THE SPINE OF THE FLEMING-VIOT PROCESS
DRIVEN BY BROWNIAN MOTION

BY KRZYSZTOF BURDZY\(^1\) AND JÁNOS ENGLÄNDER\(^2,\)\(^*\)

\(^1\)Department of Mathematics, University of Washington, burdzy@uw.edu
\(^2\)Department of Mathematics, University of Colorado, \(^*\)janos.englander@colorado.edu

We show that the spine of the Fleming-Viot process driven by Brownian motion in a bounded Lipschitz domain with Lipschitz constant less than 1 converges to Brownian motion conditioned to stay in the domain forever.

1. Introduction. Under suitable assumptions, a branching process can be decomposed into a spine and side branches. Heuristically speaking, the spine has the distribution of the driving process conditioned on non-extinction. We will prove this claim for the “Fleming-Viot branching process” introduced in [5]. In this paper, individuals follow independent Brownian motions and are killed on the boundary of a bounded Lipschitz set.

A Fleming-Viot process is an extreme case of the Moran model introduced in [24] (see [15, Def. 5.12] for the modern discussion). In the Moran model, individuals branch at a (bounded) intensity. In our model, an individual branches only when some other individual hits the “boundary” of the state space.

Our main result on the asymptotic spine distribution is limited to Fleming-Viot processes driven by Brownian motion. We conjecture that an analogous result holds for every Fleming-Viot process (perhaps under mild technical assumptions). An analogous theorem was proved for Fleming-Viot processes driven by continuous time Markov processes on finite spaces in [2]. That article contained extra results on the branching structure, missing from the present paper, namely, it was proved that the rate of branching along the spine converges to twice the rate for a generic particle and the distribution of a side branch converges to the distribution of a branching process with the critical branching rate. Weak convergence of spines to the driving process conditioned to survive forever has been recently proved in [29, Cor. 5.3] in the case when the driving process is a normally reflected diffusion in a compact domain with soft killing (see [29, Sec. 1.3]).

The literature on branching processes is huge so we will mention only a few key publications. Most of them contain extensive reference lists. A precursor of our model can be found in a paper by Moran [24]. The book by Jagers [22] is a classical treatise on branching processes and their applications to biology. A modern review of continuous time and space branching can be found in a book by Etheridge [15]. The “Evans’ immortal particle picture” was introduced in [16]. The “look-down” process was defined by Donnelly and Kurtz in [12]. Modern approaches to the spine can be found in [14] and [21]. Some of the most profound analysis of the genealogical structure of the Moran and related models appeared in [18, 11, 19, 28].

Our paper is organized as follows. Our proof is complicated so we start with a non-technical review of the proof strategy in Section 2. Section 3 contains basic definitions and a review of known results. Section 4 contains the statement of Theorem 4.1, our main result, and its proof. The proof is based on many estimates that are relegated to Section 5. We present a generalization of our main result to non-Brownian processes in Section 6. Section 7 contains a short and informal review of the results on spines of superprocesses.

MSC2020 subject classifications: Primary 60J80, 60J65.
Keywords and phrases: Fleming-Viot process; branching process; spine; h-transform.
2. Heuristic outline of the argument. Consider a fixed population of \( n \geq 2 \) individuals that move independently according to Brownian motion. They all start inside a domain \( \Lambda \) and are killed when they exit it. When an individual is killed, another individual, chosen randomly (uniformly), branches into two individuals. Therefore, the population size is always equal to \( n \).

Assume that \( \Lambda \) is a Lipschitz domain with a Lipschitz constant smaller than 1. It has been proven that there exists a single trajectory inside the branching structure, referred to as a “spine,” which extends over the entire time interval \([0, \infty)\) and never hits the boundary of \( \Lambda \) (see Sections 3.1 and 3.2).

Our argument involves “Brownian motion conditioned to stay in \( \Lambda \) forever.” To clarify this concept, one can appeal to the main theorem in [25]: the distribution of Brownian motion conditioned to stay in \( \Lambda \) for a long time is close to that of a process obtained from Brownian motion by a space-time Doob’s \( h \)-transform. We state and prove our version of Pinsky’s result in Lemma 5.7 since we require a specific order of quantifiers.

The main result of our paper is stated in Theorem 4.1: as the population size \( n \) increases to infinity, the spine processes converge in distribution to Brownian motion conditioned to stay in \( \Lambda \) forever.

Below, we give a heuristic overview of the main ideas in the proof. The proof consists of multiple steps. Two of them are the following. Fix \( \varepsilon > 0 \).

(A) (convergence of conditioned Brownian motions). For sufficiently large \( s > 0 \), there is a \( t_0 = t_0(s) > s \) such that for all \( t > t_0 \), the following holds. Let BB (“Brownian bridge”) be Brownian motion conditioned to stay in \( \Lambda \) up to time \( t \) and conditioned further to have a distribution \( \nu \) at time \( t \). Then, restricted to \([0, s]\), the (Prohorov) distance between the distribution of BB and that of Brownian motion conditioned to stay in \( \Lambda \) forever is less than \( \varepsilon \), uniformly over all probability distributions \( \nu \) on \( \Lambda \).

The second result we prove roughly states that for large \( n \), the spine trajectory up to time \( t \) is close to a particular Brownian bridge.

(B) (convergence of spines). Let \( \nu_n^\varepsilon \) denote the distribution of the \( n \)-spine (spine of the process with \( n \) individuals) at time \( t \) on \( \Lambda \). Given \( t > 0 \), for all sufficiently large \( n \) (i.e., all \( n > n_0(\varepsilon, t) \)), the distribution of the trajectory of the \( n \)-spine on \([0, t]\) is closer than \( \varepsilon \) to the distribution of Brownian motion conditioned to stay inside \( \Lambda \) up to time \( t \) and further conditioned on having distribution \( \nu_n^\varepsilon \) at time \( t \).

Although we do not know \( \nu_n^\varepsilon \), Theorem 4.1 follows from (A) and (B) because the estimate in (A) is uniform in \( \nu \).

Our proof of (B) relies heavily on [30]. The result in [30] states that when the population size \( n \) is large, the distribution of the location of a randomly (uniformly) chosen individual at time \( t \) is close to the distribution of the position at time \( t \) of the driving process conditioned to stay in \( \Lambda \) until time \( t \). Villemonais’ theorem applies to branching populations driven by very general Markov processes. This has been used in [2] to show that Villemonais’ theorem can be applied not only to the positions of individuals at time \( t \), but also to the whole trajectories (genealogies) of individuals alive at time \( t \). In other words, if time \( t > 0 \) is fixed and the population size \( n \) is large, and one chooses an individual alive at time \( t \) randomly (uniformly), then the distribution of the genealogy of this individual on \([0, t]\) is close to the distribution of the driving process on \([0, t]\) conditioned to stay in \( \Lambda \) until time \( t \).

Unfortunately, this very general theorem of Villemonais is not directly applicable in our setting. This is because, given the location of all individuals at time \( t \), the individual residing on the spine is determined by the Fleming-Viot process on the interval \([t, \infty)\), and there is no reason why the spine position should be chosen uniformly among all individuals at time \( t \) (in hindsight, our main theorem implies that it is not). Hence, in order to surmount this difficulty and make Villemonais’ result applicable to our problem, we will need to consider a randomly picked particle with the (unknown to us) distribution of the spine at time \( t \).
Villemonais proved a quantitative version of his theorem in [30], which we use to show that the genealogical trajectory of an individual chosen randomly (uniformly) in a small cube $Q \subset \Lambda$ at time $t$ has a distribution close to the distribution of Brownian motion conditioned to stay in $\Lambda$ on the time interval $[0, t]$ and reach $Q$ at time $t$. Then we prove that, for a fixed small cube, the probability of the spine passing through any individual inside the cube at time $t$ is very close to the probability of it passing through any other individual present in the same cube at that time. This is due to the fact that, given the positions of individuals in a cube at time $t$ and their positions at a slightly later time $t + \Delta t$, Brownian motions that are conditioned to connect these sets of points are almost equally likely to choose any permutation—this follows from the form of the multidimensional Gaussian distribution. Therefore, on the scale of a single small cube, the spine position is chosen almost uniformly from all individuals present in the cube at time $t$ and, therefore, the Villemonais’ estimate mentioned at the beginning of the paragraph can be applied.

We will now discuss a delicate aspect of the above reasoning. We condition the spine to be in $Q$ at a time $t + \Delta t > 0$. The event we condition on, referring to the spine, is determined by the evolution of the process only on $[t + \Delta t, \infty)$, and not on $[0, t)$. Therefore the conditioning has no bearing on the past behavior of the trajectories of the particles in $Q$. This is why for large $n$, the spine, conditioned this way is still close (in distribution) to Brownian motion conditioned on not hitting the boundary before $t$ and ending up in $Q$ at time $t$.

To use the above to determine spine’s position, we divide the domain $\Lambda$ into small cubes. Since we do not know where the spine passes at time $t$, all we can say is that the distribution of the spine on $[0, t]$ is close to the distribution of Brownian motion conditioned to stay in $\Lambda$ on $[0, t]$ and to have some (unknown to us) distribution at time $t$. But this is sufficient in view of (A).

In light of the preceding paragraphs, this completes the main argument for (B) and hence for Theorem 4.1. However, there are several loose ends that need to be taken care of.

First, trajectories may branch on the small time interval $[t, t + \Delta t]$ mentioned above, but we show that this effect is negligible.

Second, the argument based on small cubes applies only to those cubes that are far from the boundary (relative to the cube size) so that we can assume that the trajectories of individuals starting from the cube at time $t$ have a negligible chance of hitting the boundary of $\Lambda$ during the time interval $[t, t + \Delta t]$. The cubes that are close to the boundary are addressed by proving that the spine is unlikely to be near the boundary of $\Lambda$. We accomplish this by showing that there are no trajectories in the branching structure that remain close to the boundary of $\Lambda$ for an extended period of time. Therefore, as the spine is one of these trajectories, it must stay away from the boundary for much of the time.

3. Notation, definitions and known results. This section is based on [2].

Our main theorem will be concerned with Fleming-Viot processes driven by Brownian motion in $\mathbb{R}^d$. Nevertheless we need to consider Fleming-Viot processes with an abstract underlying state space because our arguments will be based on “dynamical historical processes” which are Fleming-Viot processes driven by Markov processes with values in function spaces.

Let $\Gamma$ be a topological space and let $\Lambda$ be a Borel proper subset of $\Gamma$. We will write $\Lambda^c = \Gamma \setminus \Lambda$. Let $\{B_t, t \geq 0\}$ be a continuous time strong Markov process with state space $\Gamma$ whose almost all sample paths are right continuous. For $s \geq 0$, let

$$\tau_{\Lambda,s} = \inf \{t > s : B_t \in \Lambda^c\}, \quad \tau_{\Lambda} = \tau_{\Lambda,0}.$$  \tag{3.1}

We assume that $\Lambda^c$ is absorbing, i.e., $B_t = B_{\tau_{\Lambda,s}}$ for all $t \geq \tau_{\Lambda,s}$, a.s.

We make the following assumptions.

(A1) \( \mathbb{P} (s < \tau_{\Lambda,s} < \infty \mid B_s = x) = 1 \) for all \( x \in \Lambda \) and \( s \geq 0 \).

(A2) For every \( x \in \Lambda \) and \( s \geq 0 \), the conditional distribution of \( \tau_{\Lambda,s} \) given \( \{B_s = x\} \) has no atoms.

Consider an integer \( n \geq 2 \) and a family \( \{U_k^i, 1 \leq i \leq n, k \geq 1\} \) of jointly independent \nrandom variables such that \( U_k^i \) has the uniform distribution on the set \( \{1, \ldots, n\} \setminus \{i\} \).

We will use induction to construct a Fleming-Viot type process \( X^n_t = (X_t^1, \ldots, X^n_t), t \geq 0 \), with values in \( \Lambda^n \). Let \( \tau_0 = 0 \) and consider the (possibly random) initial configuration \( (X^1_0, \ldots, X^n_0) \in \Lambda^n \). Let

\begin{equation}
X^{1,1}_t, \ldots, X^{1,n}_t, \quad t \geq 0,
\end{equation}

be independent and have the transition probabilities of the process \( B \). We assume that processes in (3.2) are independent of the family \( \{U_k^i, 1 \leq i \leq n, k \geq 1\} \). Let

\[ \tau_1 = \inf \left\{ t > 0 : \exists i_1 \leq n \ X^{1,i_1}_t \in \Lambda^c \right\}. \]

By assumption (A2), no pair of processes can exit \( \Lambda \) at the same time, so the index \( i \) in the above definition is unique, a.s.

For the induction step, assume that the families

\[ X^{j,1}_t, \ldots, X^{j,n}_t, \quad t \geq 0, \]

and the stopping times \( \tau_j \) have been defined for \( j \leq k \). For each \( j \leq k \), denote by \( i_j \) the unique index such that \( X^{j,i_j}_{\tau_j} \in \Lambda^c \). Let

\[ X^{k+1,m}_{\tau_k} = X^{k,m}_{\tau_k} \quad \text{for} \ m \neq i_k, \]

and

\[ X^{k+1,i_k}_{\tau_k} = X^{k,U^{i_k}_k}_{\tau_k}. \]

In words: it is the particle indexed by \( (k, i_k) \), i.e., \( X^{k,i_k} \), which hits the boundary and then jumps on another, randomly chosen particle inside. Let the conditional joint distribution of

\[ X^{k+1,1}_t, \ldots, X^{k+1,n}_t, \quad t \geq \tau_k, \]

given \( \{X^{j,m}_t, 0 \leq t \leq \tau_j, 1 \leq m \leq n\}, j \leq k \), and \( \{U_k^i, 1 \leq i \leq n, k \geq 1\} \), be that of \( n \) independent processes with transition probabilities of \( B \), starting from \( X^{k+1,m}_{\tau_k}, 1 \leq m \leq n \). Let

\[ \tau_{k+1} = \inf \left\{ t > \tau_k : \exists i \leq n \ X^{k+1,i}_t \in \Lambda^c \right\}. \]

We define \( X^m_t := (X^1_t, \ldots, X^n_t) \) by

\[ X^m_t = X^{k,m}_t, \quad \text{for} \ \tau_{k-1} \leq t < \tau_k, \ k \geq 1, \ m = 1, 2, \ldots, n. \]

Note that the process \( X^n \) is well defined only up to the time

\[ \tau_{\infty} := \lim_{k \to \infty} \tau_k. \]

We will say that \( X^k \) experiences branching on the interval \( [s_1, s_2] \) if some other process \( X^m \) jumps from \( \partial \Lambda \) to \( X^k \) at a time \( s \in [s_1, s_2] \). Note that the particle that is jumping is not considered to experience branching at that time.

If \( B \) is Brownian motion then the processes \( X^m, m = 1, \ldots, n \), are “driven” by independent copies \( B^m \) of \( B \) that can be distilled from \( X^m \)'s in the following way:

\begin{equation}
B^m_t = X^m_t - X^m_0 - \sum_{\tau \leq t} (X^m_{\tau} - X^m_{\tau^-}), \quad t \geq 0.
\end{equation}
3.1. Dynamical historical processes and spine. Heuristically speaking, for each \( k \in \{1, \ldots, n\} \), the “dynamical historical process” \( \{H^k_t(s), 0 \leq s \leq t\} \) (to be defined rigorously below) represents the unique path in the branching structure of the Fleming-Viot process which goes from \( X^k_t \) to one of the points \( X^1_0, \ldots, X^n_0 \) along the trajectories of \( X^1, \ldots, X^n \) and does not jump at times \( \tau_k \).

Let \( \mathcal{A} \) be the family of all sequences of the form \( ((a_1, b_1), (a_2, b_2), \ldots, (a_k, b_k)) \), where \( a_i \in \{1, \ldots, n\} \) and \( b_i \in \mathbb{N} \) for all \( i \). If \( a = ((a_1, b_1), (a_2, b_2), \ldots, (a_k, b_k)) \) then we will write \( a + (m, r) \) to denote \( ((a_1, b_1), (a_2, b_2), \ldots, (a_k, b_k), (m, r)) \). We will define a function \( \mathcal{L} : \{1, \ldots, n\} \times [0, \tau_\infty) \rightarrow \mathcal{A} \). We interpret \( \mathcal{L}(i, s) \) as a label of \( X^i_t \) so, by abuse of notation, we will write \( \mathcal{L}(X^i_t) \) instead of \( \mathcal{L}(i, s) \). We let \( \mathcal{L}(X^i_t) = ((i, 0)) \) for all \( 0 \leq s < \tau_1 \) and \( 1 \leq i \leq n \).

If \( \mathcal{L}(X^i_t) = a \) for \( \tau_{k-1} \leq s < \tau_k \), \( i \neq i_k \) and \( i \neq U^i_k \) then we let \( \mathcal{L}(X^i_t) = a \) for \( \tau_k \leq s < \tau_{k+1} \). Suppose that \( i = U^i_k \) and \( \mathcal{L}(X^i_t) = a \) for \( \tau_{k-1} \leq s < \tau_k \). Then we let \( \mathcal{L}(X^i_t) = a + (i, k) \) and \( \mathcal{L}(X^i_t) = a + (i_k, k) \) for \( \tau_k \leq s < \tau_{k+1} \).

