# The Trace of Spatial Brownian Motion is Capacity-equivalent to the Unit Square

Robin Pemantle[1] Yuval Peres[2,3] Jonathan W. Shapiro[2,4]

August 4, 1999

**Abstract**

We show that with probability 1, the trace $B[0,1]$ of Brownian motion in space, has positive capacity with respect to exactly the same kernels as the unit square. More precisely, the energy of occupation measure on $B[0,1]$ in the kernel $f(|x-y|)$, is bounded above and below by constant multiples of the energy of Lebesgue measure on the unit square. (The constants are random, but do not depend on the kernel.) As an application, we give almost-sure asymptotics for the probability that an $\alpha$-stable process approaches within $\epsilon$ of $B[0,1]$, conditional on $B[0,1]$.

The upper bound on energy is based on a strong law for the approximate self-intersections of the Brownian path.

We also prove analogous capacity estimates for planar Brownian motion and for the zero-set of one-dimensional Brownian motion.

**Keywords:** Brownian motion, capacity, energy, occupation measure, local time.

[1]Department of Mathematics, University of Wisconsin, Madison, WI 53706. Supported in part by National Science Foundation grant # DMS 9300191, by a Sloan Foundation Fellowship, and by a Presidential Faculty Fellowship.

[2]Department of Statistics, University of California, Berkeley, California 94720.

[3]Research partially supported by NSF grant # DMS-9404391.

[4]Research supported by a Line and Michel Loève Fellowship.

## 1 Introduction and main results

It is well-known that for $d \geq 2$, the range of $d$-dimensional Brownian motion has Hausdorff dimension 2, but its 2-dimensional measure is almost surely 0. Hausdorff dimension is defined via Hausdorff measures, but has an equally important interpretation (due to Frostman [10]) as the critical parameter for positivity of Riesz capacities. Exact Hausdorff measure is one much-studied means of specifying more precisely the size of a "small" set (see Taylor [25] for a comprehensive survey in the context of random sets); exact capacity is a different one, that is directly relevant to intersections of the small set with other random sets. Cieselski and Taylor [6] found the exact Hausdorff measure for the trace of Brownian motion in space, which quantifies to what extent the trace is "smaller" than the plane. Here we show that with probability 1, the spatial Brownian trace has positive capacity exactly in the same kernels as the plane. Theorem 1.1 is a quantitative version of this; Theorems 1.2 and 1.6 give analogous statements for planar Brownian motion and for the zero-set of 1-dimensional Brownian motion, respectively. The latter theorem sharpens an integral test due to Kahane and Hawkes.

For a decreasing kernel function $f : [0, \infty) \to [0, \infty]$, define the **energy** of a Borel measure $\nu$ on $\mathbf{R}^d$ with respect to $f$ by

$$\mathcal{E}_f(\nu) = \int_{R^d} \int_{R^d} f(|x - y|) \, \boldsymbol{d}\nu(x) \, \boldsymbol{d}\nu(y)$$

and the **capacity** of a Borel set $\Lambda \subset \mathbf{R}^d$ with respect to $f$ by

$$\mathrm{Cap}_f(\Lambda) = \left[ \inf_{\nu(\Lambda)=1} \mathcal{E}_f(\nu) \right]^{-1} .$$

Thus $\mathrm{Cap}_f(\Lambda) > 0$ if and only if there exists a Borel measure $\nu$ supported on $\Lambda$ such that $\mathcal{E}_f(\nu) < \infty$. When $f(r) = r^{-\alpha}$, we write $\mathrm{Cap}_\alpha$ for $\mathrm{Cap}_f$, and then the "capacitary dimension" $\sup\{\alpha : \mathrm{Cap}_\alpha(\Lambda) > 0\}$ of a Borel set $\Lambda$ is equal to its Hausdorff dimension (see, e.g., Carleson [4] or Kahane [12], page 133).

In the sequel we assume that all kernel functions $f$ considered are (weakly) decreasing and satisfy $\lim_{r \downarrow 0} f(r) = f(0)$ if this limit is finite.

Pemantle and Peres [18] introduced a notion of "capacity-equivalence", which we specialize to $\mathbf{R}^d$:

**Definition 1** *The sets $A, B \subset \mathbf{R}^d$ are* **capacity-equivalent** *if there exist positive constants $C_1, C_2$ such that*

$$C_1 \mathrm{Cap}_f(B) \le \mathrm{Cap}_f(A) \le C_2 \mathrm{Cap}_f(B) \qquad \text{for all } f.$$

Let $(B_t(\omega) : 0 \le t \le 1)$ be $d$-dimensional Brownian motion started at 0, and consider its range $B[0,1] = \{x \in \mathbf{R}^d : B_t = x \text{ for some } 0 \le t \le 1\}$. It is classical, and follows easily from Theorem 2.1 below (see the discussion around (9)), that for any kernel $f$, if $m$ denotes Lebesgue measure on $[0,1]^2$, then

$$\mathcal{E}_f(m) \le C \left[\mathrm{Cap}_f([0,1]^2)\right]^{-1},$$

where $C$ is an absolute constant. In particular, $[0,1]^2$ has positive capacity with respect to the kernel function $f$ if and only if

$$\int_{0+} f(r) r\, \boldsymbol{d}r < \infty.$$

Theorem 1.1 implies, *a fortiori*, that with probability 1, the same criterion holds for $B[0,1]$ in dimension $d \ge 3$, uniformly over kernels.

**Theorem 1.1** *For $d \ge 3$, the Brownian trace $B[0,1]$ is a.s. capacity-equivalent to $[0,1]^2$. More precisely, with probability 1 there exist random constants $C_1, C_2 > 0$ such that*

$$C_1 \mathrm{Cap}_f([0,1]^2) \le \mathrm{Cap}_f(B[0,1]) \le C_2 \mathrm{Cap}_f([0,1]^2) \qquad \text{for all } f. \tag{1}$$

In dimension 2, the recurrence of $(B_t)$ leads to a slight modification:

**Theorem 1.2** *For any decreasing $f$, denote $\widetilde{f}(r) = f(r) \log \frac{1}{r}$. For planar Brownian motion, with probability 1 there exist random constants $C_1, C_2 > 0$ such that*

$$C_1 \mathrm{Cap}_{\widetilde{f}}([0,1]^2) \le \mathrm{Cap}_f(B[0,1]) \le C_2 \mathrm{Cap}_{\widetilde{f}}([0,1]^2) \qquad \text{for all } f. \tag{2}$$

Our main interest in capacity is that for many stochastic processes, particularly Markov processes (see [5], [9] and the references therein) and certain fractal percolation processes (see [18]), hitting probabilities of sets are equivalent to their capacities.

The next theorem exploits this equivalence, as well as the fact that our almost-sure capacity estimates hold uniformly over all kernels. Aizenman [1] showed that if $[B]$ and $[B']$ are the traces of two independent $d$-dimensional Brownian motions started apart, then

$$\mathbf{P}\Big[\mathrm{dist}([B],[B']) < \epsilon\Big] \asymp \begin{cases} \epsilon^{d-4} & \text{if } d > 4 \\ (\log \frac{1}{\epsilon})^{-1} & \text{if } d = 4 \end{cases}$$

as $\epsilon \downarrow 0$. (Earlier, Lawler [14] had obtained precise asymptotics for the analogous problem for two random walks on $\mathbf{Z}^4$. See Albeverio and Zhou [2] for a recent refinement of Aizenman's estimates.) Theorem 2.6 of [19] contains the following generalization of Aizenman's result: if $[X^\alpha]$ and $[B]$ denote the traces of an independent $\alpha$-stable process and Brownian motion, started apart, then

$$\mathbf{P}\Big[\mathrm{dist}([B],[X^\alpha]) < \epsilon\Big] \asymp \begin{cases} \epsilon^{d-\alpha-2} & \text{if } \alpha < d-2 \\ (\log \frac{1}{\epsilon})^{-1} & \text{if } \alpha = d-2 \end{cases}$$

as $\epsilon \downarrow 0$.

We derive an almost-sure version of these estimates, uniform over $\alpha$, conditional on the Brownian motion $B$. For $0 < \alpha \leq 2$, let $\mathbf{P}^\alpha_x$ be the law of a symmetric $\alpha$-stable process $(X^\alpha_t)$ in $\mathbf{R}^d$ started at $x$, so that

$$\mathbf{E}^\alpha_x\, e^{i\lambda\cdot(X_t - x)} = e^{-|\lambda|^\alpha t}$$

for $\lambda \in \mathbf{R}^d$, and let $f^{(\alpha)}(|x-y|) = c(\alpha)|x-y|^{\alpha-d}$ be the corresponding potential density. We always consider $B$ and $X^\alpha$ to be independent. Write $[B] = B[0,1]$ and $[X^\alpha] = X^\alpha[0,\infty)$.

