right] &\le \rho_{\gamma,\alpha'}(K_0^*) + \int_{K_0^*}^{\infty} \rho'_{\gamma,\alpha'}(x)\ \mathrm{P}\left\{\frac{\Psi_T^*}{T} > x\right\} dx \\
&\le \rho_{\gamma,\alpha'}(K_0^*) \\
&\quad + \gamma \int_{K_0^*}^{\infty} \left(\frac{x(2\log x - \alpha')}{(\log x)^{\alpha'+1}}\right) \exp\left\{\frac{\gamma x^2}{(\log x)^{\alpha'}} - \frac{cTx^2}{(\log Tx)^{\alpha}}\right\} dx.
\end{aligned}$$

This is finite and decreasing in $T \ge 1$—the exponent α' is less than $2e$—so evaluating it when T is T_0 provides a uniform bound $\beta(\gamma, \alpha', \alpha)$ on the expectation for all $T \ge T_0$:

$$\mathrm{E}\left[\rho_{\gamma,\alpha'}\left(\frac{\Psi_T^*}{T}\right)\right] \le \beta(\gamma, \alpha', \alpha).$$

For $0 \le T < T_0$, $\Psi_T^* \le \Psi_{T_0}^*$, so

$$\begin{aligned}
\sup_{T\ge 0}\ \mathrm{E}\left[\rho_{\gamma,\alpha'}\left(\frac{\Psi_T^*}{T\vee 1}\right)\right] &\le \sup_{T\ge 0}\ \mathrm{E}\left[\rho_{\gamma,\alpha'}\left(\frac{\Psi_{T_0}^*}{T\vee 1}\right) \vee \exp\bigl(\gamma e^{2e-2}\bigr)\right] \\
&\le \beta(\gamma T_0^2, \alpha', \alpha) + \exp\bigl(\gamma e^{2e-2}\bigr).
\end{aligned}$$

The cutoff at $\exp(\gamma e^{2e-2})$ is there to make the function $\rho_{\gamma,\alpha'}$ monotonic. Since α was an arbitrary number between 2 and $2e$, this implies that

$$\sup_{T\ge 0}\ \mathrm{E}\left[\rho_{\gamma,\alpha'}\left(\frac{\Psi_T^*}{T\vee 1}\right)\right] < \infty$$

for all $\alpha' > 2$. Finally, we use the fact that $\Phi_T^* \le \Psi_T^* + BT$ to complete the proof.

REFERENCES

[1] ABRAMOWITZ, M. and STEGUN, I. (1965). *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

[2] ALEXANDER, K. S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12** 1041–1067.

[3] CARMONA, R. (1994). Transport properties of Gaussian vector fields. Lectures given at the Rennes SMF School on Random Media, Oct. 10–12, 1994.

[4] CARMONA, R. and CEROU, F. (1999). Transport by incompressible random velocity fields: simulations and mathematical conjectures. In *Stochastic Partial Differential Equations: Six Perspectives* (R. Carmona and B. Rozovskii, eds.) 153–181. Amer. Math. Soc., Providence, RI.

[5] CRANSTON, M. SCHEUTZOW, M. and STEINSALTZ, D. (1999). Linear expansion of isotropic Brownian flows. *Electron. Comm. in Probab.* **4** 91–101.

[6] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam.

[7] ISICHENKO, M. B. (1992). Percolation, statistical topography and transport in random media. *Rev. Mod. Phys.*

[8] KUNITA, H. (1990). *Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Math.* **24**. Cambridge Univ. Press.

[9] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach Spaces*. Springer, New York.

[10] POLLARD, D. (1989). Asymptotics via empirical processes. *Statist. Sci.* **4** 341–366.

[11] POLLARD, D. (1990). *Empirical Processes: Theory and Applications*. Hayward, CA.

[12] REVUZ, D. and YOR, M. (1990). *Continuous Martingales and Brownian Motion*, 1st ed. Springer, New York.

[13] THOMASSEN, C. (1983). Infinite graphs. In *Selected Topics in Graph Theory 2* (L. Beineke and R. Wilson, eds.) 129–160. Academic Press, London.

[14] VAN DER VAART, A. and WELLNER, J. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

[15] ZIRBEL, C. and ÇINLAR, E. (1996). Dispersion of particle systems in Brownian flows. *Adv. in Appl. Probab.* **28** 53–74.

M. CRANSTON
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ROCHESTER
ROCHESTER, NEW YORK 14627
E-MAIL: cran@math.rochester.edu

M. SCHEUTZOW
TECHNISCHE UNIVERSITÄT BERLIN
MA 7-5, STR. DES 17. JUNI 135
10623 BERLIN
GERMANY
E-MAIL: ms@math.tu-berlin.de

D. STEINSALTZ
TECHNISCHE UNIVERSITÄT BERLIN
MA 7-5, STR. DES 17. JUNI 135
10623 BERLIN
GERMANY
E-MAIL: dstein@math.tu-berlin.de