Suppose that \( \mathcal{L}(X^i_t) = ((a_1, b_1), (a_2, b_2), \ldots, (a_k, b_k)) \) for some \( k \geq 1 \). The assumption (A1) on the driving process \( B \) implies that \( X^i_t \) will “hit” \( \Lambda^c \) at some time greater than \( t \) (more precisely, \( \ell = i_j \) for some \( j > b_k \)) with probability 1. Before that time, it may also happen that some other \( X^j_t \) will jump onto \( X^i_t \); more precisely, it may happen that \( \ell = U^j_k \) for some \( j > b_k \). Let \( \tau' \) be the minimum of all such times. From the definition of \( \mathcal{L} \) we easily infer that \( 0 = b_1 < b_2 < \ldots < b_k \) and \( \tau_{b_k} \leq t \), so that \( 0 < \tau_{b_1} < \ldots < \tau_{b_k} \leq t < \tau' \). For \( \tau_{b_m} \leq s < \tau_{b_{m+1}} \) with \( 1 \leq m \leq k \) we let \( H^i_t(s) = X^a_{s-s} \), and for \( \tau_{b_m} \leq s \leq t \) we let \( H^i_t(s) = X^a_{s} \).

We will call \( \{H^i_t(s), 0 \leq s \leq t\} \) a dynamical historical process (DHP) corresponding to \( X^i_t \). Note that \( H^i_t \) is defined for \( 1 \leq \ell \leq n \) and \( 0 \leq t < \tau_\infty \).

The spine process will be defined below the statement of Theorem 3.1. Roughly speaking, the spine is the unique DHP that extends from time 0 to time \( \tau_\infty \). The existence and uniqueness of the spine was proved in [20, Thm. 4] under very restrictive assumptions on the driving process \( B \) and under the assumption that the lifetime \( \tau_\infty \) is infinite. It was proved in [2] that the claim holds under minimal reasonable assumptions, that is, the strong Markov property of the driving process and non-atomic character of the exit time distributions.

**Theorem 3.1.** Fix some \( n \geq 2 \), suppose that \( B \) satisfies assumptions (A1)-(A2) and \( X^n_0 \in \Lambda^n \), a.s. Then, a.s., there exists a unique infinite sequence \( ((a_1, b_1), (a_2, b_2), \ldots) \) such that its every finite initial subsequence is equal to \( \mathcal{L}(X^i_s) \) for some \( 1 \leq i \leq n \) and \( s \geq 0 \).

In the notation of the theorem, we define the spine of \( X^n \) by \( J^n_t(s) = X^n_{s-s} \) for \( \tau_{b_m} \leq s < \tau_{b_{m+1}}, m \geq 1 \). We will write \( \chi(s) = a_m \).

3.2. Brownian motion-driven Fleming-Viot process. From now on we will assume that the driving process \( B \) is Brownian motion in \( \mathbb{R}^d \). We will assume that \( \Lambda \subset \mathbb{R}^d \) is an open bounded connected Lipschitz domain with the Lipschitz constant less than 1. This means that every point in \( \partial \Lambda \) has a neighborhood where \( \partial \Lambda \) can be represented as the graph of a Lipschitz function with the Lipschitz constant less than 1 in some orthonormal coordinate system. Under these assumptions \( \tau_\infty = \infty \), a.s. (see [3, 20]). We mention parenthetically that if the driving process is Brownian motion, \( \Lambda \) is a polytope and \( n = 2 \) then we also have \( \tau_\infty = \infty \), a.s. (see [3]). However, it was proved in [4] that \( \tau_\infty < \infty \), a.s., for every \( n \), for some Fleming-Viot processes driven by one-dimensional diffusions.

3.3. Dynamical historical process as a Fleming-Viot process. Let \( C([0,t], \Gamma) \) denote the space of continuous functions with values in \( \Gamma \), with the supremum norm. For a function
\( f : C([0, t], \Gamma) \to \mathbb{R} \), let

\[
\mathcal{H}_n^t(f) = \frac{1}{n} \sum_{k=1}^{n} f(H_k^t).
\]  

(3.4)

Let \( \mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_0^k} \), i.e., \( \mu_n \) denotes the empirical distribution of \( X_0^k \).

Recall definition (3.1) and let

\[
\tilde{P}^\mu_{t}(A) = \mathbb{P} \left\{ \{B_s, 0 \leq s \leq t\} \in A \mid \tau_\Lambda > t \right\}, \quad A \subset C([0, t], \Lambda),
\]

assuming that \( \mathbb{P}(B_0 \in A_1) = \mu(A_1) \) for \( A_1 \subset \Lambda \). Note that this does not imply that \( \tilde{P}^\mu_{t}(B_0 \in A_1) = \mu(A_1) \) for \( A_1 \subset \Lambda \).

In the case when \( \mu = \mu_n \), we will write \( \tilde{P}_t \) instead of \( \tilde{P}^\mu_{t} \). The corresponding expectations will be denoted \( \tilde{E}_t \) and \( \tilde{E}^\mu_{t} \). We will write \( \tilde{E}_t(f) \) instead of \( \tilde{E}_t(f(\{B_s, 0 \leq s \leq t\})) \).

The following theorem is a corollary of [30, Thm. 2.2].

**Theorem 3.2.** For \( t \geq 0 \) and any measurable function \( f : C([0, t], \Gamma) \to \mathbb{R} \) with \( \|f\|_\infty \leq 1 \),

\[
\mathbb{E} \left| \mathcal{H}_n^t(f) - \tilde{E}_t(f) \right| \leq 2 \left( 1 + \sqrt{2} \right) \left( \mathbb{E} \left( \mathbb{P}_{\mu_n} (\tau_\Lambda > t)^{-2} \right) \right)^{1/2} n^{-1/2},
\]

where \( \mathbb{P}_{\mu_n} \) represents the distribution of the driving process \( B \) with the initial distribution \( \mu_n \).

**Proof.** It has been shown in the proof of [2, Thm. 4.2] that DHP can be identified with a space-time time-homogeneous Fleming-Viot process. Hence, our theorem follows directly from [30, Thm. 2.2].

3.4. **Conditioned Brownian motion.** A major monograph discussing conditioned Brownian motion is [13]. The topic and the book are rather technical so the reader may find the presentation of the basic facts about conditioned Brownian motion and conditioned space-time Brownian motion in the introduction of [6] more accessible.

Recall that \( \Lambda \subset \mathbb{R}^d \) is a bounded Lipschitz domain. Let \( \varphi > 0 \) denote the first eigenfunction of \( (-\frac{1}{2})\Delta \), where \( \Delta \) denotes the Dirichlet Laplacian in \( \Lambda \). Let \( \lambda > 0 \) be the corresponding eigenvalue. Then the space-time Brownian motion \( (B_t, t) \) conditioned by the parabolic function \( h(x, t) = e^{\lambda t} \varphi(x) \) stays in \( \Lambda \times \mathbb{R} \) forever. The spatial component of this process can be considered to be “Brownian motion conditioned to stay in \( \Lambda \) forever” because it is the weak limit, as \( t \to \infty \), of Brownian motions conditioned not to exit \( \Lambda \) in \([0, t]\) (see [25]). We will use \( \tilde{P}^\mu \) to denote the distribution of Brownian motion conditioned to stay in \( \Lambda \) forever, with the initial distribution \( c\varphi(x)\mu(dx) \), where \( c > 0 \) is the normalizing constant.

4. **Weak convergence of spines.** Recall that we assume that the driving process \( B \) is Brownian motion and \( \Lambda \subset \mathbb{R}^d \) is an open bounded connected Lipschitz domain with the Lipschitz constant less than 1.

Our main result is as follows.

**Theorem 4.1.** Suppose that \( \mu \) is a probability measure supported in a set \( \Lambda_1 \subset \Lambda \) such that \( \text{dist}(\Lambda_1, \Lambda^c) > 0 \). Consider a sequence of Fleming-Viot processes \( \mathbf{X}^n \) in \( \Lambda \) driven by Brownian motion. Assume that the measures \( \mu_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_0^k} \) are supported in \( \Lambda_1 \) and converge weakly to \( \mu \) as \( n \to \infty \). Then the distributions of spines \( \{J^t_n, t \geq 0\} \) converge to \( \tilde{P}^\mu \).


Remark 4.2. (i) We believe that the theorem holds for all bounded Lipschitz domains, not only those with the Lipschitz constant less than 1. At present it is not known whether the lifetime $\tau_{\infty}$ is finite for the Fleming-Viot process driven by Brownian motion in any Lipschitz domain. However, it is implicit in arguments in [5] that for every Euclidean domain $\Lambda$ (not necessarily Lipschitz), $\tau_{\infty} \to \infty$ in distribution as the number $n$ of particles goes to infinity. This is enough to extend Theorem 4.1 to all Lipschitz domains. We omit the proof of this stronger result to avoid another layer of technicalities. We cannot get rid of the assumption of the Lipschitz character of the domain $\Lambda$ because it is an essential ingredient in Lemma 5.3.

(ii) The proof of Theorem 4.1 is based on a large number of estimates specific to Brownian motion. While the overall structure of the proof, outlined in Section 2, is very general and, therefore, could be applied to any driving Markov process, Brownian motion is the only process for which the estimates needed in the proof are readily available in the literature, to our best knowledge.

(iii) One can generate examples of driving processes for which the theorem holds by "re-labeling" the state space as follows. If $\Lambda'$ is a set and $F : \Lambda \to \Lambda'$ is a one-to-one function then \{F(X_t), t \geq 0\}, $n \geq 1$, are Fleming-Viot processes with the state space $\Lambda'$, driven by the image of Brownian motion by $F$. There are some technical details that need to be taken care of. We show how this can be done in some specific cases in Section 6.

(iv) We assumed that the measures $\frac{1}{n} \sum_{k=1}^{n} \delta_{X_k^0}$ are supported in $\Lambda_1 \subset \Lambda$ such that $\text{dist}(\Lambda_1, \Lambda^c) > 0$ because, to apply Theorem 3.2 in our argument, we need the following bound: for each fixed $t > 0$,

$$\limsup_{n \to \infty} \left( \mathbb{E} \left( \mathbb{P}_{\mu, n}(\tau_\Lambda > t)^{-2} \right) \right)^{1/2} < \infty.$$

It is easy to see that for every fixed $t > 0$, the function $x \mapsto \mathbb{P}(\tau_\Lambda > t \mid B_0 = x)$ is continuous and strictly positive inside $\Lambda$. Hence, $\inf_{x \in \Lambda} \mathbb{P}(\tau_\Lambda > t \mid B_0 = x) > 0$. This implies (4.1).

The following example shows that (4.1) fails for some natural initial distributions. Suppose that $\Lambda$ has a smooth boundary, $\mu$ is the uniform probability distribution in $\Lambda$ and $X_k^0$, $k = 1, \ldots, n$, are i.i.d. with the distribution $\mu$. By an argument similar to that in Lemma 5.3,

$$\mathbb{P}(\tau_\Lambda > t \mid B_0 = x) \leq c_1 \text{dist}(x, \partial \Lambda), \quad x \in \Lambda,$$

where $c_1$ depends on $\Lambda$ and $t$. It follows that

$$\mathbb{E} \left( \mathbb{P}_{\mu, n}(\tau_\Lambda > t)^{-2} \right) \geq \frac{1}{|\Lambda|} \int_{\Lambda} \mathbb{P}(\tau_\Lambda > t \mid B_0 = x)^{-2} dx \geq \frac{1}{|\Lambda|} \int_{\Lambda} c_1^{-2} \text{dist}(x, \partial \Lambda)^{-2} dx \geq c_2 \int_{0+} s^{-2} ds = \infty.$$

Therefore, (4.1) fails in this case.

(v) It is conceivable that for a fixed finite number of individuals, the spine of the Fleming-Viot process has the distribution of the process conditioned not to hit the boundary of the domain. Although this seems to be highly unlikely, proving that this is not the case does not seem to be easy. Apparently there are only two examples showing that this is not the case, one in [2, Sect. 6] and another one in [7]. They both deal with Fleming-Viot processes with only $n = 2$ individuals. In the first case, the process has a finite state space; in the second case, the driving process is Brownian motion in $[0, \infty)$.

(vi) We will use an informal notation for Radon-Nikodym derivatives. Here is a typical example: $\mathbb{P}(B_t^\prime \in dx) / \mathbb{P}(B_t^\prime \in dx)$. In this example $B^\prime$ and $B''$ are processes with values in $\mathbb{R}^d$, and $x \in \mathbb{R}^d$. The ratio of two "probabilities" represents the value of the Radon-Nikodym derivative of the distribution of $B_t^\prime$ on $\mathbb{R}^d$ with respect to the distribution of $B_s''$ on $\mathbb{R}^d$, evaluated at $x$. Most applications of this notation will be more complex than this simple example but the interpretation will remain the same.
**Proof of Theorem 4.1.** The proof is based on a large number of estimates that are relegated to Section 5.

It will suffice to show that for every fixed \( t_1 > 0 \), the distributions of \( \{ J_t^\mu, 0 \leq t \leq t_1 \} \) converge weakly to \( \bar{P}^\mu \) truncated in the obvious way to the interval \([0, t_1]\). To see this, note that convergence of finite dimensional distributions on \([0, \infty)\) is implied by the same type of convergence on all compact subintervals of \([0, \infty)\). It follows from Proposition 1.5 in Chapter XIII in [27] that tightness on \([0, \infty)\) is implied by tightness on compact time intervals. So we fix an arbitrary \( t_1 > 0 \).

Let \( 0 < \gamma < \infty \) be a constant satisfying the condition in Lemma 5.3. We will argue that there exist \( \alpha, \delta, \xi > 0 \) satisfying the following conditions,

\[
\begin{align*}
\alpha &\leq (1/2 - 2\delta + 3\gamma \delta/4)/ (\gamma + 2d), \\
0 &< \xi < 2\alpha - 3\delta/2, \\
\alpha &> \delta.
\end{align*}
\]

Note that \( d \) and \( \gamma \) are fixed at this point. Let \( \delta_0 > 0 \) be so small that \((1/2 - 2\delta + 3\gamma \delta/4)/(\gamma + 2d) \geq (1/4)/(\gamma + 2d)\) for \( 0 < \delta \leq \delta_0 \). Let \( \alpha_0 = (1/4)/(\gamma + 2d) \). If \( 0 < \delta \leq \delta_0 \) and \( 0 < \alpha \leq \alpha_0 \) then (4.2) holds. We let \( \alpha = \min(\alpha_0, \delta_0)/2 \) and \( \delta = \alpha/2 \) so that (4.4) holds true. Finally we note that with this choice of \( \alpha \) and \( \delta \), \( 2\alpha - 3\delta/2 > 0 \) so we can find \( \xi \) such that (4.3) is satisfied.

Fix an arbitrarily small \( \varepsilon > 0 \). Fix some \( u = u(t_1, \varepsilon) > t_1 \) which is greater than \( s_1 \) in Lemma 5.3, greater than \( s_1 \) in Lemma 5.5, and greater than \( s_1 \) in Lemma 5.7.

We will now refer to the notation introduced before and in Lemma 5.10. This includes \( C_j \), an atypical event, and \( k_1 \). Let \( n_1 \) be as in Lemma 5.10, relative to \( \varepsilon \) and \( u \) fixed above. From now on we will consider only \( n > n_1 \). Let \( t_2 = u + jn^{-2\alpha + \delta} \), where \( j \geq 0 \) is the smallest integer such that \( P(C_j) \leq \varepsilon \) in the notation of Lemma 5.10. Note that \( j \) and, therefore, \( t_2 \) depend on \( n \). In the notation of Lemma 5.10, \( \Delta t = n^{-2\alpha + \delta}, k_1 = [1/\Delta t] \) and \( j \leq k_1 \) so \( t_2 \in [u, u + 1] \).

Let \( t_2 = t_2 + n^{-2\alpha + \delta} \).

For \( z = (z_1, \ldots, z_d) \in \mathbb{R}^d \) let

\[
Q(z, r) = \left\{ (y_1, \ldots, y_d) \in \mathbb{R}^d : \max_{1 \leq k \leq d} |y_k - z_k| \leq r \right\}.
\]

Let \( Q \) be the family of all cubes \( \Omega(z, r, n^{-\alpha}) \) such that every \( z_k \) is an integer multiple of \( n^{-\alpha} \), and

\[
\text{dist}(\Omega, \mathbb{R}^d) \geq 3n^{-\alpha + 3\delta}/4.
\]

If \( 0 < s < t \) and \( \omega \in C[0, t] \) then \( \omega|_s \in C[0, s] \) will denote the truncation of \( \omega \) to the interval \([0, s]\).

Fix a continuous non-negative function \( f : C[0, \infty) \to \mathbb{R} \) with \( ||f||_{\infty} \leq 1 \), which depends only on the values of the process on \([0, t_1]\), i.e., if \( \omega', \omega'' \in C[0, \infty) \) and \( \omega'|_{t_1} = \omega''|_{t_1} \) then \( f(\omega') = f(\omega'') \). In the same spirit, we can apply \( f \) to \( \omega \in C[0, s] \) for any \( s \geq t_1 \). Assume that \( \mathbb{E} t_1(f) > 0 \). By Lemma 5.8, \( \mathbb{E} t_2(f) > 0 \).

For \( Q \in \mathcal{Q} \), let \( \omega_{t_2} \in \mathcal{Q} \) and \( f_{Q} = f \mathbf{1}_{G_{\mathcal{Q}}} \). Recall definition (3.4) and let

\[
\begin{align*}
A_1 &= \bigcap_{Q \in \mathcal{Q}} \left\{ \left| \mathcal{H}^\mu_{t_2} (1_{G_{\mathcal{Q}}} - \mathbb{E} t_2(G_{\mathcal{Q}})) \right| \leq n^{-\alpha d + \gamma(-\alpha + 3\delta)/4} \right\}, \\
A_2 &= \bigcap_{Q \in \mathcal{Q}} \left\{ \left| \mathcal{H}^\mu_{t_2} (f_Q) - \mathbb{E} t_2(f_Q) \right| \leq n^{-\alpha d + \gamma(-\alpha + 3\delta)/4} \right\}.
\end{align*}
\]
For a cube $Q \in \mathcal{Q}$, let $\mathcal{M}_Q = \{ j : X^j_{t_2} \in Q \}$.