**Theorem 1.3** *Suppose $d \geq 3$. Let*

$$\begin{aligned} m(x,B) &= \inf_{y\in[B]} |x-y| \\ M(x,B) &= \sup_{y\in[B]} |x-y|. \end{aligned}$$

*Then for some constants $c_d, c'_d > 0$ the following is true: For a.e. Brownian path $B$ and all $x \in \mathbf{R}^d$, there exists $\epsilon_0 = \epsilon_0(B,x)$ such that, for all $0 < \epsilon < \epsilon_0$,*

$$c_d\, M(x,B)^{\alpha-d} \leq \frac{\mathbf{P}^\alpha_x\Big[\mathrm{dist}\Big([B],[X^\alpha]\Big) < \epsilon \,\Big|\, B\Big]}{\alpha(d-\alpha-2)\,\epsilon^{d-\alpha-2}} \leq c'_d\, m(x,B)^{\alpha-d}$$

*for all $0 < \alpha < d-2$ such that $\alpha \leq 2$, and when $d = 3,4$ also*

$$c_d\, M(x,B)^{-2} \leq \frac{\mathbf{P}^\alpha_x\Big[\mathrm{dist}\Big([B],[X^\alpha]\Big) < \epsilon \,\Big|\, B\Big]}{(d-2)\,\Big(\log\frac{1}{\epsilon}\Big)^{-1}} \leq c'_d\, m(x,B)^{-2}$$

*for $\alpha = d-2$.*

REMARK: Note the uniformity in $\alpha$ in the statement above. Even for a fixed $\alpha$, the proof of Theorem 1.3, given in section 4, requires estimating the capacity of a fixed sample path $[B]$ in infinitely many kernels simultaneously.

Theorems 1.1 and 1.2 say nothing about which measures supported on $B[0,1]$ have low energy with respect to different kernels. It turns out that, up to a random constant not dependent on the kernel, one measure fits all kernels. Let $\mu$ denote the **occupation measure** of $(B_t)$, defined by

$$\mu(\Lambda) = \int_0^1 \mathbf{1}_\Lambda(B_t)\, dt$$

for Borel sets $\Lambda \subset \mathbf{R}^d$. Clearly $\mu$ has total mass 1 and is supported on $B[0,1]$. Roughly speaking, for questions of capacity, $\mu$ plays the same role for $B[0,1]$ that Lebesgue measure plays for $[0,1]^2$. More precisely, the lower bounds on $\mathrm{Cap}_f(B[0,1])$ in (1) and (2) follow directly via Theorem 2.1 from the next theorem, which says that, with probability one, the energy of $\mu$ on $B[0,1]$ is bounded by a random constant times the energy of Lebesgue measure on the unit square, uniformly over kernels.

**Theorem 1.4** *With probability one, there exists a $C = C(\omega)$ such that*

$$\mathcal{E}_f(\mu) \le C \begin{cases} \int_0^1 r f(r)\, dr & , \ d \ge 3 \\ \int_0^1 r \log \frac{1}{r} f(r)\, dr & , \ d = 2 \end{cases} \quad \text{for all } f. \tag{3}$$

A key tool for the proof of the above theorems is a simple formula for energy proved in Benjamini and Peres [3] (for logarithmic energy) and in Pemantle and Peres [18] (for general kernels), which we state later as Theorem 2.1. As we will show, the upper bounds on capacity given in Theorems 1.1 and 1.2 follow easily from known asymptotics for the volumes of Wiener sausages. The lower bounds on capacities are, as we illustrate in section 3, easily proved for *fixed* kernels, but the fact that, with probability one, these bounds hold uniformly over kernels, is new. The proofs use Theorem 2.1 together with Theorem 1.5 below. The proof of Theorem 1.3, given in section 4, is similar, and uses the additional deterministic fact that the capacity of an $\epsilon$-sausage is equivalent to the capacity of the original set with respect to an $\epsilon$-smoothed kernel (Proposition 4.1), together with the equivalence of capacities and hitting probabilities for stable processes.

For $\sigma > 0$ and $y \in \mathbf{R}^d$, define

$$g_\sigma(y) = \exp(-|y|^2/2\sigma^2),$$

where $|\cdot|$ denotes Euclidean norm. Let

$$\begin{aligned} S_\sigma &= \int_0^1 \boldsymbol{d}t_1 \int_0^1 \boldsymbol{d}t_2 \; g_\sigma(B_{t_1} - B_{t_2}) \\ &= \int_{R^d}\int_{R^d} g_\sigma(x-y)\, \boldsymbol{d}\mu(x)\, \boldsymbol{d}\mu(y)\,. \end{aligned}$$

When suitably scaled, $S_\sigma$ may be interpreted as measuring the "approximate self-intersections" of the Brownian path. The case $d = 2$ of the following theorem follows from Varadhan's renormalization of $S_\sigma$ (see section 6 for details).

**Theorem 1.5 (A strong law for approximate self-intersections)** *For $d \geq 3$,*

$$\frac{S_\sigma}{\sigma^2} \to \frac{4}{d-2} \qquad \textit{as } \sigma \downarrow 0\ ,\ \textit{a.s.}$$

*In dimension 2,*

$$\frac{S_\sigma}{\sigma^2 \log \frac{1}{\sigma}} \to 4 \qquad \textit{as } \sigma \downarrow 0\ ,\ \textit{a.s.}$$

To explain the connection to the energy estimates in Theorem 1.4, we start with the observation that the ratio $\mu(Q)/\text{side}(Q)^2$ cannot be uniformly bounded as $Q$ ranges over all cubes, since the 2-dimensional Hausdorff measure of the Brownian trace vanishes. The $\mu$-weighted average of this ratio, taken over the collection $\mathcal{D}_n$ of all dyadic cubes $Q$ of side $2^{-n}$, is

$$4^n \sum_{Q \in \mathcal{D}_n} \mu(Q)^2. \tag{4}$$

Theorem 1.5 implies that, in dimension $d \geq 3$, these weighted averages are bounded uniformly in $n$. (See the inequality (15) in Subsection 3.1.) Theorem 2.1 is then used to express the energy of $\mu$ as a positive linear combination of the averages (4), and thus to compare it to the energy of Lebesgue measure on the unit square.

## 1.1 The zero set

We have analogous results for the zero set of one-dimensional Brownian motion, $Z = \{t \in [0,1] : B_t = 0\}$. These results are technically easier than the corresponding ones for the Brownian trace, and led us to the latter. It is classical that $Z$ a.s. has Hausdorff dimension 1/2 (again, with zero measure in that dimension), so here a natural comparison set is the "middle-1/2 Cantor set"

$$K = \left\{ \sum_{n=1}^{\infty} b_n 4^{-n} : b_n = 0, 3 \right\}.$$

$K$ is a standard example of a set of Hausdorff dimension 1/2, that has positive and finite measure in that dimension.

**Theorem 1.6** *The Brownian zero-set $Z$ is a.s. capacity-equivalent to the middle-1/2 Cantor set $K$. More precisely, with probability one there exist random $C_1, C_2 > 0$, such that*

$$C_1 \mathrm{Cap}_f(K) \le \mathrm{Cap}_f(Z) \le C_2 \mathrm{Cap}_f(K) \qquad \text{for all } f. \tag{5}$$

Let $(\ell(t) : 0 \le t \le 1)$ be Brownian local time at zero, normalized so that, by results of Lévy, $\ell(t)$ has the same law as the running maximum $\max_{\tau \le t} B_\tau$. We abuse notation slightly and also let $\ell$ denote the measure, supported on $Z$, for which it is the distribution function.

The lower bound on $\mathrm{Cap}_f(Z)$ in (5) is implied by the following energy estimate:

**Theorem 1.7** *With probability one there exists $C = C(\omega)$ such that:*

$$\mathcal{E}_f(\ell) \le C \int_0^1 f(r) r^{-1/2} \, \boldsymbol{d}r, \quad \text{for all } f. \tag{6}$$

In the first (1968) edition of [12], Kahane established that, for a fixed $f$ of "positive type", finiteness of the integral in (6) is sufficient for the Brownian zero set $Z$ to a.s. have positive capacity with respect to $f$ (see [12] page 236, Theorem 2). This is the first "exact capacity" result we are aware of. Hawkes ([11] Theorem 5) proved the converse (finiteness of the integral is necessary for positive capacity) under a slightly stronger assumption (log-convexity) on the kernel $f$. In view of the expression (10) for the capacity of $K$, Theorem 1.6 is a uniform version of this result of Kahane and Hawkes; it also shows that the side conditions on the kernel are not needed. In the last section we describe a different random set that illustrates why the uniformity in the kernel is not automatic.

## 2 Upper bounds on capacities

The following representation of energy from [18] is basic for most of the results in this paper. Its proof is based on a trick from [3]. Let $\mathcal{D}_n$ denote the collection of all dyadic cubes $Q = [j_1 2^{-n}, (j_1+1)2^{-n}) \times \ldots \times [j_d 2^{-n}, (j_d+1)2^{-n})$ for $(j_1, \ldots, j_d) \in \mathbf{Z}^d$.

**Theorem 2.1 ([18], Theorem 3.1)** *Let $f : [0,\infty) \to [0,\infty]$ be a weakly decreasing function. Then for any Borel measure $\nu$ supported on the unit cube $[0,1]^d$,*

$$\mathcal{E}_f(\nu) \asymp \sum_{n=0}^{\infty} (f(2^{-n}) - f(2^{1-n})) \sum_{Q \in \mathcal{D}_n} \nu(Q)^2, \tag{7}$$

*where $\asymp$ means that the ratio of the two quantities is bounded between two positive constants depending only on $d$.*

Remark: The proof of this in [18] assumes that $f(0+) = \infty$ and that $\nu$ has no atoms, but these assumptions can be avoided as long as $f(0) = f(0+)$. If $\nu$ has atoms at the points $\{x_j\}_{j \geq 1}$, then there is a contribution of $\sum_j f(0)\nu(x_j)^2$ to the energy $\mathcal{E}_f(\nu)$ coming from the diagonal. On the right-hand side of (7), we get the same contribution.

We first note an easy general upper bound on capacity, which is essentially the same as Theorem IV.2 in Carleson [4]. Let $N_n(\Lambda)$ be the number of dyadic cubes $Q \in \mathcal{D}_n$ (as defined in Theorem 2.1) that intersect a Borel set $\Lambda \subset \mathbf{R}^d$. Then there is a constant $c > 0$, depending only on the ambient dimension $d$, such that for any probability measure $\nu$ supported on $\Lambda$, and any kernel $f$, we have

$$\begin{aligned} \mathcal{E}_f(\nu) \;&\geq\; c \sum_n \left(f(2^{-n}) - f(2^{1-n})\right) \sum_{Q \in \mathcal{D}_n} \nu(Q)^2 \\ &\geq\; c \sum_n \left(f(2^{-n}) - f(2^{1-n})\right) N_n(\Lambda)^{-1}. \end{aligned}$$

Therefore

$$\mathrm{Cap}_f(\Lambda) \leq c^{-1} \left[ \sum_n \left(f(2^{-n}) - f(2^{1-n})\right) N_n(\Lambda)^{-1} \right]^{-1}. \tag{8}$$

If for some $c$, the set $\Lambda$ carries a positive measure $\nu$ such that $\nu(Q) \leq c N_n(\Lambda)^{-1}$ for all $Q \in \mathcal{D}_n$ and all $n$, then this bound is sharp (up to a constant factor independent of $f$).

Thus we get

$$\mathrm{Cap}_f([0,1]^2) \asymp \left[\sum_n (f(2^{-n}) - f(2^{1-n}))4^{-n}\right]^{-1} \asymp \left[\int_0^1 f(r)r\,\boldsymbol{d}r\right]^{-1} \tag{9}$$

and similarly, for the middle-half Cantor set

$$\mathrm{Cap}_f(K) \asymp \left[\sum_n (f(2^{-n}) - f(2^{1-n}))2^{-n/2}\right]^{-1} \asymp \left[\int_0^1 f(r)r^{-1/2}\,\boldsymbol{d}r\right]^{-1}, \tag{10}$$

where $\asymp$ means that the ratio of the two sides is bounded above and below by positive absolute constants. The minimum energies are attained within a constant factor by Lebesgue measure in the case of $[0,1]^2$, and, for $K$, by the measure that makes the digits $(b_n) \sim$ i.i.d Bernoulli(1/2), when $K$ is represented as $\{\sum_{n=1}^\infty b_n 4^{-n} : b_n = 0, 3\}$.

**Proof of Theorems 1.1 and 1.2 - upper bound:** Strong laws for volumes of Wiener sausages (see [16] Chapter VI and the references therein) imply that, with probability one, there exist random $C_1, C_2 \in (0,\infty)$ such that for all $n$,

$$\begin{aligned} C_1 &\leq \tfrac{N_n(B[0,1])}{4^n} \leq C_2 \quad \text{for} \quad d \geq 3 \\ C_1 &\leq \tfrac{n \cdot N_n(B[0,1])}{4^n} \leq C_2 \quad \text{for} \quad d = 2. \end{aligned} \tag{11}$$

Substituting the above into (8) and comparing with (9) gives, with probability one,

$$\mathrm{Cap}_f(B[0,1]) \leq C(\omega) \begin{cases} \mathrm{Cap}_{\tilde{f}}([0,1]^2) \ , \ d = 2 \\ \mathrm{Cap}_f([0,1]^2) \ , \ d \geq 3 \end{cases} \qquad \text{for all } f,$$

where $\tilde{f}$ is defined in the statement of Theorem 1.2. □

**Proof of Theorem 1.6 - upper bound:** We need an analog of (11) for $Z$. This is provided by Kingman's [13] construction of local time, which we sketch here for the Brownian case. Recall Lévy's classical result (see, e.g., [21] page 447)

$$\delta^{1/2}\tilde{N}_\delta \to \left(\frac{2}{\pi}\right)^{1/2} \ell(1) \quad , \text{ as } \delta \downarrow 0, \text{ a.s.} \tag{12}$$

where $\tilde{N}_\delta$ is the number of maximal intervals $I_j$ of $[0,1]\backslash Z$ having length greater than $\delta$. Now if

$$Z^\delta = \{u \in [0,1] : B_t = 0 \text{ for some } t \text{ with } |u - t| < \delta/2\} \ ,$$

and $m$ denotes Lebesgue measure on $\mathbf{R}^+$, then, using the fact that $m(Z) = 0$ a.s., we obtain

$$m(Z^\delta) = \sum_j [m(I_j) \wedge \delta] + O(\delta)\,,$$

where the sum extends over all maximal intervals in $[0,1] \setminus Z$. By Fubini's theorem, this sum can be written as $\int_0^\delta \widetilde{N}_\epsilon \, \boldsymbol{d}\epsilon + O(\delta)$. Together with (12), this implies that

$$\delta^{-1/2} m(Z^\delta) \to 2 \left(\frac{2}{\pi}\right)^{1/2} \ell(1) \qquad \text{as } \delta \downarrow 0, \text{ almost surely.}$$

Thus for suitable absolute constants $c_1, c_2 > 0$, there almost surely exists a random integer $n^*$, such that

$$c_1 \ell(1) \leq 2^{-n/2} N_n(Z) \leq c_2 \, \ell(1) \qquad \text{for all}\, n \geq n^* . \tag{13}$$

The upper bound on $\mathrm{Cap}_f(Z)$ now follows from the general upper bound (8) and the estimate (10). □

## 3 Lower bounds on capacities

**Remark:** For a *fixed* kernel, it is easy to see that finiteness of the integral on the right hand side of (3) or (6) implies that the left hand side is finite. We show this for (3) in the case $d = 3$; the other proofs are similar. Recall that for any non-negative Borel function $h : \mathbf{R}^3 \to \mathbf{R}$,

$$\mathbf{E} \int_0^\infty h(B_t) \, \boldsymbol{d}t = \frac{1}{2\pi} \int_{R^3} h(x) \frac{\boldsymbol{d}x}{|x|} \,, \tag{14}$$

where $|\cdot|$ is the Euclidean norm and $\boldsymbol{d}x$ denotes Lebesgue measure. By the Markov property, we have

$$\mathbf{E}\mathcal{E}_f(\mu) \ \leq \ 2\mathbf{E} \int_0^1 f(|B_t|) \, \boldsymbol{d}t$$

Since $f$ is monotone decreasing, $f(|x|) \leq f(|x|) \mathbf{1}_{\{|x| \leq 1\}} + f(1)$. Invoking (14), we get

$$\mathbf{E}\mathcal{E}_f(\mu) \ \leq \ 2 \left( \frac{1}{2\pi} \int_0^1 f(r) \cdot 4\pi r^2 \, \frac{\boldsymbol{d}r}{r} + f(1) \right) \ \leq \ (4+4) \int_0^1 f(r) \, r \, \boldsymbol{d}r,$$

where the last step used the monotonicity of $f$ again. □

### 3.1 The Brownian trace

Proof of Theorem 1.4: Recall that $\mathcal{D}_n$ is the collection of dyadic squares of side $2^{-n}$. For $\sigma = 2^{-n}$ we have, by the definition of $S_\sigma$, that

$$S_\sigma \geq \sum_{Q \in \mathcal{D}_n} \int_Q \int_Q g_\sigma(x-y) \, \boldsymbol{d}\mu(x) \, \boldsymbol{d}\mu(y) \,.$$

All the integrands on the right hand side are bounded below by a positive constant $c = c(d)$ which does not depend on $n$. Hence by Theorem 1.5, there is a random constant $C' = C'(\omega)$ such that, with probability one, for all $n$

$$\sum_{Q \in \mathcal{D}_n} \mu(Q)^2 \leq c^{-1} S_{2^{-n}} \leq C' \begin{cases} 4^{-n} & , \quad d \geq 3 \\ n4^{-n} & , \quad d = 2 \end{cases} . \tag{15}$$

Thus, by Theorem 2.1, with probability 1

$$\mathcal{E}_f(\mu) \quad \leq \quad C' c_1 \sum_{n=n_0}^{\infty} \Big( f(2^{-n}) - f(2^{1-n}) \Big) \begin{cases} 4^{-n} & , \quad d \geq 3 \\ n4^{-n} & , \quad d = 2 \end{cases} ,$$

where $n_0 = n_0(\omega)$ is defined by $2^{-n_0} \geq$ diameter $(B[0,1]) > 2^{-n_0-1}$, and $c_1$ depends only on $d$. Since $f$ is monotone decreasing, by adjusting $C'$ we may replace $n_0$ by 1 in the above sum, and we obtain (3) after a summation by parts. $\square$

### 3.2 The zero set

We first prove a proposition, which, loosely speaking, will play the role that Theorem 1.5 did in the previous proof. Recall that $\ell(\cdot)$ denotes local time at 0.