Let $|\mathcal{M}_Q|$ denote the cardinality of $\mathcal{M}_Q$. We will use the notation $\mathcal{M}_Q = \{i_1, \ldots, i_N\}$. Note that $\Lambda = \Lambda(Q) = |\mathcal{M}_Q|$ is a random integer. By abuse of notation, $|\cdot|$ will denote the Euclidean norm in (4.9) below and later in the paper. Recall (3.3) and let

\begin{equation}
A_3 = \bigcap_{1 \leq i \leq n} \left\{ \sup_{s,t \in [t_2,t_3]} |B^i_s - B^i_t| < 2n^{-\alpha +3\delta/4} \right\},
\end{equation}

\begin{equation}
A_4 = A_1 \cap A_2 \cap A_3.
\end{equation}

Let $\mathcal{F}_t = \sigma \{ X^j_{s}, 0 \leq s \leq t \}$. Let $\pi$ be a random (i.e., uniform) permutation of $\{1, \ldots, n\}$, independent of $X^n$ and

$$
\mathcal{F}_t^+ = \sigma \left\{ \left( X^{\pi(1)}_s, X^{\pi(2)}_s, \ldots, X^{\pi(n)}_s \right), s \geq t \right\}.
$$

Let $\mathcal{G}_1$ be the smallest $\sigma$-field generated by $\mathcal{F}_{t_2}^+, \mathcal{F}_{t_3}^+$, and random sets $\bigcup_{j \in \mathcal{M}_Q} \{X^j_{t_2}\}$ for all $Q \in \mathcal{Q}$. The information in the $\sigma$-field $\mathcal{G}_1$ includes the information on the locations of all particles at all times $t \geq t_3$, but with labels of the particles missing (scrambled). For every $Q \in \mathcal{Q}$, $\mathcal{G}_1$ contains the information on the locations of $X^j_{t_2}$ for $j \in \mathcal{M}_Q$, but once again with the information on the labels missing. As a consequence, $\mathcal{G}_1$ contains the information about the location of the spine at $t_3$, and, for every $Q \in \mathcal{Q}$, on whether the spine is equal to $X^j_{t_3}$ for some $j \in \mathcal{M}_Q$ but without specifying $j$.

In view of (4.6), if $j \in \mathcal{M}_Q$, $Q \in \mathcal{Q}$ and $A_3$ holds then $X^j$ does not hit $\partial \Lambda$ in the time interval $[t_2, t_3]$ and, therefore, it does not jump during this interval. Thus, the joint conditional distribution of $\{X^j_i - X^j_{t_2}, t_2 \leq t \leq t_3\}$, $i \in \mathcal{M}_Q$, given $A_4$ and $\mathcal{F}_{t_2}$ is that of independent Brownian motions $\{B^i_t - B^i_{t_2}, t_2 \leq t \leq t_3\}$, $i \in \mathcal{M}_Q$, conditioned by

$$
\bigcap_{i \in \mathcal{M}_Q} \left\{ \sup_{s,t \in [t_2,t_3]} |B^i_s - B^i_t| < 2n^{-\alpha +3\delta/4} \right\}.
$$

Let

\[ \Lambda' = \left\{ (x^1, x^2, z^1, z^2) \in \Lambda^4 : x^1, x^2 \in Q, z^1, z^2 \in \Lambda, |x^m - z^\ell| \leq n^{-\alpha +3\delta/4} \text{ for } m, \ell = 1, 2 \right\}. \]

We will explain how Lemma 5.1 implies that

\begin{equation}
\mathbb{P} \left( \{X^j_i \in dz^1, X^k_i \in dz^2, X^j_i \in dx^1, X^k_i \in dx^2 \} \cap A_3 \mid \mathcal{F}_{t_2} \right) \leq \exp \left( 4\sqrt{dn}^{-\delta/4} \right),
\end{equation}

for all $(x^1, x^2, z^1, z^2) \in \Lambda'$, $Q \in \mathcal{Q}$, $j,k \in \mathcal{M}_Q$, and sufficiently large $n$. Processes $X^j$ and $X^k$ play the same role as $B'$ and $B''$ in Lemma 5.1. The events $F(\cdot)$ that appeared in Lemma 5.1 are replaced (implicitly) in (4.11) by $A_3$ because $A_3 \subset A_4$. Compared to Lemma 5.1, there is extra conditioning in (4.11), specifically on the event $A_4$ which is the event $A_1 \cap A_2 \subset A_4$. The extra conditions are independent of the trajectories of $X^j$ and $X^k$ on the interval $[t_2, t_3]$ given $\mathcal{F}_{t_2}$.

Let

\begin{equation}
\Lambda'' = \Lambda''(Q) = \{(z^1, \ldots, z^N) \in \Lambda^N : |x - z^j| \leq n^{-\alpha +3\delta/4} \text{ for } j = 1, \ldots, N, x \in Q \}.
\end{equation}
Consider \( x^1, \ldots, x^N \in Q \) and \( (z^1, \ldots, z^N) \in \Lambda'' \). We will argue that for any \( j_1, j_2 \in \mathcal{M}_Q \), and sufficiently large \( n \),

\[
\mathbb{P} \left( \left\{ X^j_{t_3} \in dx^j, j \in \mathcal{M}_Q \right\} \cap \left\{ X^{j_1}_{t_3} \in dz^{j_1} \right\} \cap \left\{ X^{j_2}_{t_3} \in dz^{j_2} \right\} \right) \cap \mathcal{A}_4 | \mathcal{F}_{t_2} 
\times \mathbb{P} \left( \left\{ X^j_{t_2} \in dx^j, j \in \mathcal{M}_Q \right\} \cap \left\{ X^{j_1}_{t_2} \in dz^{j_1} \right\} \cap \left\{ X^{j_2}_{t_2} \in dz^{j_2} \right\} \right) \right) \left( 4 \sqrt{dn^{-\delta/4}} \right) .
\]

(4.13)

The last formula is a more elaborate version of (4.11). Indices \( j_1 \) and \( j_2 \) in (4.13) play the same role as \( j \) and \( k \) in (4.11). Processes \( X^j \) with \( j \in \mathcal{M}_Q \setminus \{j_1, j_2\} \) are also included in (4.13). This does not affect the validity of the formula because the extra conditioning related to these processes is the same in the numerator and denominator (recall that Brownian motions \( B^i, 1 \leq i \leq n \), are jointly independent).

Consider once again \( x^1, \ldots, x^N \in Q \) and \( (z^1, \ldots, z^N) \in \Lambda'' \). Suppose that

\[
\mathcal{M}_Q = \{j_1, \ldots, j_N\},
\]

\[
\{x^1, x^2, \ldots, x^N\} = \{X^{j_1}_{t_3}, X^{j_2}_{t_3}, \ldots, X^{j_N}_{t_3}\},
\]

\[
\{z^1, z^2, \ldots, z^N\} = \{X^{j_1}_{t_2}, X^{j_2}_{t_2}, \ldots, X^{j_N}_{t_2}\}.
\]

For \( 1 \leq i \leq N \), let \( \Theta(i) = r \) if there is \( j_k \) such that \( X^{j_k}_{t_3} = z^i \) and \( X^{j_k}_{t_2} = x^r \). Note that \( \Theta \) is a permutation of \( \{1, \ldots, N\} \).

Consider \( 1 \leq i, r_1, r_2 \leq N \). Let \( \Pi_{t_2} \) be the set of permutations \( \sigma \) of \( \{1, \ldots, N\} \) such that \( \sigma(i) = r_t \), for \( t = 1, 2 \). If \( \sigma \in \Pi_1 \) and \( \sigma(i_1) = r_2 \) then we let \( R(\sigma) \) be a permutation such that \( R(\sigma)(i) = r_2, R(\sigma)(i_1) = r_1 \), and \( R(\sigma)(k) = \sigma(k) \) for \( k \neq i, i_1 \). It is easy to see that \( R : \Pi_1 \to \Pi_2 \) is a bijection. It follows from (4.13) that for every \( \sigma \in \Pi_1 \),

\[
\mathbb{P} \left( \left\{ \Theta = \sigma \right\} \cap \mathcal{A}_4 | \mathcal{G}_1 \right) \leq \exp \left( 4 \sqrt{dn^{-\delta/4}} \right).
\]

The conditioning \( \sigma \)-field has changed from \( \mathcal{F}_{t_2} \) to \( \mathcal{G}_1 \), relative to (4.13). The latter \( \sigma \)-field contains information about the set of locations of \( X^i \)'s at time \( t_3 \) for \( i \in \mathcal{M}_Q \), without information about which location corresponds to which process \( X^i \). Hence, the last formula compares probabilities of different “permutations” or assignments of \( X^i \)'s to different locations.

Summing over all \( \sigma \in \Pi_1 \) in the last formula yields for every \( i = 1, \ldots, N \) and \( r_1, r_2 = 1, \ldots, N \),

\[
\mathbb{P} \left( \left\{ \Theta(i) = r_1 \right\} \cap \mathcal{A}_4 | \mathcal{G}_1 \right) \leq \exp \left( 4 \sqrt{dn^{-\delta/4}} \right).
\]

(4.14)

Recall (4.12) and let

\[
A^Q_3(z^1, \ldots, z^N) = \left\{ \bigcup_{j \in \mathcal{M}_Q} \{X^j_{t_3}\} = \{z^1, \ldots, z^N\} \right\},
\]

(4.15)

\[
A^Q_3 = \bigcup_{(z^1, \ldots, z^N) \in \Lambda''(Q)} A^Q_3(z^1, \ldots, z^N),
\]

(4.16)
(4.17) \[ A_0^Q = A_4 \cap A_5^Q, \]

equand note that the event \( A_5^Q(z^1, \ldots, z^N) \) concerns the equality of two unordered sets. The event \( A_5^Q \) is the intersection of \( A_4 \) and the event that processes \( X^j \) for \( j \in \mathcal{M}_Q \) stay within a distance \( n^{-\alpha+3\delta/4} \) from all points of \( Q \) at time \( t_3 \).

If \( A_5^Q(z^1, \ldots, z^N) \) holds then for \( 1 \leq m \leq N \), let \( k(m) \) be such that \( X^{k(m)}_{t_3} = z^m \). It follows from (4.14) that for any \( 1 \leq m_1, m_2 \leq N \) and \( j_1, j_2 \in \mathcal{M}_Q \), and sufficiently large \( n \),

\[
\begin{align*}
\mathbb{P}\left( \left\{ H^{k(m_1)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_4 \cap A_5^Q(z^1, \ldots, z^N) \mid \mathcal{G}_1 \right) & \leq \exp \left( 4\sqrt{d} n^{-\delta/4} \right), \\
\mathbb{P}\left( \left\{ H^{k(m_2)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_4 \cap A_5^Q(z^1, \ldots, z^N) \mid \mathcal{G}_1 \right) & \leq \exp \left( 4\sqrt{d} n^{-\delta/4} \right), \\
\int_{A''} \mathbb{P}\left( \left\{ H^{k(m_1)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_4 \cap A_5^Q(z^1, \ldots, z^N) \mid \mathcal{G}_1 \right) dz^1 \ldots dz^N & \geq \exp \left( -4\sqrt{d} n^{-\delta/4} \right)
\end{align*}
\]

The last inequality implies that

\[
|\mathcal{M}_Q| \mathbb{P}\left( \left\{ H^{k(m_2)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_6^Q \mid \mathcal{G}_1 \right) = \sum_{1 \leq m_1 \leq N} \mathbb{P}\left( \left\{ H^{k(m_1)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_6^Q \mid \mathcal{G}_1 \right)
\]

\[
\geq \exp \left( -4\sqrt{d} n^{-\delta/4} \right) \sum_{1 \leq m_1 \leq N} \mathbb{P}\left( \left\{ H^{k(m_1)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_6^Q \mid \mathcal{G}_1 \right)
\]

\[
= \exp \left( -4\sqrt{d} n^{-\delta/4} \right) \mathbb{P}\left( A_6^Q \mid \mathcal{G}_1 \right) = \exp \left( -4\sqrt{d} n^{-\delta/4} \right) 1_{A_6^Q}.
\]

We have shown that for all \( 1 \leq m \leq N \), \( j \in \mathcal{M}_Q \) and sufficiently large \( n \),

\[
\mathbb{P}\left( \left\{ H^{k(m)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_6^Q \mid \mathcal{G}_1 \right) \geq \frac{1}{|\mathcal{M}_Q|} \exp \left( -4\sqrt{d} n^{-\delta/4} \right) 1_{A_6^Q}.
\]

Suppose that \( R \) is an \( \mathcal{F}^{+}_{t_3} \)-measurable random variable taking values in \( \{ X^1_{t_3}, \ldots, X^n_{t_3} \} \) and let

\[
A_7^R = \bigcup_{j \in \mathcal{M}_Q} \{ R = X^j_{t_3} \}.
\]

By abuse of notation, if \( R = X^j_{t_3} = z^m \), we will write \( H^R_{t_3} \) or \( H^m_{t_3} \) instead of \( H^{k(m)}_{t_3} \). Note that \( R \) and \( A_7^R \) are \( \mathcal{G}_1 \)-measurable so (4.19) implies that

\[
\mathbb{P}\left( \left\{ H^{k(m)}_{t_3}(t_2) = X^j_{t_2} \right\} \cap A_6^Q \cap A_7^R \mid \mathcal{G}_1 \right) \geq \frac{1}{|\mathcal{M}_Q|} \exp \left( -4\sqrt{d} n^{-\delta/4} \right) 1_{A_6^Q \cap A_7^R}.
\]
Recall that \( f : C([0, \infty)) \rightarrow [0, 1] \) is a continuous function and it only depends on the values of the process on \([0, t_1]\). We will explain how the last formula and the definition (3.4) imply that

\[
(4.21) \quad \mathbb{E} \left( \mathcal{H}_t^n(1_{G_Q}) f \left( H^R_{t_z} \right) 1_{A^Q_i \cap A^Q_j} \mid G_1 \right)
\]

\[
= \sum_{j \in M_Q} \mathbb{E} \left( \frac{|M_Q|}{n} f \left( H^j_{t_z} \right) 1_{A^Q_j \cap A^Q_i} \mid G_1 \right) \mathbb{P} \left( \{ H^R_{t_z} = X^j_{t_z} \} \cap A^Q_j \cap A^Q_i \mid G_1 \right)
\]

\[
\geq \frac{|M_Q|}{n} \sum_{j \in M_Q} \mathbb{E} \left( f \left( H^j_{t_z} \right) 1_{A^Q_j \cap A^Q_i} \mid G_1 \right) \frac{1}{|M_Q|} \exp \left( -4 \sqrt{dn}^{-\delta/4} \right) 1_{A^Q_i \cap A^Q_j}
\]

\[
= \frac{1}{n} \sum_{j \in M_Q} \mathbb{E} \left( f \left( H^j_{t_z} \right) 1_{A^Q_j \cap A^Q_i} \mid G_1 \right) \exp \left( -4 \sqrt{dn}^{-\delta/4} \right)
\]

There is no indicator \( 1_{A^Q_i \cap A^Q_j} \) at the end of the last line because it is implicitly included in the conditional expectation, since \( A^Q_i \cap A^Q_j \) is \( G_1 \)-measurable. The first equality follows from the “total probability formula,” i.e., we sum over all events \( \{ H^R_{t_z} = X^j_{t_z} \} \cap A^Q_j \cap A^Q_i \) because they form a partition of \( A^Q_i \cap A^Q_j \). Since \( f \) depends only on the part of the trajectory in \([0, t_1]\), the paths of \( H^R_{t_z} \) and \( H^j_{t_z} \) agree on \([0, t_1]\) if \( \{ H^R_{t_z} = X^j_{t_z} \} \cap A^Q_j \cap A^Q_i \) holds. The quotient \( |M_Q|/n \) is an alternative way of writing \( \mathcal{H}_t^n(1_{G_Q}) \). The sum can be moved under the expectation in the last step because the events \( \{ j \in M_Q \} \) belong to \( G_1 \).

One of the major steps in the proof will be a derivation of a bound for \( \mathbb{E} \left( f \left( H^R_{t_z} \right) 1_{A^Q_i \cap A^Q_j} \right) \).