**Proposition 3.1** *Consider the quadratic variation of $\ell$ at scale $\delta$:*

$$L_\delta \quad = \quad \sum_{j=0}^{\lceil \delta^{-1} \rceil} [\ell((j+1)\delta) - \ell(j\delta)]^2 .$$

*With probability 1, there exists a random $C = C(\omega)$ such that*

$$L_\delta \leq C\delta^{1/2} \qquad \text{for all } \delta > 0. \tag{16}$$

PROOF: We consider separately the summands for odd and even $j$ in $L_\delta$. Denote one-dimensional Brownian motion by $B_t$. For fixed $\delta > 0$, let $j_1, j_2, \ldots,$ be a left-to-right enumeration of all the <u>odd</u> $j \geq 1$ such that $B_t = 0$ for some $t$ in the interval $[(j-1)\delta\, ,\, j\delta]$. Let $M(\delta) := \max\{i \,:\, j_i\delta \leq 1 + \delta\}$ be the number of these intervals which intersect $[0,1]$.

Define stopping times $T_i = \inf\{t \in [(j_i - 1)\delta\, ,\, j_i\delta] \,:\, B_t = 0\}$, and let $X_i := \ell(T_i + \delta) - \ell(T_i)$. The strong Markov property at the times $T_i$ implies that, for fixed $\delta$, the variables $\{X_i\}_{i \geq 1}$

are i.i.d. with the law of $\ell(\delta)$, which is the same as the law of $|B_\delta|$. In particular $X_i^2$ have mean $\delta$ and exponentially decaying tails. Thus the partial sums $Y_k(\delta) := \sum_{i=1}^k X_i$ satisfy

$$\mathbf{P}\Big(Y_k(\delta) > 2k\delta\Big) \le e^{-ck} \text{ for some constant } c > 0. \tag{17}$$

By the argument leading to (13), with probability 1 there exists a $\delta^* = \delta^*(\omega)$ such that

$$M(\delta) \le c_2 \ell(1)\delta^{-1/2} \quad \text{for all } \delta < \delta^*, \tag{18}$$

with $c_2$ an absolute constant.

Denote $Y^{(n)} := Y_{M(2^{-n})}(2^{-n})$. Since $k = k(n) = c_2\ell(1)2^{n/2}$ is eventually larger than $n$, we see that

$$\begin{aligned}&\mathbf{P}\Big[Y^{(n)} > 2c_2\ell(1)2^{-n/2}\ \textbf{i.o.}\Big] \ \le \\ &\mathbf{P}\Big[M(2^{-n}) > c_2\ell(1)2^{n/2}\ \textbf{i.o.}\Big] \ + \ \mathbf{P}\Big[\text{for infinitely many } n,\ \exists k > n \ : \ Y_k(2^{-n}) > 2k2^{-n}\Big].\end{aligned}$$

The first probability in the sum vanishes by (18), and the second by (17) and Borel-Cantelli. Thus a.s. there is a random constant $A = A(\omega)$ such that $Y^{(n)} \le A2^{-n/2}$ for all $n$. Now $Y^{(n)}$ is an upper bound for the sum over all odd indices $j$ in the quadratic variation $L_{2^{-n}}$, and the even indices are handled similarly. Consequently $2^{n/2}L_{2^{-n}}$ is a.s. bounded by a random constant.

To go from the powers of $1/2$ to general $\delta$, observe that any interval $I$ can be covered by three shorter dyadic intervals, say $J_1, J_2, J_3$. Clearly $\ell(I)^2 \le 3(\ell(J_1)^2 + \ell(J_2)^2 + \ell(J_3)^2)$. Therefore, if $2^{1-n} > \delta \ge 2^{-n}$ then $L_\delta \le 6L_{2^{-n}}$. This concludes the proof. □

Proof of Theorem 1.7: Follow the proof of Theorem 1.4 given in section 3.1, replacing $\mu$ by $\ell$ and using Proposition 3.1. □

## 4 Probabilities of $\epsilon$-approach

In this section we prove Theorem 1.3. The next deterministic proposition states that the capacity of an $\epsilon$-sausage is equivalent to the capacity of the original set with respect to an $\epsilon$-smoothed kernel. More precisely, given a kernel function $f$ and $\epsilon > 0$, let

$$\overline{f}(\epsilon) = \epsilon^{-d} d \int_0^\epsilon f(s)\, s^{d-1}\, \boldsymbol{ds},$$

and define

$$f_\epsilon(r) = \begin{cases} f(r) & \text{if } r \geq \epsilon \\ \overline{f}(\epsilon) & \text{if } r < \epsilon \end{cases}.$$

Note that $f_\epsilon$ is decreasing, since $f$ is. Also, $\overline{f}(\epsilon) < \infty$ provided that $\mathrm{Cap}_f(\mathbf{R}^d) > 0$, which we may always assume.

For a Borel set $\Lambda \subset \mathbf{R}^d$, we denote the $\epsilon$-sausage about $\Lambda$ by

$$\Lambda_\epsilon = \{x : |x - y| < \epsilon \quad \text{for some } y \in \Lambda\}.$$

Recall that "$\asymp$" ("is comparable to") means that the two quantities are within finite positive constant multiples of each other, the constants depending only on the dimension $d$. Similarly, the expression "$a \precsim b$" will mean "$a \leq c_d\, b$". We also use the notation $Q \in \mathcal{D}_n$ for dyadic cubes introduced at the beginning of section 2.

**Proposition 4.1** *For any Borel set $\Lambda \subset \mathbf{R}^d$, kernel function $f$, and $\epsilon > 0$, we have*

$$\mathrm{Cap}_f(\Lambda_\epsilon) \asymp \mathrm{Cap}_{f_\epsilon}(\Lambda). \tag{19}$$

Proof: It clearly suffices to prove the proposition for compact $\Lambda$. We first show that the left-hand side of (19) is, up to a constant factor, greater than the right. Given a probability measure $\nu$ on $\Lambda$, it is natural to smooth it by convolving with normalized Lebesgue measure on a ball of radius $\epsilon$. It will be even easier to control a discrete version of this convolution. Choose $m_\epsilon$ and $n_\epsilon$ so that

$$\begin{aligned} 2^{-m_\epsilon} &< \epsilon \leq 2^{-m_\epsilon+1} \qquad \text{and} \\ \sqrt{d}2^{-n_\epsilon} &< \epsilon \leq \sqrt{d}2^{-n_\epsilon+1}. \end{aligned}$$

Observe that the definition of $\overline{f}(\epsilon)$ and the monotinicity of $f$ imply that

$$\overline{f}(\epsilon) \asymp \sum_{n \geq m_\epsilon} f(2^{-n}) 2^{d(m_\epsilon - n)}. \tag{20}$$

Define a smoothed probability measure $\nu_\epsilon$ by

$$\boldsymbol{d}\nu_\epsilon \Big|_Q = 2^{n_\epsilon d}\, \nu(Q)\, \boldsymbol{d}x \Big|_Q, \qquad \text{for } Q \in \mathcal{D}_{n_\epsilon},$$

where $\boldsymbol{d}x$ denotes Lebesgue measure.