Recall the definition (3.4). The inequality below holds because \( A^Q_i \cap A^Q_j \subset A_2 \) and we have (4.8).

\[
(4.22) \quad \left( \mathbb{E}_{t_z} (f_Q) - n^{-\alpha+\gamma(-\alpha+3\delta/4)-\delta} \right) \mathbb{P}(A^Q_i \cap A^Q_j)
\]

\[
= \mathbb{E} \left( \left( \mathbb{E}_{t_z} (f_Q) - n^{-\alpha+\gamma(-\alpha+3\delta/4)-\delta} \right) 1_{A^Q_i \cap A^Q_j} \right)
\]

\[
\leq \mathbb{E} \left( \mathcal{H}_t^n(f_Q) 1_{A^Q_i \cap A^Q_j} \right) = \frac{1}{n} \mathbb{E} \left( \sum_{i \in M_Q} f \left( H^i_{t_z} \right) 1_{A^Q_i \cap A^Q_j} \right).
\]

If follows from the definition of \( M_Q \) and the fact that \( f \) depends only on the values of the process in \([0, t_1]\) that

\[
(4.23) \quad \sum_{i \in M_Q} f \left( H^i_{t_z} \right) 1_{A^Q_i \cap A^Q_j} = \sum_{i \in M_Q} f \left( H^i_{t_z} \right) 1_{A^Q_i \cap A^Q_j}.
\]

Recall that \( R \) and, therefore, \( 1_{A^Q_i} \) are \( G_1 \)-measurable. In the following calculation, we use (4.23) in the first equality. The first inequality follows from (4.21). The second inequality follows from the fact that \( A^Q_i \subset A_1 \) and (4.7).

\[
\frac{1}{n} \mathbb{E} \left( \sum_{i \in M_Q} f \left( H^i_{t_z} \right) 1_{A^Q_i \cap A^Q_j} \right) = \frac{1}{n} \mathbb{E} \left( \sum_{i \in M_Q} f \left( H^i_{t_z} \right) 1_{A^Q_i \cap A^Q_j} \right)
\]
By Lemma 5.6,

\[
\tilde{p}_{t_2}(G_Q) \geq c_1 n^{-d+\gamma(-\alpha+3\delta/4)}, \quad \tilde{E}_{t_2}(f_Q) \geq c_2 n^{-d+\gamma(-\alpha+3\delta/4)},
\]

where \(c_2 = c_3 \tilde{E}_{t_2}(f)\). Thus (4.24) yields for large \(n\),

\[
\mathbb{E}\left( f\left(H_{t_3}^R\right) 1_{A_0^Q \cap A_\gamma^Q} \right) \geq \frac{\tilde{E}_{t_2}(f_Q) (1 - n^{-\delta/4}) \mathbb{P}(A_0^Q \cap A_\gamma^Q)}{\tilde{p}_{t_2}(G_Q) (1 + n^{-\delta/4}) \exp\left(4\sqrt{d} n^{-\delta/4}\right)}.
\]  

One can show in a similar manner that

\[
\mathbb{E}\left( f\left(H_{t_3}^R\right) 1_{A_0^Q \cap A_\gamma^Q} \right) \leq \tilde{E}_{t_2}(f \mid B_{t_2} \in Q) (1 - n^{-\delta/8}) \mathbb{P}(A_0^Q \cap A_\gamma^Q).
\]  

By Lemma 5.5 and the assumption that \(\|f\|_{\infty} \leq 1\), for every \(Q \in \mathcal{Q}\),

\[
1 - \varepsilon \leq \frac{\tilde{E}_{t_2}(f \mid B_{t_2} \in Q)}{\tilde{E}_{t_2}(f)} \leq 1 + \varepsilon,
\]

\[
\tilde{E}_{t_2}(f) - \varepsilon \leq \tilde{E}_{t_2}(f \mid B_{t_2} \in Q) \leq \tilde{E}_{t_2}(f) + \varepsilon.
\]

We use this estimate, (4.25) and (4.26) to see that

\[
\mathbb{E}\left( f\left(H_{t_3}^R\right) 1_{A_0^Q \cap A_\gamma^Q} \right) \geq \left(\tilde{E}_{t_2}(f) - \varepsilon\right) (1 - n^{-\delta/8}) \mathbb{P}(A_0^Q \cap A_\gamma^Q),
\]

\[
\mathbb{E}\left( f\left(H_{t_3}^R\right) 1_{A_0^Q \cap A_\gamma^Q} \right) \leq \left(\tilde{E}_{t_2}(f) + \varepsilon\right) (1 + n^{-\delta/8}) \mathbb{P}(A_0^Q \cap A_\gamma^Q).
\]
Recall that $R$ is an $\mathcal{F}_t^n$-measurable random variable taking values in $\{X^{1}_{t_3}, \ldots, X^{n}_{t_3}\}$ and $A^Q_7 = \bigcup_{j \in M_6} \{R = X^{j}_{t_3}\}$. Let $A_8 = \bigcup_{Q \in Q} (A^Q_6 \cap A^Q_7)$. It follows from (4.27) that
\[
\mathbb{E}(f(H^R_t) 1_{A_8}) \geq \left(\mathbb{E}_t^2(f) - \varepsilon\right)(1 - n^{-\delta/8}) \mathbb{P}(A_8).
\]
This and $\|f\|_\infty \leq 1$ imply that,
\[
\mathbb{E}(f(H^R_t)) \geq \mathbb{E}(f(H^R_t) 1_{A_8}) \geq \mathbb{E}_t^2(f)(1 - n^{-\delta/8}) \mathbb{P}(A_8) - \varepsilon
= \mathbb{E}_t^2(f) \mathbb{P}(A_8) - \mathbb{E}_t^2(f)n^{-\delta/8} \mathbb{P}(A_8) - \varepsilon
\geq \mathbb{E}_t^2(f) - \mathbb{E}_t^2(f)(1 - \mathbb{P}(A_8)) - n^{-\delta/8} - \varepsilon
\geq \mathbb{E}_t^2(f) - \mathbb{P}(A^c_8) - n^{-\delta/8} - \varepsilon.
\]

For the following upper bound, we use (4.28) and the fact that $\|f\|_\infty \leq 1$,
\[
\mathbb{E}(f(H^R_t) 1_{A_8}) \leq \left(\mathbb{E}_t^2(f) + \varepsilon\right)(1 + n^{-\delta/8}) \mathbb{P}(A_8).
\]
This and $\|f\|_\infty \leq 1$ imply that
\[
\mathbb{E}(f(H^R_t)) \leq \mathbb{E}(f(H^R_t) 1_{A_8}) + \mathbb{P}(A^c_8)
\leq \mathbb{E}_t^2(f)(1 + n^{-\delta/8}) \mathbb{P}(A_8) + \mathbb{P}(A^c_8) + 2\varepsilon
= \mathbb{E}_t^2(f) \mathbb{P}(A_8) + \mathbb{E}_t^2(f)n^{-\delta/8} \mathbb{P}(A_8) + \mathbb{P}(A^c_8) + 2\varepsilon
\leq \mathbb{E}_t^2(f) + \mathbb{P}(A^c_8) + n^{-\delta/8} + 2\varepsilon.
\]

The last two estimates show that
\[
(4.29) \quad \left| \mathbb{E}(f(H^R_t)) - \mathbb{E}_t^2(f) \right| \leq \mathbb{P}(A^c_8) + n^{-\delta/8} + 2\varepsilon.
\]

Recall the definition of $\chi$ following the statement of Theorem 3.1 and note that $H^\chi(t_3)(t) = J^n(t)$ for $t \in [0, t_3]$. Hence we can apply (4.29) to $R = \chi(t_3)$ to obtain,
\[
(4.30) \quad \left| \mathbb{E}(f(J^n)) - \mathbb{E}_t^2(f) \right| \leq \mathbb{P}(A^c_8) + n^{-\delta/8} + 2\varepsilon.
\]

Recall definitions (4.10) of $A_4$, (4.17) of $A^Q_6$ and (4.20) of $A^Q_7$. We have
\[
A_8 = \bigcup_{Q \in Q} (A^Q_6 \cap A^Q_7) = \bigcup_{Q \in Q} (A_1 \cap A^Q_3 \cap A^Q_7) = A_1 \cap A_2 \cap A_3 \cap \bigcup_{Q \in Q} (A^Q_6 \cap A^Q_7),
\]
so
\[
(4.31) \quad A^c_8 \subset A^c_1 \cup A^c_2 \cup A^c_3 \cup \left(\bigcap_{Q \in Q} A^Q_3\right)^c \cup \left(A_1 \cap A_2 \cap A_3 \cap \bigcap_{Q \in Q} A^Q_3 \cap \left(\bigcup_{Q \in Q} A^Q_7\right)^c\right).
\]

In words, for the event $A_8$ to fail, one of the following must be true:

(i) at least one of the following events fails: $A_1$, $A_2$, $A_3$, $\bigcap_{Q \in Q} A^Q_3$, or

(ii) the event $A_1 \cap A_2 \cap A_3 \cap \bigcap_{Q \in Q} A^Q_3$ holds and $\bigcup_{Q \in Q} A^Q_7$ fails.

By Lemmas 5.2 and 5.9,
\[
(4.32) \quad \lim_{n \to \infty} \mathbb{P}\left(A_1 \cap A_2 \cap A_3 \cap \bigcap_{Q \in Q} A^Q_3\right) = 1.
\]
Suppose the event in (4.32) holds but $\bigcup_{Q \in \mathcal{Q}} A^Q_7$ fails. Then $J^n_{t_3} = X^n_t$ for some $j \notin \bigcup_{Q \in \mathcal{Q}} \mathcal{M}_Q$. This implies, in view of (4.5) and (4.6), that $\text{dist}(X^n_t, \Lambda^n) \leq 4n^{-\alpha+3\delta/4}$ for large $n$. If $X^j$ did not jump in the interval $[t_2, t_3]$ then $J^n_{t_2} = X^n_t$ and $\text{dist}(J^n_{t_2}, \Lambda^n) \leq 4n^{-\alpha+3\delta/4}$. The last event is called $C^n_j$ in (5.33), with $s_j$ playing the role of $t_2$. If $X^j$ jumps in the interval $[t_2, t_3]$ then the spine $J^n_j$ passes through a branch point in $[t_2, t_3]$. This event is called $C^n_j$ in (5.34), with $[s_j, s_{j+1}]$ playing the role of $[t_2, t_3]$. It follows from (5.33)-(5.34) and Lemma 5.10 that

$$\limsup_{n \to \infty} \mathbb{P} \left( \bigcup_{Q \in \mathcal{Q}} A_1 \cap A_2 \cap A_3 \cap \bigcap_{Q \in \mathcal{Q}} \bigg( \bigcup_{Q \in \mathcal{Q}} A^Q_7 \cap \left( \bigcup_{Q \in \mathcal{Q}} A^Q_7 \right) \bigg) \right) \leq \varepsilon.$$ 

This, (4.30), (4.31) and (4.32) imply that

$$\limsup_{n \to \infty} \left| \mathbb{E} \left( f(J^n) \right) - \mathbb{E}_{t_2}(f) \right| \leq 3 \varepsilon.$$

By Lemma 5.7 and $\|f\|_{\infty} \leq 1$,

$$\limsup_{n \to \infty} \left| \mathbb{E} \left( f(J^n) \right) - \mathbb{E}_t^{\mu}(f) \right| \leq 4 \varepsilon.$$

Since $\varepsilon > 0$ is arbitrarily small,

$$\lim_{n \to \infty} \mathbb{E} \left( f(J^n) \right) = \mathbb{E}_t^{\mu}(f).$$

The function $f$ is an arbitrary continuous non-negative function $f : C[0, \infty) \to \mathbb{R}$ with $\|f\|_{\infty} \leq 1$ and $\mathbb{E}_{t_1}(f) > 0$, which depends only on the values of the process on $[0, t_1]$, and $t_1 > 0$ is also arbitrary, so the theorem follows.

5. Estimates. We will use notation presented in Section 3 and in the proof of Theorem 4.1.

5.1. Large deviations and likelihood ratio.

Lemma 5.1. Suppose that $\alpha, \delta, t_2$ and $t_3$ are as in (4.2)-(4.4) and the paragraphs following these conditions. For a process $B$ with values in $\mathbb{R}^d$, let

$$F(B) = \left\{ \sup_{s,t \in [t_2,t_3]} |B_t - B_s| < 2n^{-\alpha+3\delta/4} \right\}.$$

Suppose that $B'$ and $B''$ are independent Brownian motions in $\mathbb{R}^d$. If $x^1, x^2 \in Q \in \mathcal{Q}$ and $|x^j - z^k| \leq n^{-\alpha+3\delta/4}$ for $j, k = 1, 2$ then

$$\mathbb{P} \left( \{ B_{t_3}^j \in dz^2, B_{t_2}^j \in dz^1, B_{t_2}^j \in dx^1, B_{t_2}^j \in dx^2 \} \cap F(B') \cap F(B'') \right) \leq \exp \left( 4 \sqrt{d} n^{-\delta/4} \right).$$

Proof. Step 1. We will prove that if $B$ is Brownian motion in $\mathbb{R}^d$ then there exists $c_1$ such that for any $x$ and $z$ such that $|x - z| \leq n^{-\alpha+3\delta/4}$ and sufficiently large $n$, the following bounds hold for the Radon-Nikodym derivative,

$$1 - \exp \left( -c_1 n^{\delta/2} \right) \leq \frac{\mathbb{P}(B_{t_3}^j \in dz \mid F(B), B_{t_2} = x)}{\mathbb{P}(B_{t_3}^j \in dz \mid B_{t_2} = x)} \leq 1 + \exp \left( -c_1 n^{\delta/2} \right).$$
In the first step, we will write $F$ instead of $F(B)$. Standard estimates show that there exists $c_2$ such that for all $n \geq 2$,
\[
P(F^c) \leq \mathbb{P} \left( \sup_{s,t \in [t_2,t_3]} |B_t - B_s| \geq 2n^{-\alpha + 3\delta/4} \right) \leq \exp \left( -c_2n^{2(-\alpha + 3\delta/4)}/(t_3 - t_2) \right)
\]
\begin{equation}
(5.2) \quad \leq \exp \left( -c_2n^{2(-\alpha + 3\delta/4)/n - 2\alpha + \delta} \right) = \exp \left( -c_2n^{\delta/2} \right).
\end{equation}
Recall that $|x - z| \leq n^{-\alpha + 3\delta/4}$. Let $B(v, r)$ denote a ball with center $v$ and radius $r$. If $B_{t_2} = x$ then dist $(z, \partial B(B_{t_2}, 2n^{-\alpha + 3\delta/4})) \geq n^{-\alpha + 3\delta/4}$. Let $\tau$ be the hitting time of the sphere $S := \partial B(B_{t_2}, 2n^{-\alpha + 3\delta/4})$. The last estimate, (5.2) and an application of the strong Markov property at time $\tau$ imply that
\begin{equation}
(5.3) \quad \mathbb{P}(B_{t_3} \in dz, F^c | B_{t_2} = x) \leq \mathbb{P}(\tau < t_3 | B_{t_2} = x) \sup_{y \in S, s \in [0, t_3 - t_2]} \mathbb{P}(B_s \in dz | B_0 = y)
\end{equation}
\[
\leq \mathbb{P}(F^c) \sup_{y \in S, s \in [0, t_3 - t_2]} \mathbb{P}(B_s \in dz | B_0 = y) \leq \exp \left( -c_2n^{\delta/2} \right) \mathbb{P}(B_{t_3} \in dz | B_{t_2} = x).
\]
We have
\begin{equation}
(5.4) \quad \frac{\mathbb{P}(B_{t_3} \in dz | F, B_{t_2} = x)}{\mathbb{P}(B_{t_3} \in dz | B_{t_2} = x)} = \frac{\mathbb{P}(B_{t_3} \in dz, F | B_{t_2} = x)}{\mathbb{P}(B_{t_3} \in dz | B_{t_2} = x) \mathbb{P}(F | B_{t_2} = x)}
\end{equation}
\[
= \frac{\mathbb{P}(B_{t_3} \in dz | B_{t_2} = x) - \mathbb{P}(B_{t_3} \in dz, F^c | B_{t_2} = x)}{\mathbb{P}(F | B_{t_2} = x)}
\end{equation}
\[
= \frac{1}{\mathbb{P}(F | B_{t_2} = x)} \left( 1 - \frac{\mathbb{P}(B_{t_3} \in dz, F^c | B_{t_2} = x)}{\mathbb{P}(B_{t_3} \in dz | B_{t_2} = x)} \right).
\]
This and (5.3) yield
\begin{equation}
(5.5) \quad \frac{\mathbb{P}(B_{t_3} \in dz | F, B_{t_2} = x)}{\mathbb{P}(B_{t_3} \in dz | B_{t_2} = x)} \geq 1 - \frac{\mathbb{P}(B_{t_3} \in dz, F^c | B_{t_2} = x)}{\mathbb{P}(B_{t_3} \in dz | B_{t_2} = x)} \geq 1 - \exp \left( -c_2n^{\delta/2} \right).
\end{equation}
It follows from (5.4) and (5.2), along with translation invariance, that for some $c_3$ and sufficiently large $n$,
\[
\frac{\mathbb{P}(B_{t_3} \in dz | F, B_{t_2} = x)}{\mathbb{P}(B_{t_3} \in dz | B_{t_2} = x)} \leq \frac{1}{\mathbb{P}(F | B_{t_2} = x)} = \frac{1}{\mathbb{P}(F)} = \frac{1}{1 - \mathbb{P}(F^c)}
\end{equation}
\[
\leq \frac{1}{1 - \exp \left( -c_2n^{\delta/2} \right)} \leq 1 + \exp \left( -c_3n^{\delta/2} \right).
\]
This estimate and (5.5) imply (5.1).

**Step 2.** We will prove that if $B'$ and $B''$ are independent Brownian motions in $\mathbb{R}^d$, $x^1, x^2 \in Q \subset \mathbb{Q}$ and $|x^1 - z^k| \leq n^{-\alpha + 3\delta/4}$ for $j, k = 1, 2$ then
\begin{equation}
(5.6) \quad \frac{\mathbb{P}(B'_t \in dz^2, B'_t \in dx^1)}{\mathbb{P}(B'_t \in dz^1, B'_t \in dx^1)} \cdot \frac{\mathbb{P}(B''_t \in dz^2, B''_t \in dx^2)}{\mathbb{P}(B''_t \in dz^2, B''_t \in dx^2)} \leq \exp \left( 2\sqrt{dn}^{\delta/4} \right).
\end{equation}
Suppose that $x^1, x^2 \in Q, Q$. Then $|x^1 - x^2| \leq \sqrt{dn}^{-\alpha}$. If in addition $|x^j - z^k| \leq n^{-\alpha + 3\delta/4}$ for $j, k = 1, 2$ then
\[
\frac{\mathbb{P}(B'_t \in dz^2, B'_t \in dx^1)}{\mathbb{P}(B'_t \in dz^1, B'_t \in dx^1)} \cdot \frac{\mathbb{P}(B''_t \in dz^2, B''_t \in dx^2)}{\mathbb{P}(B''_t \in dz^2, B''_t \in dx^2)}
\]
\[
\mathbb{P}(B_s' \in dz, B_{s+}'' \in dz | B_{s}'' = x^1, B_{s}'' = x^2) \\
= \mathbb{P}(B_s' \in dz, B_{s+}'' \in dz | B_{s}'' = x^1, B_{s}'' = x^2) (z^1, z^2) \\
= \frac{\exp \left( -|x^2 - z^1|^2/(2(t_3 - t_2)) \right)}{\exp \left( -|x^1 - z^1|^2/(2(t_3 - t_2)) \right)} \cdot \frac{\exp \left( -|x^1 - z^2|^2/(2(t_3 - t_2)) \right)}{\exp \left( -|x^2 - z^2|^2/(2(t_3 - t_2)) \right)} \\
= \exp \left( -|x^2 - x^1|^2 - |x^1 - z^1|^2 - |x^1 - z^2|^2 + |x^2 - z^2|^2 \right) / (2(t_3 - t_2)) \cdot \exp \left( -|x^2 - x^1|^2 - |x^1 - z^1|^2 - |x^1 - z^2|^2 + |x^2 - z^2|^2 \right) / (2(t_3 - t_2)) \\
= \exp \left( 2\sqrt{dn^{-\alpha}n^{-3+\delta/4}}/n^{-2+\alpha} \right) = \exp \left( 2\sqrt{dn^{-\delta/4}} \right). 
\]

**Step 3.** We use independence of \(B'\) and \(B''\), (5.1) and (5.6) to obtain for large \(n\),
\[
\begin{align*}
&\mathbb{P}(\{B'_s \in dz, B'_t \in dx^1, B'_t \in dx^2 \} \cap F(B') \cap F(B'')) \\
= \frac{\mathbb{P}(\{B'_s \in dz, B'_t \in dx^1, B'_t \in dx^2 \} \cap F(B')) \cdot \mathbb{P}(\{B'_t \in dx^1, B'_t \in dx^2 \} \cap F(B''))}{\mathbb{P}(\{B'_s \in dz, B'_t \in dx^1, B'_t \in dx^2 \} \cap F(B'))} \\
\leq \frac{\mathbb{P}(B'_s \in dz, B'_t \in dx^1) \cdot \mathbb{P}(B'_t \in dx^1, B'_t \in dx^2) \cdot (1 + \exp(-c_3n^{\delta/2}))^2}{\mathbb{P}(B'_s \in dz, B'_t \in dx^1) \cdot \mathbb{P}(B'_t \in dx^1, B'_t \in dx^2) \cdot (1 - \exp(-c_3n^{\delta/2}))^2} \\
\leq \exp \left( 2\sqrt{dn^{-\delta/4}} \right) \left( \frac{1 + \exp(-c_3n^{\delta/2})}{1 - \exp(-c_3n^{\delta/2})} \right)^2 \leq \exp \left( 4\sqrt{dn^{-\delta/4}} \right). 
\end{align*}
\]

Recall definitions of \(A_3, \Lambda''(Q), A^Q_5(z, \ldots, z^N)\) and \(A^Q_5\) stated in (4.9), (4.12), (4.15) and (4.16).

**Lemma 5.2.** (i) \(\lim_{n \to \infty} \mathbb{P}(A_3) = 1\).

(ii)
\[
\lim_{n \to \infty} \mathbb{P} \left( \bigcap_{Q \in Q} A^Q_5 \right) = 1.
\]

**Proof.** (i) We use (5.2) to see that
\[
\mathbb{P}(A_3^Q) \leq \sum_{1 \leq j \leq n} \mathbb{P} \left( \sup_{s,t \in [t_2, t_3]} \left| B^j_s - B^j_t \right| \geq 2n^{-\alpha+3\delta/4} \right) \leq n \exp \left( -c_1n^{\delta/2} \right).
\]

Since \(\delta > 0\), (i) follows.
(ii) The number of \( Q \in \mathcal{Q} \) is bounded by \( n^{\alpha_d|\Lambda|} \). Note that (5.2) applies (with a different constant) when \( 2n^{-\alpha+38/4} \) is replaced with \( (1/2)n^{-\alpha+38/4} \). Thus

\[
\mathbb{P} \left( \left( \bigcap_{Q \in \mathcal{Q}} A^Q \right)^c \right) = \mathbb{P} \left( \bigcup_{Q \in \mathcal{Q}} \left( \bigcup_{(z^1, \ldots, z^N) \in \Lambda^n(Q)} A^Q_{\delta} (z^1, \ldots, z^N) \right)^c \right) \\
\leq n^{\alpha_d|\Lambda|} n \exp \left(-c_2 n^{k/2}\right).
\]

Part (ii) follows since \( \delta > 0 \).

\[\Box\]

5.2. Conditioned space-time Brownian motion. Let \( \mathbb{P}_{x,y,s} \) denote the distribution of \( \{B_t, 0 \leq t \leq s\} \) where \( B \) is Brownian motion starting from \( B_0 = x \), conditioned by \( \{B_s = y\} \), and conditioned to stay inside \( \Lambda \) on the interval \([0, s]\). The distribution \( \mathbb{P}_{x,y,s} \) can be thought of as the distribution of the space component of the space-time Brownian motion \((B_t, t)\) conditioned by the parabolic function \( h \) in \( \Lambda \times [0, s] \), equal to 0 everywhere on the boundary of \( \Lambda \times [0, s] \) except for \((y, s) \). Such processes are known as \( h \)-processes.

Let \( p_t(x, y) \) denote the transition density for Brownian motion killed upon exiting \( \Lambda \). Let \( \lambda > 0 \) be the first eigenvalue for \( -(\frac{1}{2}) \Delta \) with Dirichlet boundary conditions in \( \Lambda \) and let \( \varphi > 0 \) be the corresponding eigenfunction. A bounded Lipschitz domain is intrinsically ultrcontractive (IU). We will need only one result on IU domains, cited below, so we will not define IU domains here; instead we ask the reader to consult, e.g., [10, 1]. It follows from [1, (1.8)] that for any \( \eta > 0 \) there exists \( u \) such that for \( s \geq u \) and \( x, y \in \Lambda \),

\[
1 - \eta \leq \frac{p_s(x, y)}{e^{-x} \varphi(x) \varphi(y)} \leq 1 + \eta.
\]

**Lemma 5.3.** Let \( \tilde{p}_t(x, y) \) denote the transition density for Brownian motion conditioned to stay in \( \Lambda \) on \([0, t]\). There exists \( s_1 \) and \( c_1, c_2, \gamma \in (0, \infty) \) such that for all \( t \geq s_1 \) and \( x, y \in \Lambda \),

\[
c_1 \text{ dist}(y, \Lambda) \gamma \leq \tilde{p}_t(x, y) \leq c_2.
\]

Results of this type are well known but we could not find an exact reference for (5.9).

**Proof of Lemma 5.3.** Since \( \Lambda \) is a bounded Lipschitz domain with the Lipschitz constant less than 1, we can find \( \rho > 0 \), a finite number \( k_1 \) of points \( x^k \in \partial \Lambda \), orthonormal coordinate systems \( CS_k \) and Lipschitz functions \( \psi_k : \mathbb{R}^{d-1} \rightarrow \mathbb{R} \) with the Lipschitz constant less than 1, satisfying the following conditions. For \( y \in \partial \Lambda \) and \( r > 0 \), let

\[
\Gamma_k(y, r) = \{ y + (x_1, \ldots, x_d) : x_1^2 + \cdots + x_{d-1}^2 \leq r^2, x_d^2 \leq r^2 \}, \quad \text{in } CS_k.
\]

Moreover, in \( CS_k \),

\[
\Lambda \cap \Gamma_k(x^k, \rho) = \{ (x_1, \ldots, x_d) \in \Gamma_k(x^k, \rho) : x_d > \psi_k(x_1, \ldots, x_{d-1}) \}, \quad k = 1, \ldots, k_1,
\]

\[
\{ x \in \Lambda : \text{dist}(x, \Lambda^c) \leq \rho/4 \} \subset \bigcup_{1 \leq k \leq k_1} \Gamma_k(x^k, \rho/2).
\]

Fix \( CS_k \) and \( y \in \partial \Lambda \cap \Gamma_k(x^k, \rho/2) \). Let \( n_k = (0, \ldots, 0, 1) \) in \( CS_k \). Let \( L_y = \{ y + an_k : a > 0 \} \) and let \( C_{y, a} \) be the open cone consisting of all half-lines with the endpoint at \( y \) forming the angle less than \( \alpha \) with \( L_y \). Note that

\[
\Gamma_k(y, \rho/2) \cap \Lambda \subset \Gamma_k(x^k, \rho) \cap \Lambda,
\]

\[
C_{y, \pi/8} \cap \Gamma_k(y, \rho/2) \subset \Lambda \cap \Gamma_k(y, \rho/2).
\]
It is well known (see, e.g., “Application” on page 192 in [8]) that there exists a positive harmonic function \( h_1(x), x \in C_{y,\pi/8} \), such that \( h_1(x) = 0 \) for \( x \in \partial C_{y,\pi/8} \) and \( h_1(x) = |x - y|^\gamma f(\theta) \) for some \( \gamma > 0 \) and a function \( f \), where \( \theta \) is the angle between the line segment \( 0, x \) and \( L_y \). We can and will assume that \( f(0) = 1 \).

Let \( h_2(x) \) be a positive harmonic function in \( \Lambda \cap \Gamma_k(y, \rho/2) \), with the boundary values \( h_2(x) = h_1(x) \) for \( x \in C_{y,\pi/8} \cap \partial \Gamma_k(y, \rho/2) \cap \Lambda \) and \( h_2(x) = 0 \) for \( x \in \partial (\Lambda \cap \Gamma_k(y, \rho/2)) \setminus C_{y,\pi/8} \). Then \( h_2(x) \geq h_1(x) \) for \( x \in \partial (C_{y,\pi/8} \cap \Gamma_k(y, \rho/2)) \). It follows by the elliptic maximum principle that \( h_2(x) \geq h_1(x) \) for \( x \in L_y \cap \Lambda \cap \Gamma_k(y, \rho/2) \). Hence,

\[
\begin{align*}
(5.10) & \quad h_2(x) \geq |x - y|^\gamma, \quad x \in L_y, \\
(5.11) & \quad h_2(x) \leq c_3 := \sup \{h_1(x) : x \in C_{y,\pi/8}, |x - y| \leq 2\rho\}, \quad x \in \Lambda \cap \Gamma_k(y, \rho/2).
\end{align*}
\]

Recall that \( p_t(x, z) \) denotes the transition density for Brownian motion killed upon exiting \( \Lambda \). Let \( g(t, x) = h_2(x) \) for \( x \in \Lambda \cap \Gamma_k(y, \rho/2) \) and \( t \geq 0 \). The functions \( (z, t) \to p_t(x, z) \) and \( (z, t) \to g(t, z) \) are parabolic, i.e., they are solutions to the heat equation and they have zero boundary values on \( \partial \Lambda \cap \Gamma_k(y, \rho/2) \times \mathbb{R} \). Hence, by [17, Thm. 1.6], for \( s > 0 \), there exist \( c_4 \) and \( a > b_1 > 0 \) depending only on \( \Lambda \) and \( s \), such that for \( t \geq s, 0 < b < b_1 \) and \( x \in \Lambda \),

\[
\frac{g(t, y + b n_k)}{p_t(x, y + b n_k)} \leq c_4 \frac{g(t + 2a^2, y + a n_k)}{p_{t-2a^2}(x, y + a n_k)},
\]

so, using (5.10),

\[
\begin{align*}
p_t(x, y + b n_k) & \geq c_4^{-1} \frac{g(t, y + b n_k)}{g(t + 2a^2, y + a n_k)} p_{t-2a^2}(x, y + a n_k) \\
& = c_4^{-1} \frac{h_2(y + b n_k)}{h_2(y + a n_k)} p_{t-2a^2}(x, y + a n_k) \geq c_4^{-1} \frac{b^\gamma}{h_2(y + a n_k)} p_{t-2a^2}(x, y + a n_k).
\end{align*}
\]

Let \( \Lambda_0^c \) be the set of all points of the form \( y + a_1 n_k \) with \( 0 < a_1 < a \), where \( y \) can be any point in \( \partial (\Lambda \cap \Gamma_k(x_k, \rho/2)) \) and \( k \) can be any \( 1, \ldots, k_1 \). Let \( \Lambda_a = \Lambda \setminus \Lambda_0^c \).

By (5.8), for any \( \eta > 0 \), \( a \) as above, sufficiently large \( s_1 \) and \( t \geq s_1 \),

\[
\begin{align*}
p_t(x, y + a n_k) & \geq p_t(x, y + a n_k)/2, \\
\inf_{z \in \Lambda_a} p_t(x, z) & \geq (1 - \eta) e^{-L t} \varphi(x) \inf_{z \in \Lambda_a} \varphi(z).
\end{align*}
\]

Since \( \text{dist}(\Lambda_a, \Lambda^c) > 0 \), \( c_5 := \inf_{z \in \Lambda_a} \varphi(z) > 0 \). This observation, (5.11), (5.12), (5.13) and (5.14) yield

\[
(5.15) \quad p_t(x, y + b n_k) \geq c_4^{-1} \frac{b^\gamma}{c_5} (1 - \eta) e^{-L t} \varphi(x) c_5 = c_6 b^\gamma e^{-L t} \varphi(x).
\]

By (5.8), for sufficiently large \( t \),

\[
\int_\Lambda p_t(x, z) dz \leq 2 e^{-L t} \varphi(x) \int_\Lambda \varphi(z) dz = c_7 e^{-L t} \varphi(x).
\]

It follows from this and (5.15) that

\[
\overline{p_t}(x, y + b n_k) = \frac{p_t(x, y + b n_k)}{\int_\Lambda p_t(x, z) dz} \geq \frac{c_6 b^\gamma e^{-L t} \varphi(x)}{c_7 e^{-L t} \varphi(x)} = c_8 b^\gamma.
\]

This proves the lower bound in (5.9) for \( y \in \Lambda_0^c \) because \( b \geq \text{dist}(y + b n_k, \Lambda^c) \). We can make the bound valid for all \( y \in \Lambda \) by making \( c_1 > 0 \) smaller, if necessary.
Since
\[ \tilde{p}_t(x, y) = \frac{p_t(x, y)}{\int_A p_t(x, y) dy}, \]
it follows from (5.8) that for any \( \eta_1 > 0 \) there exists a \( u > 0 \) such that for \( s \geq u \) and \( x, y \in \Lambda \),
\begin{equation}
\left( \int_{\Lambda} \varphi(x) dx \right)^{-1} (1 - \eta_1) \leq \frac{\tilde{p}_s(x, y)}{\varphi(y)} \leq \left( \int_{\Lambda} \varphi(x) dx \right)^{-1} (1 + \eta_1).
\end{equation}
The upper bound in (5.9) follows because \( \sup_{y \in \Lambda} \varphi(y) < \infty. \)

**Remark 5.4.** Suppose that \( \mathbb{P}' \) and \( \mathbb{P}'' \) are probability measures on space-time trajectories \( \{ (B_s, s), 0 \leq s \leq t \} \) such that under each of these measures, the process \( \{ B_s, 0 \leq s \leq t \} \) is the space component of a space-time Brownian motion conditioned to exit \( \Lambda \times \{ 0, t \} \) via \( \Lambda \times \{ t \} \), but with different exit distributions. It follows from the theory of \( h \)-processes (i.e., conditioned Brownian motion; see [13]) that under both \( \mathbb{P}' \) and \( \mathbb{P}'' \), the process \( \{ B_s, 0 \leq s \leq t \} \) conditioned by \( \{ B_0 = x, B_t = y \} \) has the same distribution. It follows that for \( u \leq t \), the Radon-Nikodym derivative \( d\mathbb{P}' / d\mathbb{P}'' \) on the \( \sigma \)-field \( \sigma(B_s, 0 \leq s \leq u) \) depends only on \( B_u \).

Recall that for \( x, y \in \Lambda \), \( \mathbb{P}_{x,y,s} \) denotes the distribution of \( \{ B_t, 0 \leq t \leq s \} \) where \( B \) is Brownian motion starting from \( D_0 = x \), conditioned to stay inside \( \Lambda \) on the interval \( [0, s] \) and further conditioned by \( \{ B_s = y \} \).

**Lemma 5.5.** Fix any \( t_1, \epsilon > 0 \). There exists \( s_1 > t_1 \) so large that for all \( x, y_1, y_2 \in \Lambda \) and \( t_2 \geq s_1 \),
\begin{equation}
1 - \epsilon \leq \frac{d\mathbb{P}_{x,y_1,t_2}}{d\mathbb{P}_{x,y_2,t_2}} \leq 1 + \epsilon
\end{equation}
on the \( \sigma \)-field \( \mathcal{F}_{t_1}^B := \sigma(B_t, 0 \leq t \leq t_1) \).