Suppose $\nu$ is supported on $\Lambda$; then $\nu_\epsilon$ is supported on $\Lambda_\epsilon$. Note that for every $n$ we have

$$\sum_{Q\in\mathcal{D}_n} \nu_\epsilon(Q)^2 \le 2^{d(n_\epsilon-n)} \sum_{Q\in\mathcal{D}_{n_\epsilon}} \nu(Q)^2\,;$$

indeed for $n \ge n_\epsilon$ the two sides are clearly equal, while for $n < n_\epsilon$ the inequality follows from Cauchy-Schwarz, since every $Q \in \mathcal{D}_n$ is the union of $2^{d(n_\epsilon-n)}$ cubes in $\mathcal{D}_{n_\epsilon}$. Thus using (7) to expand $\mathcal{E}_f(\nu_\epsilon)$ gives

$$\begin{aligned}\mathcal{E}_f(\nu_\epsilon) \quad &\precsim \quad \sum_{n<m_\epsilon}\Big(f(2^{-n})-f(2^{1-n})\Big)\sum_{Q\in\mathcal{D}_n}\nu(Q)^2 &(21)\\ &\quad + \sum_{n\ge m_\epsilon}\Big(f(2^{-n})-f(2^{1-n})\Big)2^{d(n_\epsilon-n)}\sum_{Q\in\mathcal{D}_{n_\epsilon}}\nu(Q)^2\,.\end{aligned}$$

Since $2^{dn_\epsilon} \asymp 2^{dm_\epsilon}$, by (20) the last line is comparable to

$$\Big(\overline{f}(\epsilon)-f(2^{1-m_\epsilon})\Big)\sum_{Q\in\mathcal{D}_{n_\epsilon}}\nu(Q)^2\,.$$

Invoking (7) again, we infer that

$$\mathcal{E}_f(\nu_\epsilon) \ \precsim\ \mathcal{E}_{f_\epsilon}(\nu)\,. \tag{22}$$

(The reverse inequality $\succsim$ also holds, but we will not need it.) The asserted inequality $\mathrm{Cap}_f(\Lambda_\epsilon)^{-1} \ \precsim\ \mathrm{Cap}_{f_\epsilon}(\Lambda)^{-1}$ now follows by taking the infimum in (22) as $\nu$ ranges over probability measures on $\Lambda$.

To obtain the reverse inequality, we use a Borel-measurable mapping $\pi : \Lambda_\epsilon \to \Lambda$, which moves every point by at most $\epsilon$. For instance, $\pi(x)$ can be defined as the lexicographically minimal $y \in \Lambda$ such that $|y - x| \le \epsilon$.

Suppose that $\nu$ is a probability measure on $\Lambda_\epsilon$, and consider the projected measure $\nu\pi^{-1}$ on $\Lambda$. As before, we have

$$\begin{aligned}\mathcal{E}_{f_\epsilon}(\nu\pi^{-1}) \ &\asymp\ \sum_{n<m_\epsilon}\Big(f(2^{-n})-f(2^{1-n})\Big)\sum_{Q\in\mathcal{D}_n}\nu\pi^{-1}(Q)^2\\ &\quad + \Big(\overline{f}(\epsilon)-f(2^{1-m_\epsilon})\Big)\sum_{Q\in\mathcal{D}_{m_\epsilon}}\nu\pi^{-1}(Q)^2\,. &(23)\end{aligned}$$

Now for each cube $Q \in \mathcal{D}_n$, the preimage $\pi^{-1}(Q)$ is contained in the union of the cubes $Q' \in D_n$ such that $\mathrm{dist}(Q', Q) < \epsilon$. If $n \le m_\epsilon$, then there are at most $5^d$ such cubes $Q'$, and hence by Cauchy-Schwarz,

$$\nu\pi^{-1}(Q)^2 \le 5^d \sum_{Q'\in\mathcal{D}_n}\nu(Q'\cap\pi^{-1}Q)^2\,.$$

Therefore for $n \le m_\epsilon$,

$$\sum_{Q \in \mathcal{D}_n} \nu\pi^{-1}(Q)^2 \le 5^d \sum_{Q' \in \mathcal{D}_n} \Big( \sum_{Q \in \mathcal{D}_n} \nu(Q' \cap \pi^{-1}Q)^2 \Big) \le 5^d \sum_{Q' \in \mathcal{D}_n} \nu(Q')^2 . \tag{24}$$

Combining (23) and (24), we get

$$\begin{aligned} \mathcal{E}_{f_\epsilon}(\nu\pi^{-1}) \quad &\precsim \quad \sum_{n<m_\epsilon} \Big( f(2^{-n}) - f(2^{1-n}) \Big) \sum_{Q \in \mathcal{D}_n} \nu(Q)^2 \\ &\qquad + \Big( \overline{f}(\epsilon) - f(2^{1-m_\epsilon}) \Big) \sum_{Q \in \mathcal{D}_{m_\epsilon}} \nu(Q)^2 . \end{aligned} \tag{25}$$

On the other hand, we can use Cauchy-Schwarz to bound the energy $\mathcal{E}_f(\nu)$ from below:

$$\begin{aligned} \mathcal{E}_f(\nu) \quad &\succsim \quad \sum_{n<m_\epsilon} \Big( f(2^{-n}) - f(2^{1-n}) \Big) \sum_{Q \in \mathcal{D}_n} \nu(Q)^2 \\ &\qquad + \sum_{n \ge m_\epsilon} \Big( f(2^{-n}) - f(2^{1-n}) \Big) 2^{m_\epsilon - n} \sum_{Q \in \mathcal{D}_{m_\epsilon}} \nu(Q)^2 . \end{aligned} \tag{26}$$

By using (20) to compare (25) and (26), we see that

$$\mathcal{E}_{f_\epsilon}(\nu\pi^{-1}) \precsim \mathcal{E}_f(\nu),$$

and taking the infimum over probability measures $\nu$ on $\Lambda_\epsilon$ completes the proof. □

Next, we recall the well-known quantitative version of the classical equivalence between the capacity of a set and its probability of being hit by a stable process. As in the introduction, let $\mathbf{P}_x^\alpha$ denote the law of a symmetric $\alpha$-stable process $(X_t^\alpha)$ started at $x \in \mathbf{R}^d$ with potential density $f^{(\alpha)}(|x-y|) = c(\alpha)\,|x-y|^{\alpha-d}$ and trace $[X^\alpha]$.

**Proposition 4.2** *(see, e.g., [24] Lemma 2, or [19] Proposition 3.2) Let $\Lambda$ be any Borel subset of $\mathbf{R}^d$, and suppose there are positive numbers $k$ and $K$ such that $k \le f^{(\alpha)}(|x-y|) \le K$ for all $y \in \Lambda$. Then*

$$k\,\mathrm{Cap}_{f^{(\alpha)}}(\Lambda) \le \mathbf{P}_x^\alpha\Big[[X^\alpha] \cap \Lambda \ne \emptyset\Big] \le K\,\mathrm{Cap}_{f^{(\alpha)}}(\Lambda).$$

PROOF OF THEOREM 1.3: Recall the notation $\overline{f}(\epsilon)$ and $f_\epsilon$ introduced at the beginning of this section. The proof begins similarly to that of Theorem 1.4. By Theorem 1.5, for some fixed constants $c$ and $c' > 0$, with probability 1 there exists $n_* = n_*(\omega)$ such that

$$c4^{-n} \le S_{2^{-n}} \le c'4^{-n} \qquad \text{for } n > n_*. \tag{27}$$

By (7) and (15), we have

$$\mathcal{E}_{f_\epsilon^{(\alpha)}}(\mu) \precsim \left(\sum_{n\le n_*} + \sum_{n>n_*}\right)\left(f_\epsilon^{(\alpha)}(2^{-n}) - f_\epsilon^{(\alpha)}(2^{-n+1})\right) S_{2^{-n}}. \tag{28}$$

Assume that $\epsilon < 2^{-n_*}$. Then the first sum is clearly $\le f^{(\alpha)}(2^{-n_*})$. On the other hand, a simple integration shows that

$$\overline{f^{(\alpha)}}(\epsilon) = \frac{c(\alpha)}{\alpha}\,\epsilon^{\alpha-d} \quad \text{for } \epsilon > 0. \tag{29}$$

Assume now that $\alpha < d-2$. Substituting (27) into the second sum in (28), summing by parts (as in the proof of Theorem 1.4), and letting $\epsilon \downarrow 0$ shows that

$$\mathcal{E}_{f_\epsilon^{(\alpha)}}(\mu) \precsim \frac{c(\alpha)}{\alpha(d-\alpha-2)}\,\epsilon^{2+\alpha-d}$$

for all $\epsilon$ less than some $\epsilon_0(\omega)$. So, by Proposition 4.1,

$$\mathrm{Cap}_{f^{(\alpha)}}([B]_\epsilon) \asymp \mathrm{Cap}_{f_\epsilon^{(\alpha)}}([B]) \succsim \frac{\alpha(d-\alpha-2)}{c(\alpha)}\,\epsilon^{d-2-\alpha}$$

for $\epsilon < \epsilon_0$. Since $\mathrm{dist}([X^\alpha],[B]) < \epsilon$ if and only if $X^\alpha$ hits $[B]_\epsilon$, the above estimate and Proposition 4.2 establish the desired lower bound on $\mathbf{P}_x^\alpha\left[\mathrm{dist}([X^\alpha],[B]) < \epsilon \,\middle|\, B\right]$. A similar calculation handles the case $\alpha = d-2$. The proof of the upper bound is entirely analogous, using the general upper bound on capacity (8) and the strong law for volumes of Wiener sausages alluded to above (11) instead of Theorem 1.5. $\square$

# 5 Proof of the strong law for $S_\sigma$ (Theorem 1.5)

We prove Theorem 1.5 only for the case $d \ge 3$. Our elementary method also works with only minor modifications for $d = 2$, but since this case follows from Varadhan's renormalization, which has received at least four proofs ([26, 22, 15, 28]), we omit it here. **Throughout this section, we assume $d \ge 3$.**

The argument follows classical lines: Estimate the first two moments, use Chebyshev's inequality to obtain convergence along a subsequence, and interpolate. However, showing that the variance of $S_\sigma$ is of lower order than the squared mean requires some care, so we include the details.