**Proof.** Recall that \( \tilde{p}_t(x, y) \) denotes the transition density for Brownian motion conditioned to stay in \( \Lambda \) on \( [0, t] \). By the Markov property applied at \( t_1 \), the density of \( B_{t_2} \) under \( \mathbb{P}_{x,y_1,t_2} \) evaluated at \( z \) is
\[ \frac{\tilde{p}_{t_1}(x, z) \tilde{p}_{t_2-t_1}(z, y_1)}{\tilde{p}_{t_2}(x, y_1)}, \]
and the analogous formula holds for \( \mathbb{P}_{x,y_2,t_2} \). This and Remark 5.4 imply that
\begin{equation}
\frac{d\mathbb{P}_{x,y_1,t_2}}{d\mathbb{P}_{x,y_2,t_2}} \bigg|_{\mathcal{F}_{t_1}^B} = \frac{\tilde{p}_{t_2-t_1}(z, y_1) \tilde{p}_{t_2}(x, y_2)}{\tilde{p}_{t_2}(x, y_1) \tilde{p}_{t_2-t_1}(z, y_2)},
\end{equation}
If we assume that \( B \) is defined on the canonical probability space then we can identify \( \{ B_t, 0 \leq t \leq t_2 \} \) with \( \{ \omega_t, 0 \leq t \leq t_2 \} \), where \( \omega \in C([0, t_2], \Lambda) \), and rewrite (5.18) as
\[ \frac{d\mathbb{P}_{x,y_1,t_2}}{d\mathbb{P}_{x,y_2,t_2}} \bigg|_{\mathcal{F}_{t_1}^B} (\omega) = \frac{\tilde{p}_{t_2-t_1}(\omega_{t_1}, y_1) \tilde{p}_{t_2}(x, y_2)}{\tilde{p}_{t_2}(x, y_1) \tilde{p}_{t_2-t_1}(\omega_{t_1}, y_2)}. \]
Combining this with (5.16) completes the proof. \( \square \)
Lemma 5.6. Recall from (3.5) that $\overline{P}_{t_2}$ refers to the Wiener measure conditioned on staying inside $\Lambda$ up to $t_2$. We have assumed that $f$ depends only on the values of the process in $[0, t_1]$ and $\overline{E}_{t_2}(f) > 0$. Assume that $t_2 - t_1 \geq s_1$ where $s_1$ is given in Lemma 5.3. Then there exist $c_1$ and $c_2$ such that for every $Q \in \mathcal{Q}$,

\begin{align}
(5.19)\quad c_1 n^{-\alpha+\gamma(-\alpha+3\delta/4)} \leq \overline{P}_{t_2}(G_Q) \leq c_2 n^{-\alpha},
\end{align}

\begin{align}
(5.20)\quad c_1 \overline{E}_{t_2}(f) n^{-\alpha+\gamma(-\alpha+3\delta/4)} \leq \overline{E}_{t_2}(f) \leq c_2 \overline{E}_{t_2}(f) n^{-\alpha}.
\end{align}

Proof. We will apply Remark 5.4. The volume of $Q$ is $n^{-\alpha}$. Thus (5.19) follows from the assumption (4.6) and Lemma 5.3.

Let $E_{x,y,t}$ denote the expectation corresponding to the probability measure $\mathbb{P}_{x,y,t}$ discussed in Lemma 5.5. Using that lemma, for an arbitrary fixed $z \in \Lambda$,

\begin{align*}
\overline{E}_{t_2}(f_Q) &= \int_Q \int_{\Lambda_1} E_{x,y,t_2}(f) d\mu_n(dx) \overline{P}_{t_2}(dy) \leq \int_Q \int_{\Lambda_1} (1 + \varepsilon) E_{x,z,t_2}(f) d\mu_n(dx) \overline{P}_{t_2}(dy) \\
&= (1 + \varepsilon) \overline{P}_{t_2}(G_Q) \int_{\Lambda_1} E_{x,z,t_2}(f) d\mu_n(dx) \\
&= (1 + \varepsilon) \overline{P}_{t_2}(G_Q) \int_{\Lambda_1} \int_{\Lambda_1} E_{x,y,t_2}(f) d\mu_n(dx) \overline{P}_{t_2}(dy) \\
&\leq (1 + \varepsilon)^2 \overline{P}_{t_2}(G_Q) \int_{\Lambda_1} \int_{\Lambda_1} E_{x,y,t_2}(f) d\mu_n(dx) \overline{P}_{t_2}(dy) = (1 + \varepsilon)^2 \overline{P}_{t_2}(G_Q) \overline{E}_{t_2}(f).
\end{align*}

Similarly,

\[ \overline{E}_{t_2}(f_Q) \geq (1 - \varepsilon)^2 \overline{P}_{t_2}(G_Q) \overline{E}_{t_2}(f). \]

Now (5.20) follows from (5.19).

Recall that $\mu$ and $\mu_n$ are probability measures supported in $\Lambda_1 \subset \Lambda$ and $\mu_n$ converge weakly to $\mu$ as $n \to \infty$. The next lemma follows essentially from the main theorem in [25] but we need a specific order of quantifiers which does not seem to follow directly from that theorem. Recall that $\overline{P}_{t}^\mu$ denotes the distribution of Brownian motion conditioned to stay in $\Lambda$ forever, with the initial distribution $e\phi(x)\mu(dx)$, where $e > 0$ is the normalizing constant.

Lemma 5.7. For every $\varepsilon > 0$ and $t_1 > 0$ there exists $s_1$ such that for every positive continuous function $f$ on $C[0, t_1]$ with $\|f\|_\infty \leq 1$ and $t \geq s_1$,

\[ 1 - \varepsilon \leq \liminf_{n \to \infty} \frac{\overline{E}_{t_1}^{\mu_n}(f)}{\overline{E}^{\mu_n}(f)} \leq \limsup_{n \to \infty} \frac{\overline{E}_{t_1}^{\mu_n}(f)}{\overline{E}^{\mu_n}(f)} \leq 1 + \varepsilon. \]

Proof. Let $\varphi > 0$ denote the first eigenfunction of $(-\frac{1}{2})\Delta$ in $\Lambda$ with Dirichlet boundary conditions and let $\lambda > 0$ be the corresponding eigenvalue. The function $h(x, t) = e^{\lambda t}\varphi(x)$ is parabolic in $\Lambda \times (0, \infty)$. The transition density $\tilde{p}_t(x, y)$ for Brownian motion conditioned not to exit $\Lambda$ is given by, for $t > 0$ and $x, y \in \Lambda$,

\begin{align}
\tilde{p}_t(x, y) = \frac{1}{h(x, 0)p_t(x, y)}h(y, t)e^{\lambda t}\varphi(y)\varphi(x)p_t(x, y).
\end{align}

Under both $\overline{P}_t$ and $\tilde{P}$, the process $\{B_s, 0 \leq s \leq t_1\}$ is the space component of a space-time Brownian motion conditioned to exit $\Lambda \times (0, t_1)$ via $\Lambda \times \{t_1\}$, but with different exit
distributions. We will show that for every \( \varepsilon > 0 \) there exists \( s_1 \) such that for \( t \geq s_1 \) and \( x, y, v, z \in \Lambda \),

\[
1 - \varepsilon \leq \frac{\bar{P}_t(B_{t_1} \in dy \mid B_0 = x, B_t = v)}{\bar{P}(B_{t_1} \in dy \mid B_0 = x, B_t = z)} \leq 1 + \varepsilon. \tag{5.22}
\]

By (5.21),

\[
\bar{P}_t(B_{t_1} \in dy \mid B_0 = x, B_t = v) = \frac{P_{t_1}(x, y)P_{t-t_1}(y, v)}{P_t(x, y)} dy,
\]

\[
\bar{P}_t(B_{t_1} \in dy \mid B_0 = x, B_t = z) = \frac{P_{t_1}(x, y)\hat{P}_{t-t_1}(y, z)}{P_t(x, z)} dy
\]

\[
\begin{aligned}
&= \frac{e^{\lambda t_1}\varphi(y)P_t(x, y)e^{\lambda(t-t_1)}\varphi(z)P_{t-t_1}(y, z)}{e^{\lambda t}\varphi(x)P_t(x, z)} dy \\
&= \frac{P_{t_1}(x, y)P_{t-t_1}(y, z)}{P_t(x, z)} dy.
\end{aligned}
\]

Hence,

\[
\frac{\bar{P}_t(B_{t_1} \in dy \mid B_0 = x, B_t = v)}{\bar{P}(B_{t_1} \in dy \mid B_0 = x, B_t = z)} = \frac{P_{t-t_1}(y, v)P_t(x, z)}{P_t(x, v)P_{t-t_1}(y, z)}.
\]

By (5.8), for any \( \eta > 0 \) there exists \( s_2 \) such that for \( t - t_1 \geq s_2 \) and \( x, y, v, z \in \Lambda \),

\[
\frac{\bar{P}_t(B_{t_1} \in dy \mid B_0 = x, B_t = v)}{\bar{P}(B_{t_1} \in dy \mid B_0 = x, B_t = z)} \leq \frac{(1 + \eta)^2}{(1 - \eta)^2} e^{-\lambda(t-t_1)}\varphi(y)\varphi(v)e^{-\lambda t}\varphi(x)\varphi(z) \leq (1 + \eta)^2 \frac{1 - \eta)^2}{1 - \eta^2}.
\]

This proves the upper bound in (5.22). The lower bound can be proved in an analogous way.

Suppose that \( f \) is a positive continuous function on \( C[0, t] \) with \( \|f\|_\infty \leq 1 \). Recall from Remark 5.4 that under both \( \bar{P}_t \) and \( \hat{P}_t \), the process \( \{B_s, 0 \leq s \leq t_1\} \) conditioned by \( \{B_0 = x, B_{t_1} = y\} \) has the same distribution. Hence, (5.22) implies that for every \( \varepsilon > 0 \) there exists \( s_1 \) such that for \( t \geq s_1 \) and \( x, v, z \in \Lambda \),

\[
1 - \varepsilon \leq \frac{\overline{E}_t(f \mid B_0 = x, B_t = v)}{\overline{E}(f \mid B_0 = x, B_t = z)} \leq 1 + \varepsilon.
\]

Since \( v, z \in \Lambda \) are arbitrary, for every \( \varepsilon > 0 \) there exists \( s_1 \) such that for \( t \geq s_1 \) and \( x \in \Lambda \),

\[
1 - \varepsilon \leq \frac{\overline{E}_t(f \mid B_0 = x)}{\overline{E}(f \mid B_0 = x)} \leq 1 + \varepsilon. \tag{5.24}
\]

By (5.8), for any \( \eta > 0 \), sufficiently large \( t \), normalizing constants \( c_n \) and \( c_s \),

\[
\bar{P}_t(B_0 \in dx) = c_n\mu_n(dx) \int_\Lambda P_t(x, y)dy \geq (1 - \eta)c_n\mu_n(dx) \int_\Lambda e^{-\lambda t}\varphi(x)\varphi(y)dy \\
= c_n\epsilon(1 - \eta)\varphi(x)\mu_n(dx),
\]

and similarly

\[
\bar{P}_t(B_0 \in dx) \leq c_n c_s(1 + \eta)\varphi(x)\mu_n(dx). \tag{5.26}
\]
Recall that $\mu_n$ converge weakly to $\mu$ and $\bar{\mathbb{P}}(B_0 \in dx) = c\varphi(x)\mu(dx)$ for a normalizing constant $c$. It follows from (5.26) that

$$1 \leq \liminf_{n \to \infty} \int c_n c_*(1 + \eta)\varphi(x)\mu_n(dx) = \liminf_{n \to \infty} \int \frac{c_n c_*(1 + \eta)}{c} c\varphi(x)\mu_n(dx)$$

$$= \liminf_{n \to \infty} \int \frac{c_n c_*(1 + \eta)}{c} c\varphi(x)\mu(dx) = \liminf_{n \to \infty} \frac{c_n c_*(1 + \eta)}{c},$$

so

$$\liminf_{n \to \infty} c_n \geq c/(c_*(1 + \eta)).$$

This and (5.25) imply that

$$(5.27) \quad \liminf_{n \to \infty} \bar{\mathbb{P}}_t(B_0 \in dx) \geq \liminf_{n \to \infty} c_n c_*(1 - \eta)\varphi(x)\mu_n(dx) \geq \liminf_{n \to \infty} \frac{1 - \eta}{1 + \eta} c\varphi(x)\mu_n(dx) = \frac{1 - \eta}{1 + \eta} \bar{\mathbb{P}}(B_0 \in dx).$$

Similarly,

$$(5.28) \quad \limsup_{n \to \infty} \bar{\mathbb{P}}_t(B_0 \in dx) \leq \frac{1 + \eta}{1 - \eta} \bar{\mathbb{P}}(B_0 \in dx).$$

We have assumed in Theorem 4.1 that $\mu$ is supported in a set $\Lambda_1 \subset \Lambda$ such that $\text{dist}(\Lambda_1, \Lambda^c) > 0$. It follows from [9, Cor. 1] that if $x^k \in \Lambda_1$ and $x^k \to x^\infty \in \Lambda_1$ then $\bar{\mathbb{E}}_t(f \mid B_0 = x^k) \to \bar{\mathbb{E}}_t(f \mid B_0 = x^\infty)$. A standard coupling argument can be applied to construct processes $B^n$ with the same transition probabilities as those of $B$, on the same probability space, with the initial distributions $\mu_n$ for $B^n$ and $\mu$ for $B$, and such that $|B^n_0 - B_0| \to 0$, a.s., as $n \to \infty$. These observations, (5.24) and (5.28) imply that

$$\lim sup_{n \to \infty} \bar{\mathbb{E}}_t(f) = \lim sup_{n \to \infty} \int \bar{\mathbb{E}}_t(f \mid B_0 = x)\bar{\mathbb{P}}_t(B_0 \in dx)$$

$$\leq \lim sup_{n \to \infty} \int \bar{\mathbb{E}}_t(f \mid B_0 = x) \frac{1 + \eta}{1 - \eta} \bar{\mathbb{P}}(B_0 \in dx)$$

$$\leq \lim sup_{n \to \infty} \int (1 + \varepsilon)\bar{\mathbb{E}}(f \mid B_0 = x) \frac{1 + \eta}{1 - \eta} \bar{\mathbb{P}}(B_0 \in dx) = (1 + \varepsilon) \frac{1 + \eta}{1 - \eta} \bar{\mathbb{E}}(f).$$

Similarly, one can use (5.24) and (5.27) to obtain

$$\lim inf_{n \to \infty} \bar{\mathbb{E}}_t(f) \geq (1 - \varepsilon) \frac{1 - \eta}{1 + \eta} \bar{\mathbb{E}}(f).$$

Since $\eta > 0$ can be arbitrarily small, the lemma follows. \hfill $\square$

**Lemma 5.8.** Recall that we have assumed that $\bar{\mathbb{E}}_t(f) > 0$. For every $t \geq t_1$, $\bar{\mathbb{E}}_t(f) > 0$.

**Proof.** It follows from Lemma 5.3 and (5.23) that the distributions of $B_{t_1}$ under $\bar{\mathbb{P}}_t$ and $\bar{\mathbb{P}}_{t_1}$ have strictly positive densities in $\Lambda$, say $p_s(x)$ and $p_{s*}(x)$. Let $\Lambda_\ast \subset \Lambda$ and $c_1, c_2 > 0$ be such that $p_s(x)/p_{s*}(x) > c_1$ and $\bar{\mathbb{E}}_{t_1}(f1_{\{B_{t_1} \in \Lambda_\ast\}}) > c_2$.

According to Remark 5.4, under both $\bar{\mathbb{P}}_t$ and $\bar{\mathbb{P}}_{t_1}$, the process $\{B_s, 0 \leq s \leq t_1\}$ conditioned by $\{B_0 = x, B_{t_1} = y\}$ has the same distribution. Hence

$$\bar{\mathbb{E}}_t(f) \geq \bar{\mathbb{E}}_t(f1_{\{B_{t_1} \in \Lambda_\ast\}}) \geq c_1 \bar{\mathbb{E}}_{t_1}(f1_{\{B_{t_1} \in \Lambda_\ast\}}) \geq c_1 c_2 > 0.$$

$\square$
5.3. Villemois’s estimate. Recall that we have assumed that \( u \) is fixed and \( t_2 \leq u + 1 \). This easily implies that for some \( c_1 > 0 \) and all probability measures \( \mu_n \) supported in \( \Lambda_1 \),

\[
\mathbb{P}_{\mu_n}(\tau_\Lambda > t_2) \geq c_1,
\]

where \( \mathbb{P}_{\mu_n} \) represents the distribution of the driving process \( B \) with the initial distribution \( \mu_n \).

Recall definitions of \( A_1, A_2 \) and \( \mathcal{H}_t^n \) stated in (4.7)-(4.8) and (3.4).

**Lemma 5.9.** \( \lim_{n \to \infty} \mathbb{P}(A_1 \cap A_2) = 1. \)

**Proof.** By Theorem 3.2 and (5.29),

\[
\mathbb{P} \left( \left| \mathcal{H}_{t_2}^n \left( 1_{G_Q} \right) - \mathbb{P}_{t_2}(G_Q) \right| \geq n^{-\alpha d + \gamma(-\alpha + 3\delta/4) - \delta} \right) \\
\leq \mathbb{E} \left( \left| \mathcal{H}_{t_2}^n \left( 1_{G_Q} \right) - \mathbb{P}_{t_2}(G_Q) \right| \right) n^{\alpha d - \gamma(-\alpha + 3\delta/4) + \delta} \\
\leq 2 \left( 1 + \sqrt{2} \right) c_1^{-1} n^{-1/2} n^{\alpha d - \gamma(-\alpha + 3\delta/4) + \delta}.
\]

Recall from (4.2) that we have assumed that \( \alpha \leq (1/2 - 2\delta + 3\gamma\delta/4)/(\gamma + 2d) \). Hence, \(-1/2 + \alpha d - \gamma(-\alpha + 3\delta/4) + \delta \leq -\alpha d - \delta \). This and (5.30) show that for some \( c_2 \),

\[
\mathbb{P} \left( \left| \mathcal{H}_{t_2}^n \left( 1_{G_Q} \right) - \mathbb{P}_{t_2}(G_Q) \right| \geq n^{-\alpha d + \gamma(-\alpha + 3\delta/4) - \delta} \right) \leq c_2 n^{-\alpha d - \delta}.
\]

Since \( \Lambda \) is bounded set, the number of \( Q \) in \( Q \) is bounded by \( c_3 n^{\alpha d} \). It follows from the definition (4.7) of \( A_1 \) and (5.31) that

\[
\mathbb{P}(A_1^c) \leq \sum_{Q \in Q} \mathbb{P} \left( \left| \mathcal{H}_{t_2}^n \left( 1_{G_Q} \right) - \mathbb{P}_{t_2}(G_Q) \right| \geq n^{-\alpha d + \gamma(-\alpha + 3\delta/4) - \delta} \right) \leq c_3 n^{\alpha d} c_2 n^{-\alpha d - \delta}
\]

\[
(5.32) = c_3 c_2 n^{-\delta}.
\]

Hence, \( \lim_{n \to \infty} \mathbb{P}(A_1) = 1 \). The proof for \( A_2 \) is completely analogous. \( \square \)

5.4. Paths close to the boundary and branching. Our next lemma is comprised of two claims, corresponding to events \( C_j^1 \) and \( C_j^2 \) defined below. The first step of the proof is the only common aspect of the two claims.

We believe that some “unusual” events are very unlikely to occur at an arbitrarily chosen fixed time \( u \). This we cannot prove. But we will prove that for a fixed \( u \), there is a time with this property in \([u, u + 1]\).

Consider any \( u > 0 \), let \( \Delta t = n^{-2\alpha + \delta} \), \( k_1 = \lfloor 1/\Delta t \rfloor \) and \( s_j = u + j\Delta t \) for \( j = 0, \ldots, k_1 \). Recall that \( J^n \) is the spine of \( X^n \). Let

\[
C_j^1 = \{ \text{dist}(J^n_{s_j}, \Lambda^c) \leq 4n^{-\alpha + 3\delta/4} \},
\]

\[
C_j^2 = \{ J^n \text{ passes through a branch point in } [s_j, s_{j+1}] \},
\]

\[
C_j = C_j^1 \cup C_j^2.
\]

**Lemma 5.10.** For every \( \varepsilon > 0 \) and \( u > 0 \) there exists \( n_1 \) so large that for \( n \geq n_1 \) there exists \( j \in \{0, \ldots, k_1 \} \) such that \( \mathbb{P}(C_j) \leq \varepsilon. \)
**Proof.** Step 1. Fix \( \epsilon \in (0,1) \) and \( u > 0 \). Suppose that for a given \( n \), there is no \( j \in \{0, \ldots, k_1\} \) such that \( \mathbb{P}(C_j) \leq \epsilon \). Then

\[
\mathbb{E} \left( \sum_{1 \leq j \leq k_1} 1_{C_j} \right) \geq \epsilon k_1.
\]

Let

\[
p = \mathbb{P} \left( \sum_{1 \leq j \leq k_1} 1_{C_j} \leq \epsilon k_1 / 2 \right).
\]

Then

\[
\mathbb{E} \left( \sum_{1 \leq j \leq k_1} 1_{C_j} \right) \leq p \epsilon k_1 / 2 + (1 - p) k_1.
\]

Hence \( \epsilon k_1 \leq p \epsilon k_1 / 2 + (1 - p) k_1 \) and, therefore,

\[
p \leq \frac{1 - \epsilon}{1 - \epsilon / 2} < 1.
\]

It follows that

\[
\mathbb{P} \left( \sum_{1 \leq j \leq k_1} 1_{C_j} > \epsilon k_1 / 2 \right) = 1 - p \geq \frac{\epsilon / 2}{1 - \epsilon / 2} > 0.
\]

This implies that at least one of the following inequalities holds,

\[
\mathbb{P} \left( \sum_{1 \leq j \leq k_1} 1_{C_j^1} > \epsilon k_1 / 4 \right) \geq \frac{\epsilon / 4}{1 - \epsilon / 2} > 0,
\]

\[
\mathbb{P} \left( \sum_{1 \leq j \leq k_1} 1_{C_j^2} > \epsilon k_1 / 4 \right) \geq \frac{\epsilon / 4}{1 - \epsilon / 2} > 0.
\]

It will suffice to show that each of these inequalities fails for large \( n \).

Step 2. Recall that \( B(v,r) \) denotes a ball with center \( v \) and radius \( r \). Let \( B \) be Brownian motion and

\[
\mathcal{C}_j^1 = \{ \text{dist}(B_{s_j}, \Lambda^c) \leq 4n^{-\alpha + \delta/4} \}.
\]

Since \( \Lambda \) is a bounded Lipschitz domain, it is easy to see that for some \( c_1, c_2 > 0 \) and every \( x \in \Lambda \) there exists \( z \in \Lambda^c \) such that \( |z - x| \leq c_1 \text{dist}(x, \Lambda^c) \) and \( B(z, c_2 \text{dist}(x, \Lambda^c)) \subset \Lambda^c \). From here up to and including (5.39), \( \mathbb{P} \) will denote the distribution of Brownian motion starting from \( x \). By Brownian scaling, there exists \( p_1 > 0 \) such that for all \( x \),

\[
\mathbb{P} \left( \inf \{ t > 0 : B_t \in \Lambda^c \} \leq \text{dist}(x, \Lambda^c)^2 \right) \geq \mathbb{P} \left( \inf \{ t > 0 : B_t \in B(z, c_2 \text{dist}(x, \Lambda^c)) \} \leq \text{dist}(x, \Lambda^c)^2 \right) = p_1.
\]

Let

\[
j_0 = \inf \left\{ m \geq 0 : \text{dist}(B_{s_m}, \Lambda^c) \leq 4n^{-\alpha + \delta/4} \right\},
\]

\[
j_{i+1} = \inf \left\{ m > j_i : \text{dist}(B_{s_m}, \Lambda^c) \leq 4n^{-\alpha + \delta/4}, s_m \geq s_{j_i} + 4n^{-2\alpha + \delta/2} \right\}, \ i \geq 0.
\]
By (5.37) and the strong Markov property applied at \(s_j\), for every \(i\),
\[
\mathbb{P} \left( \inf \{ t > s_j : B_t \in \Lambda^c \} \leq s_{j+1} \right) \geq p_1.
\]
We apply the strong Markov property again to see that for \(k \geq 0\),
\[
(5.38) \quad \mathbb{P} \left( \inf \{ t > s_0 : B_t \in \Lambda^c \} > s_{j_k + 1} \right)
\]
\[
\leq \mathbb{P} \left( \bigcap_{j=0}^{k} \left\{ \inf \{ t > s_{j_i} : B_t \in \Lambda^c \} > s_{j_{i+1}} \right\} \right) \leq (1 - p_1)^{k+1}.
\]

If the event \(\sum_{1 \leq j \leq k} 1_{\mathcal{C}_j} > \varepsilon k_1/4\) occurred then \(s_{j} \leq u + 1\), where
\[
\ell = \frac{(\varepsilon k_1/4) \Delta t}{4n^{-2\alpha-3\delta/2}} = \frac{(\varepsilon [n^{2\alpha-\delta}/4])n^{-2\alpha+\delta}}{4n^{-2\alpha+3\delta/2}} \geq \frac{1}{16} \varepsilon n^{2\alpha-3\delta/2}.
\]

This and (5.38) imply that
\[
(5.39) \quad \mathbb{P} \left( \sum_{1 \leq j \leq k} 1_{\mathcal{C}_j} > \varepsilon k_1/6, \inf \{ t > s_0 : B_t \in \Lambda^c \} > u + 1 \right)
\]
\[
\leq \mathbb{P} (s_j \leq u + 1, \inf \{ t > s_0 : B_t \in \Lambda^c \} > u + 1) \leq \mathbb{P} (\inf \{ t > s_0 : B_t \in \Lambda^c \} > s_{j_1}) \leq (1 - p_1) \ell \leq (1 - p_1)^{\varepsilon n^{2\alpha-3\delta/2}/16}.
\]

We will apply methods and results from the proof of Theorem 1.3 in [5, p. 688] but we will use different notation.

Let \(\Lambda_2\) be such that \(\text{dist}(\Lambda, \Lambda_2) > 0\) and \(\text{dist}(\Lambda_2, \Lambda^c) > 0\). Let \(\mathcal{H} = \{k : n/4 \leq k \leq n\}\) and \(\mathcal{H}^c = \{1, \ldots, n\} \setminus \mathcal{H}\).

The set \(\mathcal{H}^c\) is needed in the argument so that we can assume that many processes \(X^k\) with \(k \in \mathcal{H}^c\) do not jump on the interval \([0, u+2]\); this is represented formally in (5.40) below. The processes that do not jump behave like Brownian motion conditioned to stay inside \(\Lambda\) and, therefore, they stay “far” from the boundary most of the time. This creates an opportunity for processes \(X^k\) with \(k \in \mathcal{H}\) to jump “far” from the boundary and hence gives them a chance to stay inside \(\Lambda\) for the rest of the time interval \([0, u+2]\). The processes in \(\mathcal{H}^c\) are not forgotten—they will be accounted for when we define \(\tilde{\mathcal{H}}\) below.

Consider \(c_3 > 0\) and let \(F = F(c_3)\) be the event that at least \(c_3n\) processes \(X^k\) with \(k \in \mathcal{H}^c\) stay within \(\Lambda_2\) on the interval \([0, u+2]\). Recall that \(X^k_0 \in \Lambda_1\) for all \(k\), a.s. By the law of large numbers, there exists \(c_3 = c_3(u) > 0\) such that
\[
(5.40) \quad \lim_{n \to \infty} \mathbb{P}(F) = 1.
\]

Informally, let \(M_k\) be the number of branching points on the tree of descendants of \(X^k\) on the interval \([0, u+2]\). Formally, let \(M_k\) be the number of branching points on DHPs \(H_{u+2}^{k}\) such that \(H_{u+2}^{k}(0) = X^k_0\). Every branching point appears on two different DHPs but we count it only once. It follows from [5, (2.11)] that for every \(r > 0\) there exists \(c_r^* < \infty\) such that for every \(k \in \mathcal{H}\) and sufficiently large \(n\),
\[
(5.41) \quad \mathbb{E}(M_k^r | F) \leq c_r^*.
\]

The proof of (5.41) is based on the branching structure and induction. Note that [5, (2.11)] is an upper bound for the \(r\)-th power of “the total number of jumps on the tree of descendants of particle \(m\)” defined just below [5, (2.10)]. But the definition given in [5] makes it clear that “jumps” include branching points on the tree of descendants of \(X^k\) defined earlier in this...
paragraph. Formula [5, (2.11)] does not contain conditioning on $F$. The conditioning on $F$ is implicit on pages 688–691, as indicated on page 688 of [5].

Recall $\xi$ introduced in (4.3). Let $r$ be so large that $\xi r > 1$. We have

$$\mathbb{P}(M_k \geq n^\xi \mid F) = \mathbb{P}(M_k^r \geq n^{\xi r} \mid F) \leq \mathbb{E}(M_k^r \mid F)n^{-\xi r} \leq c_r n^{-\xi r}.$$  

If we let

$$K = \bigcup_{k \in \mathbb{H}} \{M_k \geq n^\xi\}$$

then

$$(5.42) \quad \mathbb{P}(K \mid F) \leq c_r n^{1-\xi r}.$$  

We construct a Brownian motion $B^*$ killed on the boundary of $\Lambda$ as follows. Choose $k$ uniformly from $\mathbb{H}$. Follow $X^k$ from time 0 until the first time when $X^k$ exits $\Lambda$ or until the first branching point, whichever comes first. If the process exits $\Lambda$, stop. At a branching point, start following one of the two branches, with equal probabilities, with the random choice independent of $X^k$. Then apply the inductive construction—follow the branch until it exits $\Lambda$ or the next branching point, whichever comes first. If the process exits $\Lambda$, stop. At a branching point, start following one of the branches, with equal probabilities, with the random choice independent of $X^k$.

Let $D^*$ be the event on the first line of (5.39) but with random objects defined relative to $B^*$ in place of $B$ so that

$$(5.43) \quad \mathbb{P}(D^*) \leq (1 - p_1)^{\varepsilon n^{2\alpha - 3\delta/2}/16}.$$  

Let $D$ be the event that there is a DHP $H^k_{u+2}$ such that $H^k_{u+2}(0) = X^j_0$ for some $j \in \mathbb{H}$ and the event on the first line of (5.39) holds for this DHP in place of $B$. If $K^c \cap D$ holds then, in the above notation, $j$ will be chosen with probability $1/n$, and there will be at most $n^\xi$ choices in the construction of $B^*$ on the interval $[0, u + 2]$ so $B^*$ will follow the trajectory of $H^k_{u+2}$ all the way up to $u + 2$ with probability greater than or equal to $2^{-\xi}$. Thus

$$\mathbb{P}(D^*) \geq \mathbb{P}(D^* \mid K^c \cap D) \mathbb{P}(K^c \cap D) \geq (1/n) 2^{-\xi} \mathbb{P}(K^c \cap D).$$  

Hence, by (5.43),

$$\mathbb{P}(K^c \cap D) \leq n 2^{n\xi} \mathbb{P}(D^*) \leq n 2^{n\xi} (1 - p_1)^{\varepsilon n^{2\alpha - 3\delta/2}/16}$$

$$= \exp \left( \log n + n^\xi \log 2 + n^{2\alpha - 3\delta/2} \log(1 - p_1) \varepsilon / 16 \right).$$

This and our assumption (4.3) that $0 < \xi < 2\alpha - 3\delta/2$ imply that

$$(5.44) \quad \lim_{n \to \infty} \mathbb{P}(K^c \cap D) \leq \lim_{n \to \infty} \exp \left( \log n + n^\xi \log 2 + n^{2\alpha - 3\delta/2} \log(1 - p_1) \varepsilon / 16 \right) = 0.$$

Let $\hat{\mathbb{H}} = \{k : 1 \leq k \leq 3n/4\}$ and define $\hat{F}$, $\hat{K}$ and $\hat{D}$ relative to $\hat{\mathbb{H}}$ in the same way as $F$, $K$ and $D$ were defined relative to $\mathbb{H}$. Then, by symmetry and (5.40), (5.42) and (5.44),

$$(5.45) \quad \lim_{n \to \infty} \mathbb{P}(\hat{F}) = 1, \quad \mathbb{P}(\hat{K} \mid \hat{F}) \leq c_r n^{1-\xi r}, \quad \lim_{n \to \infty} \mathbb{P}(\hat{K}^c \cap \hat{D}) = 0.$$  

Note that

$$\left\{ \sum_{1 \leq j \leq k_1} 1_{C_j} > \varepsilon k_1/4 \right\} \subset D \cup \hat{D},$$
so
\[ P \left( \sum_{1 \leq j \leq k_1} 1_{C_j^1} > \varepsilon k_1/4 \right) \]
\[ \leq P(K^c \cap D) + P(K \cap F) + P(F^c) + P(\hat{K}^c \cap \hat{D}) + P(\hat{K} \cap \hat{F}) + P(\hat{F}^c) \]
\[ \leq P(K^c \cap D) + P(K \mid F) + P(F^c) + P(\hat{K}^c \cap \hat{D}) + P(\hat{K} \mid \hat{F}) + P(\hat{F}^c). \]
Hence, by (5.40), (5.42), (5.44) and (5.45),
\[ \lim_{n \to \infty} P \left( \sum_{1 \leq j \leq k_1} 1_{C_j^1} > \varepsilon k_1/4 \right) = 0, \]
contradicting (5.35). This completes the proof that (5.35) fails for all large \( n \).

Step 3. By (5.41), for \( \delta > 0 \) and all \( k \in H \),
\[ P(M_k \geq n^{2\alpha - 2\delta} \mid F) \leq E(M_k^r \mid F)/n^{r(2\alpha - 2\delta)} \leq c^* n^{-r(2\alpha - 2\delta)}, \]
so
\[ P \left( \bigcup_{k \in H} \{ M_k \geq n^{2\alpha - 2\delta} \} \mid F \right) \leq c^* n^{1-r(2\alpha - 2\delta)}. \]
By (4.4), \( 2\alpha - 2\delta > 0 \) so we can find \( r \) so large that \( 1 - r(2\alpha - 2\delta) < 0 \). It follows that
\[ \lim_{n \to \infty} P \left( \bigcup_{k \in H} \{ M_k \geq n^{2\alpha - 2\delta} \} \mid F \right) = 0. \]
By (5.40),
\[ \lim_{n \to \infty} P \left( \bigcup_{k \in H} \{ M_k \geq n^{2\alpha - 2\delta} \} \right) = 0. \]
One can prove in the same way that
\[ \lim_{n \to \infty} P \left( \bigcup_{1 \leq k \leq n} \{ M_k \geq n^{2\alpha - 2\delta} \} \right) = 0, \]
so
\[ \lim_{n \to \infty} P \left( \bigcup_{1 \leq k \leq n} \{ M_k \geq n^{2\alpha - 2\delta} \} \right) = 0. \]

The piece of the spine \( \{ J^n_t, 0 \leq t \leq u + 2 \} \) is a trajectory within the tree of descendants of some \( X^k \) on the interval \([0, u + 2]\). Therefore, the number \( M_J \) of branching points on \( \{ J^n_t, 0 \leq t \leq u + 2 \} \) must be less than or equal to \( M_k \) for some \( k \). If the event in (5.36) holds then \( M_J \geq \varepsilon n^{2\alpha - \delta}/12 \). Hence, (5.46) implies that (5.36) cannot hold for large \( n \).
6. A generalization. We will argue that our main result holds for some diffusions in Euclidean domains. However, in our examples, the generator of the diffusion must be associated in an appropriate way with the domain; in other words, we cannot prove that one can choose the generator and the domain independently.

Below \( \| \cdot \| \) denotes the Euclidean norm.

**Proposition 6.1.** Suppose that \( d \geq 1 \) and \( U \subset \mathbb{R}^d \) is a domain (open bounded set) with a \( C^1 \)-boundary.