## 5.1 Moment estimates

Define the joint probability densities
$p(t_1, \ldots, t_k; x_1, \ldots, x_k)$ by

$$\mathbf{P}(B_{t_1} \in A_1, \ldots, B_{t_k} \in A_k) = \int_{A_1 \times \ldots \times A_k} \boldsymbol{d}x_1 \ldots \boldsymbol{d}x_k \, p(t_1, \ldots, t_k; x_1, \ldots, x_k),$$

$A_i \subset \mathbf{R}^d$ Borel.

**Proposition 5.1**

$$\mathbf{E}S_\sigma = \frac{4}{d-2}\sigma^2 \; + \; \Theta_d(\sigma) \tag{30}$$

*where*

$$\Theta_d(\sigma) = \begin{cases} O(\sigma^3) & , \; d = 3 \\ O(\sigma^4 \log \frac{1}{\sigma}) & , \; d = 4 \\ O(\sigma^4) & , \; d \geq 5 \end{cases}$$

*as* $\sigma \downarrow 0$.

Proof: By definition,

$$\begin{aligned} \mathbf{E}S_\sigma &= 2\int_{0 \leq t_1 \leq t_2 \leq 1} \boldsymbol{d}t_1 \, \boldsymbol{d}t_2 \int_{(R^d)^2} \boldsymbol{d}x_1 \, \boldsymbol{d}x_2 \, p(t_1, t_2; x_1, x_2) \, g_\sigma(x_1 - x_2) \\ &= 2\int_0^1 \boldsymbol{d}s \, \frac{1-s}{(2\pi s)^{d/2}} \int_{R^d} \boldsymbol{d}y \, \exp\left[-\frac{1}{2}|y|^2(\frac{1}{s} + \frac{1}{\sigma^2})\right] , \end{aligned}$$

after changing variables $s \equiv t_2 - t_1$ and $y \equiv x_2 - x_1$ and integrating out first $x_1$ and then $t_1$. Therefore

$$\begin{aligned} \mathbf{E}S_\sigma &= 2\int_0^1 \boldsymbol{d}s \, (1-s) \left(\frac{\sigma^2}{\sigma^2 + s}\right)^{d/2} \\ &= 2\sigma^d \int_0^1 \frac{\boldsymbol{d}s}{(\sigma^2 + s)^{d/2}} - 2\sigma^d \int_0^1 \frac{s \, \boldsymbol{d}s}{(\sigma^2 + s)^{d/2}}. \end{aligned}$$

One readily checks that the first term equals the right-hand side of (30), while the second is easily bounded using

$$\int_0^1 \frac{s \, \boldsymbol{d}s}{(\sigma^2 + s)^{d/2}} \quad \leq \quad \int_0^1 \frac{\boldsymbol{d}s}{(\sigma^2 + s)^{\frac{d}{2} - 1}} \qquad \square$$

**Proposition 5.2 (The second moment)**

$$\mathbf{E}S_\sigma^2 = \left(\frac{4}{d-2}\sigma^2\right)^2 + \Theta'_d(\sigma) \tag{31}$$

*where*

$$\Theta'_d(\sigma) = \begin{cases} O(\sigma^5) & , \quad d=3 \\ O(\sigma^6 \log \frac{1}{\sigma}) & , \quad d=4 \\ O(\sigma^6) & , \quad d \geq 5 \end{cases}$$

*as* $\sigma \downarrow 0$.

In the calculations below, we always have $s, s_i, t \geq 0$. We will repeatedly use the following bound:

$$\begin{aligned} \iint_{s+t\leq 1} \boldsymbol{ds}\,\boldsymbol{dt} \frac{s}{(\sigma^2+s)^{d/2}} \frac{1}{(\sigma^2+t)^{d/2}} &\leq \int_0^1 \frac{\boldsymbol{ds}}{(\sigma^2+s)^{d/2-1}} \int_0^1 \frac{\boldsymbol{dt}}{(\sigma^2+t)^{d/2}} \\ &= \begin{cases} O(\sigma^{-1}) & , \quad d=3 \\ O(\sigma^{-2}\log\frac{1}{\sigma}) & , \quad d=4 \\ O(\sigma^{6-2d}) & , \quad d\geq 5 \end{cases} \end{aligned} \tag{32}$$

Call these orders of magnitude $\Psi_d(\sigma)$. Note that

$$\sigma^{2d}\Psi_d(\sigma) = \Theta'_d(\sigma). \tag{33}$$

**Proof of Proposition 5.2:**

$$\begin{aligned} \mathbf{E}S_\sigma^2 &= 8 \iiiint_{0\leq t_1 \leq \ldots \leq t_4 \leq 1} \boldsymbol{dt}_1 \ldots \boldsymbol{dt}_4 \iiiint_{(R^d)^4} \boldsymbol{dx}_1 \ldots \boldsymbol{dx}_4 \; p(t_1, \ldots, t_4; x_1, \ldots, x_4) \\ &\quad \times \{g_\sigma(x_1-x_2)g_\sigma(x_3-x_4) + g_\sigma(x_1-x_3)g_\sigma(x_2-x_4) + g_\sigma(x_1-x_4)g_\sigma(x_2-x_3)\} \\ &= 8(I_1+I_2+I_3), \end{aligned}$$

say. The calculations below show that $8I_1$ is equal to the right side of (31), and that the other integrals are of the smaller order. The latter fact makes intuitive sense: as $\sigma \downarrow 0$, the major contribution to each $I_i$ comes from the region of the time simplex where the path increments being weighted by $g_\sigma$ have small time increments. But for $I_2, I_3$, this requires that at least *three* time-increments be small simultaneously, putting us in a corner of the simplex and so losing powers of $\sigma$ asymptotically.

**Estimation of $I_1$:** Changing variables $s_i \equiv t_{i+1} - t_i$ and $y_i \equiv x_{i+1} - x_i$, and integrating out two unweighted space-time increments,

$$\begin{aligned} I_1 &= \iint_{s_1+s_3\le 1} \boldsymbol{d}s_1\, \boldsymbol{d}s_3\, \frac{(1-s_1-s_3)^2}{2} \\ &\quad \times \frac{1}{(2\pi s_1)^{d/2}} \int_{R^d} \boldsymbol{d}y_1\, \exp\left[-\frac{1}{2}|y_1|^2 \left(\frac{1}{s_1} + \frac{1}{\sigma^2}\right)\right] \\ &\quad \times \frac{1}{(2\pi s_3)^{d/2}} \int_{R^d} \boldsymbol{d}y_3\, \exp\left[-\frac{1}{2}|y_3|^2 \left(\frac{1}{s_3} + \frac{1}{\sigma^2}\right)\right] \\ &= \sigma^{2d} \iint_{s_1+s_3\le 1} \boldsymbol{d}s_1\, \boldsymbol{d}s_3\, \frac{(1-s_1-s_3)^2}{2} \frac{1}{(\sigma^2+s_1)^{d/2}} \frac{1}{(\sigma^2+s_3)^{d/2}}. \end{aligned} \tag{34}$$

Expanding (34) and using (32) and (33),

$$I_1 = \frac{1}{2}\sigma^{2d} \iint_{s_1+s_3\le 1} \boldsymbol{d}s_1\, \boldsymbol{d}s_3\, \frac{1}{(\sigma^2+s_1)^{d/2}} \frac{1}{(\sigma^2+s_3)^{d/2}} \;+\; \Theta'_d(\sigma). \tag{35}$$

To handle the first term,

$$\begin{aligned} &\int_0^1 \boldsymbol{d}s_1\, \frac{1}{(\sigma^2+s_1)^{d/2}} \int_0^{1-s_1} \boldsymbol{d}s_3\, \frac{1}{(\sigma^2+s_3)^{d/2}} \\ &= \int_0^1 \boldsymbol{d}s_1\, \frac{1}{(\sigma^2+s_1)^{d/2}} \frac{2}{d-2} \left(\sigma^{2-d} - (\sigma^2+1-s_1)^{1-\frac{d}{2}}\right). \end{aligned} \tag{36}$$

The first term of this is

$$\left(\frac{2}{d-2}\right)^2 \sigma^{4-2d} + O(\sigma^{2-d})$$

while the absolute value of the second (negative) term in (36) is bounded by integrating on $[0,1/2]$ and $[1/2,1]$ separately:

$$\int_0^1 \boldsymbol{d}s_1\, \frac{1}{(\sigma^2+s_1)^{d/2}} \frac{1}{(\sigma^2+1-s_1)^{\frac{d}{2}-1}} \le \Psi_d(\sigma)$$

with room to spare. Multiplying everything by $8 \cdot \frac{1}{2}\sigma^{2d}$ gives the right-hand side of (31).