Suppose that \( Q_i : \mathbb{R}^d \to \mathbb{R}^d \), \( Q_i = (Q_i, Q_{i2}, \ldots, Q_{id})^T \), is the gradient of a \( C^2(\mathbb{R}^d, \mathbb{R}) \)-function \( x_i \) for \( 1 \leq i \leq d \), and assume that \( a_i := \sum_{i=1}^d Q_{ij}^2 = \| Q_i \|^2 > 0 \) on \( U \) for all \( 1 \leq i \leq d \), and furthermore, that the map \((x_1, \ldots, x_d) : \mathbb{R}^d \to \mathbb{R}^d \) is invertible and the matrix \( Q := (Q_{ij}) \) is non-singular at each point of \( \mathbb{R}^d \).

Define \( H : \mathbb{R}^d \to (0, \infty) \) by

\[
H(u_1, ..., u_d) = \prod_{i=1}^d \| Q_i(u_1, ..., u_d) \|, 
\]

and the operator

\[
\mathcal{L} = \frac{1}{2H} \nabla \cdot H \nabla \phi
\]

on \( U \), where \( C : \mathbb{R}^d \to \mathbb{R}^{d \times d} \), and \( C \) is diagonal with \( c_{ii} := 1/a_i = \| Q_i \|^{-2}, 1 \leq i \leq d \).

Let \( \lambda \) denote the principal eigenvalue of \( \mathcal{L} \) on \( U \) with Dirichlet boundary conditions and let \( \phi_{\mathcal{L}, \lambda} > 0 \) denote the principal eigenfunction associated with \( \mathcal{L} \) and \( \lambda \) on \( U \). Let \( \mathcal{L}_{\phi_{\mathcal{L}, \lambda}} \) denote the generator of the \( \mathcal{L} \)-diffusion conditioned to stay in \( U \) forever, obtained by “Doob’s \( h \)-conditioning” in space-time by the parabolic function \((u, t) \to \phi_{\mathcal{L}}(u)e^{\lambda t}\).

Consider a sequence of Fleming-Viot processes driven by \( \mathcal{L} \)-diffusions and assume that their empirical initial distributions converge weakly to a measure supported on a compact subset of \( U \). The spines of these processes converge weakly to a diffusion corresponding to \( \mathcal{L}_{\phi_{\mathcal{L}, \lambda}} \).

**Proof.** (Sketch) As in the statement of the proposition, the points in \( U \) will be denoted by \((u_1, u_2, ..., u_d)\). By assumption, \( Q_i : \mathbb{R}^d \to \mathbb{R}^d \) is the gradient of some \( C^2(\mathbb{R}^d) \)-function \( x_i = x_i(u_1, ..., u_d) \), that is, \( Q_i = (Q_{i1}, Q_{i2}, \ldots, Q_{id}) = (\partial x_i/\partial u_1, ..., \partial x_i/\partial u_d) \).

Define the map \( x : \mathbb{R}^d \to \mathbb{R}^d \) by \( x = (x_1, ..., x_d) \); by assumption, it has an inverse map \( u : \mathbb{R}^d \to \mathbb{R}^d \). Informally speaking, we have two systems of coordinates on \( \mathbb{R}^d \), which we will call the \( x \)-system and the \( u \)-system.

Let \( D = \mathcal{L} \) and note that \( D \) is a \( C^1 \)-domain because, by assumption, \( Q \) is non-singular on \( \mathbb{R}^d \) so, using the Inverse Function Theorem, the map \( x : \mathbb{R}^d \to \mathbb{R}^d \) is a local diffeomorphism, under which the image of a \( C^1 \)-domain is a \( C^1 \)-domain again.

Let \( L = \Delta/2 \) be the operator in the \( x \)-system and let \( \phi \) denote the principal eigenfunction associated with \( L \) with Dirichlet boundary conditions on \( D \). Let \( L \phi \) denote the generator of Brownian motion conditioned to stay in \( D \) forever.

As is well known (or a straightforward calculation reveals), the operator \( \mathcal{L} = \frac{1}{2H} \nabla \cdot H \nabla \) in the \( u \)-system is the pullback of \( L \) in the \( x \)-system, where \( H \) and \( C \) are defined via the non-singular matrix-function \( Q \) as in the statement of the proposition. This means that \((Lf)(x) = (Lg)(u)\) for a twice differentiable \( f \) where \( g \) is the pullback of \( f \) (i.e., \( g(u) := f(x(u)) \)).

We will argue that the following claims hold true.
(1) The principal eigenvalues of the two operators agree.

(2) The pullback of $\phi$ is $\phi_{L^\phi;\lambda}$.

(3) The pullback of $L^\phi$ is $L^{\phi_{L^\phi;\lambda}}$.

(4) Brownian motion (with killing at $\partial D$) in the $x$-system becomes an $L$-diffusion (with killing at $\partial U$) in the $u$-system; the $L^\phi$-diffusion in the $x$-system becomes an $L^{\phi_{L^\phi;\lambda}}$-diffusion in the $u$-system.

These are all invariance properties (under coordinate transforms): (1) follows from the fact

$$\lambda(L;D) = \inf\{l \in \mathbb{R} \mid \exists 0 < f \in C^2(D) : (L-l)f = 0\},$$

$$\lambda(L;U) = \inf\{l \in \mathbb{R} \mid \exists 0 < g \in C^2(U) : (L-l)g = 0\},$$

while (2) follows because the infimum is in fact the minimum and for the minimal value the function $f$ (function $g$) is unique.

Next, (3) follows because $L^\phi(x) = \frac{1}{2} \phi' + \lambda$ and an analogous formula holds for $L^{\phi_{L^\phi;\lambda}}$, whereas (4) follows from (1)-(3) along with the fact that diffusions are uniquely determined via the corresponding martingale problems.

We conclude that the Fleming-Viot system on $D$ driven by Brownian motion describes the same process as the Fleming-Viot system on $U$ driven by the $L$-diffusion. In fact, we have a coupling of the two particle systems. In particular, the “$n$-spine” must refer to the same process: the unique infinite lineage. Hence, Theorem 4.1 implies that the claim concerning the limiting distribution of $n$-spines is applicable to the Fleming-Viot system on $U$ driven by the $L$-diffusion. We leave it to the reader to check that the notion of weak convergence is invariant under the smooth coordinate transformation, and a similar remark applies to the statements involving the initial empirical measures.

6.1. Examples. Below we give some examples of domains and operators which can be generated using Proposition 6.1.

1. In $d = 1$, our setup includes an arbitrary interval $U = (\alpha, \beta)$ with an operator

$$L = \frac{1}{2H} \frac{d^2}{du^2} H,$$

where $H = |Q|$ and $Q$ satisfies the assumptions of the theorem. For example, we can take $x(u) = \text{sgn}(u)|u|^a$, $a > 0$. This yields $Q(u) = ax^{a-1}$ on $U$ whenever $\alpha > 1$. Thus, writing $\gamma := 1 - a < 1$, the operator becomes the time-changed Bessel-generator

$$(L^\gamma f)(u) = \frac{2^{2\gamma}}{(1-\gamma)^2} \left( f''(u) + \frac{\gamma}{u} f'(u) \right).$$

2. Let $d = 2$. Then the proposition applies to all $C^1$-domains in $\mathbb{R}^2$ and all operators

$$L = \frac{1}{2H} \nabla \cdot HC\nabla$$

on $U$, where

$$C = \begin{bmatrix} \|
abla F\| & 0 \\ 0 & \|
abla G\| \end{bmatrix}^{-2},$$

$$H(u_1, u_2) = \|
abla F\|^2 \|
abla G\|^2,$$

and $(F, G) \in C^2(\mathbb{R}^2, \mathbb{R}^2)$ is invertible and such that

$$\frac{\partial F}{\partial u_1} \frac{\partial G}{\partial u_2} \neq \frac{\partial F}{\partial u_2} \frac{\partial G}{\partial u_1}.$$
3. Continuing the previous example, a more concrete class of operators for \( d = 2 \) can be obtained by choosing a \( C^1 \)-domain \( U \subset (0, \infty)^2 \) and \( \alpha, \beta, \gamma, \delta \in \mathbb{R} \) such that \( \alpha \delta \neq \beta \gamma \) and

\[
F(u, v) = \exp(\alpha u + \beta v), \quad G(u, v) = \exp(\gamma u + \delta v),
\]

yielding the operator \( \mathcal{L} \) on \( U \) for which

\[
(\mathcal{L} f)(u, v) = \frac{1}{2} e^{-2[\gamma^2 + \delta^2]u + 2\alpha u + 2\beta v} \nabla \cdot e^{2[\alpha u + \beta v]} \mathbf{C} \nabla f(u, v),
\]

where

\[
\mathbf{C} := \begin{bmatrix}
(\alpha^2 + \beta^2)e^{-2\alpha u - 2\beta v} & 0 \\
0 & (\gamma^2 + \delta^2)e^{-2\gamma u - 2\delta v}
\end{bmatrix}.
\]

This simplifies to

\[
(\mathcal{L} f)(u, v) = (\alpha^2 + \beta^2)e^{-2(\alpha u + \beta v)} \left( \partial^2 f/\partial u^2 + \gamma \partial f/\partial u \right)
+ (\gamma^2 + \delta^2)e^{-2(\gamma u + \delta v)} \left( \partial^2 f/\partial v^2 + \beta \partial f/\partial v \right).
\]

For example, when \( (\alpha, \beta) \) and \( (\gamma, \delta) \) are unit vectors in \( \mathbb{R}^2 \) and \( u_1 = \alpha u + \beta v \) and \( u_2 = \gamma u + \delta v \), then

\[
\mathcal{L} = e^{-2u_1} \left( \partial^2/\partial u^2 + \gamma \partial/\partial u \right) + e^{-2u_2} \left( \partial^2/\partial v^2 + \beta \partial/\partial v \right).
\]

7. Spine for the superprocess version of our Fleming-Viot process. We will briefly sketch an argument which shows that, in a sense, our main theorem has an analog for superprocesses. See [15] for an introduction to superprocesses. Let us point out an unfortunate terminological inconsistency—a Fleming-Viot superprocess (see, e.g., [15]) is not a natural superprocess analogue of our Fleming-Viot process because in the Fleming-Viot superprocess model the death rate does not depend on the spatial position of the particle, whereas in our model deaths occur only at \( \partial D \). Intuitively, when working with superprocesses, one has already passed to the infinite limit with the population size so individual particles have infinitesimally small masses.

Consider \( \varepsilon > 0 \) (the branching parameter rate) and let \( X \) be a superprocess on a bounded Euclidean domain \( D \) corresponding to the semilinear elliptic operator \( Lu - \varepsilon u^2 \) on \( D \), where \( L \) is a second order elliptic operator with smooth coefficients and the underlying motion is the \( L \)-diffusion with killing at \( \partial D \). This process becomes extinct almost surely, but the so-called \( Q \)-process, denoted \( X^{Q, \varepsilon} \), is well defined for all time. To construct \( X^{Q, \varepsilon} \) one conditions \( X \) to survive until time \( t > 0 \) and then lets \( t \to \infty \). Next we “normalize” \( X^{Q, \varepsilon} \) as follows,

\[
X^{\varepsilon}_t := \frac{X^{Q, \varepsilon}_t}{\|X^{Q, \varepsilon}_t\|}.
\]

A problem analogous to the one considered in this paper is to decide if we can identify a unique immortal particle for \( X^{\varepsilon} \) and then describe its distribution as \( \varepsilon \to 0 \). To justify the analogy, assume that the intensity of the branching parameter \( \varepsilon \) is small. If one completely ignores the effect of branching inside the domain, the particle picture is simple: the (infinitesimal) particles are killed at the boundary \( \partial D \), and at the same time their mass is instantaneously re-distributed without spatial bias in \( D \), i.e., the birth/branching locations are uniformly chosen. Of course, the above picture is only approximate because \( \varepsilon \neq 0 \).

A result analogous to our Theorem 4.1 would be that \( X^{\varepsilon} \) has a unique immortal particle with law denoted by \( Q^\varepsilon \) and \( \lim_{\varepsilon \to 0} Q^\varepsilon = Q \), where \( Q \) is the law of a diffusion corresponding to the elliptic operator \( L^\phi \). This statement is indeed true. In fact, \( X^{Q, \varepsilon} \) has a unique immortal particle with the same law \( Q \) for every \( \varepsilon > 0 \) by the spine decomposition that can be found,
for example, in the recent preprint [23] and references therein. It is shown that the $Q$-process can be decomposed into an immortal “infinitesimal” particle (the “spine”) and a collection of “local bushes” which become extinct [23, Lemmas 4.11 and 4.17]. (See also [26] for proofs.) The spine corresponds to the $h$-transformed operator $L^\phi$, where $\phi$ is the positive eigenfunction (ground state) for the principal eigenvalue of $L$ on $D$ with zero Dirichlet boundary conditions. Finally, we note that the result holds for state spaces and operators much more general than described above; see again [23] and references therein.

8. Acknowledgments. We are grateful to Rodrigo Bañuelos for very useful advice. We thank the referees for very detailed reports and many suggestions for improvement.

REFERENCES

[1] BAÑUELOS, R. (1991). Intrinsic ultracontractivity and eigenfunction estimates for Schrödinger operators. J. Funct. Anal. 100 181–206.
[2] BIENIEK, M. and BURDZY, K. (2018). The distribution of the spine of a Fleming-Viot type process. Stochastic Process. Appl. 128 3751–3777.
[3] BIENIEK, M., BURDZY, K. and FINCH, S. (2012). Non-extinction of a Fleming-Viot particle model. Probab. Theory Related Fields 153 293–332.
[4] BIENIEK, M., BURDZY, K. and PAL, S. (2012). Extinction of Fleming-Viot-type particle systems with strong drift. Electron. J. Probab. 17 no. 11, 15.
[5] BURDZY, K., HOLYST, R. and MARCH, P. (2000). A Fleming-Viot particle representation of the Dirichlet Laplacian. Comm. Math. Phys. 214 679–703.
[6] BURDZY, K. and SALISBURY, T. S. (1999). On minimal parabolic functions and time-homogeneous parabolic $h$-transforms. Trans. Amer. Math. Soc. 351 3499–3531.
[7] BURDZY, K. and TADIĆ, T. (2021). On the spine of two-particle Fleming-Viot process driven by Brownian motion. Preprint. Math ArXiv, arXiv:2111.07968.
[8] BURKHOLDER, D. L. (1977). Exit times of Brownian motion, harmonic majorization, and Hardy spaces. Advances in Math. 26 182–205.
[9] CHAUMONT, L. and URIBE BRAVO, G. (2011). Markovian bridges: weak continuity and pathwise constructions. Ann. Probab. 39 609–647.
[10] DAVIES, E. B. and SIMON, B. (1984). Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. J. Funct. Anal. 59 335–395.
[11] DEPPERSCHMIDT, A., GREVEN, A. and PFAFHELHUBER, P. (2012). Tree-valued Fleming-Viot dynamics with mutation and selection. Ann. Appl. Probab. 22 2560–2615.
[12] DONNELLY, P. and KURTZ, T. G. (1996). A countable representation of the Fleming-Viot measure-valued diffusion. Ann. Probab. 24 698–742.
[13] DOOB, J. L. (1984). Classical potential theory and its probabilistic counterpart. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] 262. Springer-Verlag, New York.
[14] ENGŁANDER, J. and KYPRIANOU, A. E. (2004). Local extinction versus local exponential growth for spatial branching processes. Ann. Probab. 32 78–99.
[15] ETHERIDGE, A. M. (2000). An introduction to superprocesses. University Lecture Series 20. American Mathematical Society, Providence, RI.
[16] EVANS, S. N. (1993). Two representations of a conditioned superprocess. Proc. Roy. Soc. Edinburgh Sect. A 123 959–971.
[17] FABES, E. B., GAROFALO, N. and SALSA, S. (1986). A backward Harnack inequality and Fatou theorem for nonnegative solutions of parabolic equations. Illinois J. Math. 30 536–565.
[18] GREVEN, A., LIMIC, V. and WINTER, A. (2005). Representation theorems for interacting Moran models, interacting Fisher-Wright diffusions and applications. Electron. J. Probab. 10 no. 39, 1286–1356 (electronic).
[19] GREVEN, A., PFAFHELHUBER, P. and WINTER, A. (2013). Tree-valued resampling dynamics martingale problems and applications. Probab. Theory Related Fields 155 789–838.
[20] GRIGORESCU, I. and KANG, M. (2012). Immortal particle for a catalytic branching process. Probab. Theory Related Fields 153 333–361.
[21] HÉNARD, O. (2013). Change of measure in the lookdown particle system. Stochastic Process. Appl. 123 2054–2083.
[22] Jagers, P. (1975). Branching processes with biological applications. Wiley-Interscience [John Wiley & Sons], London-New York-Sydney.

[23] Liu, R., Ren, Y.-X., Song, R. and Sun, Z. (2021). Subcritical superprocesses conditioned on non-extinction. Preprint. Math arXiv:2112.15184v2.

[24] Moran, P. A. P. (1958). Random processes in genetics. Proc. Cambridge Philos. Soc. 54 60–71.

[25] Pinsky, R. G. (1985). On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. Ann. Probab. 13 363–378.

[26] Ren, Y.-X., Song, R. and Sun, Z. (2020). Spine decompositions and limit theorems for a class of critical superprocesses. Acta Appl. Math. 165 91–131.

[27] Revuz, D. and Yor, M. (1999). Continuous martingales and Brownian motion, third ed. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] 293. Springer-Verlag, Berlin.

[28] Seidel, P. (2014). The historical process of the spatial Moran model with selection and mutation, PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

[29] Tough, O. (2021). Scaling Limit of the Fleming-Viot Multi-Colour Process. Preprint. Math ArXiv. arXiv:2110.05049.

[30] Villemonais, D. (2014). General approximation method for the distribution of Markov processes conditioned not to be killed. ESAIM: Probability and Statistics 18 441–467.