**Estimation of $I_2$:** With the same change of variables $s_i \equiv t_{i+1} - t_i$ and $y_i \equiv x_{i+1} - x_i$, we integrate out $y_0$ and $s_0$ to obtain

$$\begin{aligned} I_2 &= \iiint_{s_1+s_2+s_3\le 1} \boldsymbol{d}s_1\, \boldsymbol{d}s_2\, \boldsymbol{d}s_3\, (1-s_1-s_2-s_3) \\ &\quad \times \frac{1}{(2\pi s_1)^{d/2}} \frac{1}{(2\pi s_2)^{d/2}} \iint_{(R^d)^2} \boldsymbol{d}y_1\, \boldsymbol{d}y_2\, \exp -\frac{1}{2}\left(\frac{|y_1|^2}{s_1} + \frac{|y_2|^2}{s_2} + \frac{|y_1+y_2|^2}{\sigma^2}\right) \\ &\quad \times \frac{1}{(2\pi s_3)^{d/2}} \int_{R^d} \boldsymbol{d}y_3\, \exp -\frac{1}{2}\left(\frac{|y_3|^2}{s_3} + \frac{|y_2+y_3|^2}{\sigma^2}\right). \end{aligned}$$

Changing variables $t \equiv s_1 + s_2, z \equiv y_1 + y_2$ and integrating out $y_1$ and $s_1$, we get

$$
\begin{aligned}
I_2 &= \iint_{t+s_3 \le 1} \boldsymbol{d}t\, \boldsymbol{d}s_3\, (1-t-s_3)t \\
&\quad \times \frac{1}{(2\pi t)^{d/2}} \int_{R^d} \boldsymbol{d}z \, \exp\left[-\frac{1}{2}|z|^2 \left(\frac{1}{t} + \frac{1}{\sigma^2}\right)\right] \\
&\quad \times \frac{1}{(2\pi s_3)^{d/2}} \int_{R^d} \boldsymbol{d}y_3 \, \exp -\frac{1}{2}\left(\frac{|y_3|^2}{s_3} + \frac{|y_2+y_3|^2}{\sigma^2}\right). \qquad (37)
\end{aligned}
$$

We bound the last factor above (line (37)) by noticing that it is

$$
\begin{aligned}
&= (2\pi\sigma^2)^{d/2} \int_{R^d} p(s_3; y_3)\, p(\sigma^2; -y_2 - y_3)\, \boldsymbol{d}y_3 \\
&= (2\pi\sigma^2)^{d/2}\, p(s_3 + \sigma^2; -y_2) \\
&\le \frac{\sigma^d}{(\sigma^2 + s_3)^{d/2}}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
I_2 &\le \sigma^{2d} \iint_{t+s_3\le 1} \boldsymbol{d}t\, \boldsymbol{d}s_3\, (1-t-s_3)t \frac{1}{(\sigma^2+t)^{d/2}} \frac{1}{(\sigma^2+s_3)^{d/2}} \\
&\le \sigma^{2d} \iint_{t+s_3\le 1} \boldsymbol{d}t\, \boldsymbol{d}s_3\, \frac{1}{(\sigma^2+t)^{\frac{d}{2}-1}} \frac{1}{(\sigma^2+s_3)^{d/2}} \\
&= \Theta'_d(\sigma)
\end{aligned}
$$

by (32) and (33).

**Estimation of $I_3$:** Similarly,

$$
\begin{aligned}
I_3 &= \iiint_{s_1+s_2+s_3\le 1} \boldsymbol{d}s_1\, \boldsymbol{d}s_2\, \boldsymbol{d}s_3\, (1-s_1-s_2-s_3) \\
&\quad \times \frac{1}{(2\pi s_2)^{d/2}} \int_{R^d} \boldsymbol{d}y_2 \, \exp\left[-\frac{1}{2}|y_2|^2 \left(\frac{1}{s_2} + \frac{1}{\sigma^2}\right)\right] \\
&\quad \times \frac{1}{(2\pi s_1)^{d/2}} \frac{1}{(2\pi s_3)^{d/2}} \int_{(R^d)^2} \boldsymbol{d}y_1\, \boldsymbol{d}y_3 \, \exp -\frac{1}{2}\left(\frac{|y_1|^2}{s_1} + \frac{|y_3|^2}{s_3} + \frac{|y_1+y_2+y_3|^2}{\sigma^2}\right).
\end{aligned}
$$

Changing variables $t \equiv s_1 + s_3, z \equiv y_1 + y_3$ and integrating out $y_1$ and $s_1$, we have

$$
\begin{aligned}
I_3 &= \iint_{t+s_2\le 1} \boldsymbol{d}t\, \boldsymbol{d}s_2\, (1-t-s_2)t \\
&\quad \times \frac{1}{(2\pi s_2)^{d/2}} \int_{R^d} \boldsymbol{d}y_2 \, \exp\left[-\frac{1}{2}|y_2|^2 \left(\frac{1}{s_2} + \frac{1}{\sigma^2}\right)\right] \\
&\quad \times \frac{1}{(2\pi t)^{d/2}} \int_{R^d} \boldsymbol{d}z \, \exp -\frac{1}{2}\left(\frac{|z|^2}{t} + \frac{|z+y_2|^2}{\sigma^2}\right).
\end{aligned}
$$

As at (37), the last factor above is bounded by $\frac{\sigma^d}{(\sigma^2+t)^{d/2}}$, and we obtain

$$
\begin{aligned}
I_3 &\leq \sigma^{2d} \iint_{t+s_2\leq 1} \boldsymbol{d}t\, \boldsymbol{d}s_2 (1-t-s_2) t \frac{1}{(\sigma^2+s_2)^{d/2}} \frac{1}{(\sigma^2+t)^{d/2}} \\
&\leq \sigma^{2d} \iint_{t+s_2\leq 1} \boldsymbol{d}t\, \boldsymbol{d}s_2 \frac{1}{(\sigma^2+s_2)^{d/2}} \frac{1}{(\sigma^2+t)^{\frac{d}{2}-1}} \\
&= \Theta_d'(\sigma)
\end{aligned}
$$

by (32) and (33).

## 5.2 Almost-sure convergence

We need the following deterministic lemma. For any Borel measure $\nu$ on $\mathbf{R}^d$ and $\sigma > 0$, define

$$S_\sigma(\nu) = \int_{R^d}\int_{R^d} g_\sigma(|x-y|)\, d\nu(x)\, d\nu(y).$$

Thus $S_\sigma = S_\sigma(\mu)$.

**Lemma 5.3** *For any Borel measure $\nu$ on $\mathbf{R}^d$, the quantity $\sigma^{-d} S_\sigma(\nu)$ is monotone decreasing in $\sigma$. In particular, $\sigma^{-d} S_\sigma$ is a.s. monotone decreasing in $\sigma$.*

Proof: Let $\widehat{\ }$ denote the Fourier transform, so that for $\xi \in \mathbf{R}^d$

$$
\begin{aligned}
\widehat{g_\sigma}(\xi) &= (2\pi)^{-d/2} \int_{\mathbf{R}^d} g_\sigma(x) e^{-i\xi\cdot x}\, \boldsymbol{d}x \\
&= \sigma^d e^{-\sigma^2|\xi|^2/2}.
\end{aligned}
$$

Then, by Plancherel's formula,

$$
\begin{aligned}
\sigma^{-d} S_\sigma(\nu) &= \sigma^{-d} \int_{\mathbf{R}^d} \widehat{g_\sigma}(\xi) |\widehat{\nu}(\xi)|^2\, d\xi \\
&= \int_{\mathbf{R}^d} e^{-\sigma^2|\xi|^2/2} |\widehat{\nu}(\xi)|^2\, d\xi.
\end{aligned}
$$

and the lemma clearly follows. $\square$

**Proof of Theorem 1.5 :**

By Propositions 5.1 and 5.2,

$$\mathbf{Var}(S_\sigma) = \begin{cases} O(\sigma^5) & , \ d=3 \\ O(\sigma^6 \log \frac{1}{\sigma}) & , \ d=4 \\ O(\sigma^6) & , \ d \geq 5 \end{cases} .$$

By Chebyshev's inequality, for any $\epsilon > 0$,

$$\mathbf{P}\Big(|S_\sigma - \mathbf{E}S_\sigma| > \epsilon\sigma^2\Big) \leq \epsilon^{-2}\sigma^{-4}\mathbf{Var}(S_\sigma) = O(\sigma).$$

The right hand side is summable as $\sigma$ runs over the sequence $\sigma_n = n^{-2}$, so by Borel Cantelli and Proposition 5.1,

$$\sigma_n{}^{-2} S_{\sigma_n} \to \frac{4}{d-2} \qquad \text{as } n \to \infty, \text{ a.s.} \tag{38}$$

Now for arbitrary positive $\sigma < 1$, choose $n$ such that $\sigma_{n+1} < \sigma \leq \sigma_n$. Then by Lemma 5.3,

$$\sigma_n^{-d}\, S_{\sigma_n} \quad \leq \quad \sigma^{-d}\, S_\sigma \quad \leq \quad \sigma_{n+1}^{-d}\, S_{\sigma_{n+1}}$$

so that

$$(\sigma/\sigma_n)^{d-2}\, \sigma_n^{-2}\, S_{\sigma_n} \quad \leq \quad \sigma^{-2}\, S_\sigma \quad \leq \quad (\sigma/\sigma_{n+1})^{d-2}\, \sigma_{n+1}^{-2}\, S_{\sigma_{n+1}} .$$

Thus $\sigma^{-2}\, S_\sigma$ is sandwiched between two expressions which tend to $\frac{4}{d-2}$ as $\sigma \downarrow 0$, and we're done. $\square$

# 6 Concluding remarks

1. The following example shows that the uniformity in capacity-equivalence statements for random sets is not automatic. Consider the random Cantor set $\Lambda$ in $[0,1]$ constructed as follows. For each $k \geq 1$, pick a random integer $n_k$ uniformly in the interval $[3^k + k, 3^{k+1} - k]$, with all picks independent; define $\Lambda$ to be the set of all sums $\sum_{n=1}^{\infty} a_n 4^{-n}$ with

$$a_n = \begin{cases} 0 & \text{for} \quad n \in (n_k - k, n_k] \\ 0,1,2,3 & \text{for} \quad n \in (n_k, n_k + k] \\ 0,3 & \qquad \text{otherwise.} \end{cases}$$

   Then it is not hard to check that for fixed $f$, with probability one, $\mathrm{Cap}_f(\Lambda) > 0$ if and only if $\int_0^1 f(r) r^{-1/2}\, \boldsymbol{dr} < \infty$. (See [20] for details.) However, $\Lambda$ is *not* capacity-equivalent to the middle-half Cantor set; indeed there exists a random kernel $f^*$

(depending on the sample $\Lambda$) that satisfies this integrability condition but gives $\mathrm{Cap}_{f^*}(\Lambda) = 0$.

2. In 1969 Varadhan [26] proved that, in dimension two, $\sigma^{-2}(S_\sigma - \mathbf{E}S_\sigma)$ converges a.s. to a well-defined random variable. This clearly implies the planar case of our Theorem 1.5. Varadhan's renormalization has received many proofs and extensions. (See [26, 22, 15, 28, 7, 8] as well as Chapter VIII of [16] and the bibliographical notes there.) Rosen [23], Remarks II-III, gives detailed calculations which are close in spirit to ours (though more Fourier-analytic), and which could probably be extended to $d \geq 3$ to yield our Theorem 1.5. While self-intersection local time exists in dimension 3 as well as dimension 2, there seems to be no analogue of Varadhan's almost-sure renormalization there. However, Yor [27] shows that the renormalized $S_\sigma$ converge *in law* for $d = 3$. Yor establishes that $\sigma^{-3}(\log \frac{1}{\sigma})^{-1/2}(S_\sigma - \mathbf{E}S_\sigma)$ converges in law (to a Gaussian) as $\sigma \downarrow 0$; this seems tighter than the estimate $\mathbf{Var}[S_\sigma] = O(\sigma^5)$ given in (31).

3. *Does Brownian motion in three-space almost surely have the property that* all *of the orthogonal projections to planes of its trace are capacity-equivalent to each other?*

4. Let $B$ and $B'$ be two independent standard Brownian motions in $\mathbf{R}^3$. The "fractal percolation" methods of [18] and [19], which are based on the results of Lyons [17], imply the following: for any fixed kernel $f$, the capacity of the intersection $\mathrm{Cap}_f(B[0,1] \cap B'[0,1])$ is almost surely positive if $\mathrm{Cap}_f([0,1]) > 0$; otherwise with probability 1 the intersection has capacity 0 in this kernel. However, these methods do not indicate if this holds uniformly in the kernel.

   *Is the intersection $B[0,1] \cap B'[0,1]$ of two independent Brownian traces in $\mathbf{R}^3$, almost-surely capacity-equivalent to $[0,1]$?*

**Acknowledgement:** We are indebted to Russell Lyons, who first alerted us to the importance of the order of quantifiers in capacity estimates for random sets.

## References


[1] Aizenman, M. (1985). The intersection of Brownian paths as a case study of a renormalization group method for quantum field theory. *Physics* **97**, 91–110.

[2] Albeverio, S. and Zhou, X. Y. (1994). Intersection properties of Brownian motions in four dimensions. *Preprint.*

[3] Benjamini, I. and Peres, Y. (1992). Random walks on a tree and capacity in the interval. *Annals Inst. Henri Poincaré* **28**, 557–592.

[4] Carleson, L. (1967). *Selected Problems on Exceptional Sets.* Van Nostrand, Princeton, New Jersey.

[5] Chung, K. L. (1973). Probabilistic approach in potential theory to the equilibrium problem. *Ann. Inst. Fourier, Grenoble* **23**, 313–322.

[6] Cieselski, Z. and Taylor, S. J. (1962). First passage times and sojourn times for Brownian motion in space and the exact Hausdorff measure of the sample path. *Trans. Amer. Math. Soc.* **103**, 434–450.

[7] Dynkin, E. B. (1985). Random fields associated with multiple points of Brownian motion. *Journal of Functional Analysis* **62**, 397–434.

[8] Dynkin, E. B. (1988). Self-intersection gauge for random walks and for Brownian motion. *Ann. Probab.* **16**, 1–57.

[9] Fitzsimmons, P.J. and Salisbury, T. (1989). Capacity and energy for multiparameter Markov processes. *Ann. Inst. Henri Poincarè, Probab.* **25** 325–350.

[10] Frostman, O. (1935) Potential d'équilibre et capacité des ensembles. *Thesis*, Lund.

[11] Hawkes, J. (1977). Local properties of some Gaussian processes. *Zeits. Wahr. verw. Geb.* **40**, 309–315.

[12] Kahane, J. P. (1985). *Some random series of functions.* Second edition, Cambridge University Press.

[13] Kingman, J. F. C. (1973). An intrinsic description of local time. *J. London Math. Soc.* (2) **6**, 725–731.

[14] Lawler, G.F. (1982). The probability of intersection of independent random walks in four dimensions. *Commun. Math. Phys.* **86**, 539–554.

[15] Le Gall, J. F. (1985). Sur le temps local d'intersection du mouvement plan et la methode de renormalisation de Varadhan. *Séminaire de Probabilités XIX.* Lect. Notes Math. v.1123 pp.314–331. Springer-Verlag.

[16] Le Gall, J. F. (1992). Some properties of planar Brownian motion. *In* J. F. Le Gall, M. I. Freidlin, *Ecole d'ete de probabilites de St.-Flour XX, 1990.* Springer-Verlag.

[17] Lyons, R. (1992). Random walks, capacity, and percolation on trees. *Ann. Probab.* **20**, 2043–2088.

[18] Pemantle, R. and Peres, Y. (1995). Galton-Watson trees with the same mean have the same polar sets. *Ann. Probab., to appear.*

[19] Peres, Y. (1995). Intersection-equivalence of Brownian paths and certain branching processes. *Comm. Math. Phys., to appear.*

[20] Peres, Y. (1995). Remarks on capacity-equivalence and intersection-equivalence of random sets. *Preprint.*

[21] Revuz, D. and Yor, M. (1991). *Continuous martingales and Brownian motion.* Springer-Verlag.

[22] Rosen, J. (1986). Tanaka's formula and renormalization for intersections of planar Brownian motion. *Ann. Probab.* **14** 1245–1251.

[23] Rosen, J. (1986). A renormalized local time for multiple intersections of planar Brownian motion. *Séminaire de Probabilités XX.* Lect. Notes Math. v.1204 pp.515–531.

[24] Taylor, S. J. (1967). Sample path properties of a transient stable process. *J. Math. Mechanics* **16**, 1229–1246.

[25] Taylor, S. J. (1986). The measure theory of random fractals. *Math. Proc. Camb. Phil. Soc.* **100**, 383–406.

[26] Varadhan, S. R. S. (1969). Appendix to "Euclidean quantum field theory" by K. Symanzik. *In* R. Jost (ed.) *Local quantum field theory.* Academic Press.

[27] Yor, M. (1985). Renormalisation et convergence en loi pour les temps locaux d'intersection du mouvement Brownien dans $\mathbf{R}^3$. *Séminaire de Probabilités XIX.* Lect. Notes Math. v.1123 pp.350–365. Springer-Verlag.

[28] Yor, M. (1986). Précisions sur l'existence et la continuité des temps locaux d'intersection du mouvement Brownien dans $\mathbf{R}^d$. *Séminaire de Probabilités XX.* Lect. Notes Math. v.1204 pp.532–541. Springer-Verlag.