Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 48, 1–38.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP302

Invariance principle for non-homogeneous random walks

Nicholas Georgiou* Aleksandar Mijatović† Andrew R. Wade*

Abstract

We prove an invariance principle for a class of zero-drift spatially non-homogeneous random walks in $\mathbb{R}^d$, which may be recurrent in any dimension. The limit $\mathcal{X}$ is an elliptic martingale diffusion, which may be point-recurrent at the origin for any $d \geq 2$. To characterize $\mathcal{X}$, we introduce a (non-Euclidean) Riemannian metric on the unit sphere in $\mathbb{R}^d$ and use it to express a related spherical diffusion as a Brownian motion with drift. This representation allows us to establish the skew-product decomposition of the excursions of $\mathcal{X}$ and thus develop the excursion theory of $\mathcal{X}$ without appealing to the strong Markov property. This leads to the uniqueness in law of the stochastic differential equation for $\mathcal{X}$ in $\mathbb{R}^d$, whose coefficients are discontinuous at the origin. Using the Riemannian metric we can also detect whether the angular component of the excursions of $\mathcal{X}$ is time-reversible. If so, the excursions of $\mathcal{X}$ in $\mathbb{R}^d$ generalize the classical Pitman–Yor splitting-at-the-maximum property of Bessel excursions.

Keywords: non-homogeneous random walk; invariance principle; diffusion limits; excursions; skew product; rapid spinning; recurrence; transience.
AMS MSC 2010: Primary 60J05; 60J60, Secondary 60F17; 58J65; 60J55.
Submitted to EJP on April 23, 2018, final version accepted on March 27, 2019.
Supersedes arXiv:1801.07882.

1 Introduction

A large class of spatially non-homogeneous zero-drift random walks (Markov chains) on $\mathbb{R}^d$ ($d \geq 2$) was introduced in [9], where it was shown that such a walk may be transient or recurrent in any dimension $d \geq 2$. These walks are martingales with uniformly non-degenerate increments (see assumptions (A1)–(A2) below). A non-homogeneous random walk with zero drift and fixed covariance matrix exhibits the classical dichotomy of Pólya's theorem: recurrence for $d = 1, 2$ and transience for $d \geq 3$ [20, Thm 4.1.3]. The anomalous recurrence behaviour of the walks in [9] is achieved by varying the limiting increment

*Department of Mathematical Sciences, Durham University, UK.
E-mail: nicholas.georgiou@durham.ac.uk,andrew.wade@durham.ac.uk
†Department of Statistics, University of Warwick, & The Alan Turing Institute, UK.
E-mail: a.mijatovic@warwick.ac.uk

covariance matrix, as described by a matrix-valued function $\sigma^2 : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ on the unit sphere $\mathbb{S}^{d-1}$ in $\mathbb{R}^d$ (see assumptions (A3)–(A4) below). This is a genuinely many-dimensional phenomenon: essentially, a one-dimensional walk whose increments have zero mean and two moments is always recurrent [20, Thm 4.1.2].

An important element of the classical theory of spatially homogeneous random walks is the Donsker invariance principle, which describes the *scaling limit* for the class of homogeneous random walks whose increments have zero mean and positive-definite covariance in terms of Brownian motion (BM) on $\mathbb{R}^d$. This paper studies scaling limits of the non-homogeneous random walks introduced in [9]. The assumptions (A0)–(A6) that we impose are described formally in Section 2, along with some examples; we make some remarks on the motivation behind these assumptions at the start of Section 1.1. Under these assumptions, we prove that under diffusive scaling, the random walk converges weakly to a diffusion process $\mathcal{X} = (\mathcal{X}_t, t \in \mathbb{R}_+)$ whose law is determined uniquely by σ^2 via the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{X}_t = \sigma(\hat{\mathcal{X}}_t)\mathrm{d}W_t, \qquad \mathcal{X}_0 = \mathbf{x}_0 \in \mathbb{R}^d. \tag{1.1}$$

Here $\hat{\mathbf{x}}$ is the radial projection onto $\mathbb{S}^{d-1}$ of any $\mathbf{x} \in \mathbb{R}^d$ (with an arbitrary choice $\hat{\mathbf{0}} \in \mathbb{S}^{d-1}$ for the origin $\mathbf{0}$), $(W_t, t \geq 0)$ denotes a standard BM on $\mathbb{R}^d$, $\sigma : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ is a *square root* of σ^2 (i.e., $\sigma(\mathbf{u})\sigma^\top(\mathbf{u}) = \sigma^2(\mathbf{u})$ for all $\mathbf{u} \in \mathbb{S}^{d-1}$) and $\mathbf{x}_0$ a non-random point. Our first main result says that the SDE (1.1) characterizes uniquely in law a continuous strong Markov process (a diffusion), which will serve as the limit in our invariance principle.

Theorem 1.1. *Let the positive-definite symmetric matrix-valued function* $\sigma^2 : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ *satisfy* (A4)–(A6) *below. Then, for any starting point* $\mathcal{X}_0 = \mathbf{x}_0$ *in* $\mathbb{R}^d$*, weak existence and uniqueness in law hold for SDE* (1.1) *and the strong Markov property is satisfied. Moreover, the law of* $\mathcal{X}$ *does not depend on the choices of the square-root* σ *and* $\hat{\mathbf{0}} \in \mathbb{S}^{d-1}$.

The process $\mathcal{X}$ possesses certain universal properties, in some aspects resembling those of a BM on $\mathbb{R}^d$ (in the special case where σ^2 is the identity, it is BM). In particular, the norm process $\|\mathcal{X}\|$ is a constant multiple of a Bessel process of 'dimension' (parameter) $V/U > 1$, where U and V describe via (A6) the stability properties of σ^2. The key difference to the case of BM is that, due to the possible recurrence of the random walk in any dimension $d \geq 2$, the scaling limit $\mathcal{X}$ may visit the origin infinitely often (when $V/U < 2$). Since the diffusion coefficient is discontinuous at $\mathbf{0}$, the proof of the uniqueness in law requires the development of the excursion theory of $\mathcal{X}$ before the strong Markov property can be established. This step constitutes the main technical contribution of the paper (see Section 3.6 below) and provides an insight into the structure of the excursions of $\mathcal{X}$. The backbone of the excursions is provided by the excursions of the radial (Bessel) component, and the full excursion description rests on the introduction of a (non-Euclidean) Riemannian metric on $\mathbb{S}^{d-1}$ (Section 3.3 below), yielding a skew-product decomposition of the excursions of $\mathcal{X}$, which in turn entails a generalization of Stroock's representation of spherical BM [11, p. 83] (see (1.3) below). The new geometry on the sphere also yields a multi-dimensional generalization of the splitting-at-the-maximum property of Bessel excursions [23]. Furthermore, the choice of the square root of σ^2 turns out to be relevant for the pathwise uniqueness of SDE (1.1), which may fail, thus generalizing to higher dimensions the example of Stroock and Yor [27] for complex BM. These and other features of the law of $\mathcal{X}$ are described in more detail in Section 1.1 below. The proof of Theorem 1.1 is in Section 3 with an overview in Section 3.1.

Having characterized the scaling limit, we state our invariance principle. For a discrete-time process $X = (X_m, m \in \mathbb{Z}_+)$, any $n \in \mathbb{N}$ and $t \in \mathbb{R}_+$, define $\lfloor nt \rfloor := \max\{k \in$

$\mathbb{Z}_+ : k \le nt\}$ and

$$\widetilde{X}_n(t) := n^{-1/2} X_{\lfloor nt \rfloor}. \tag{1.2}$$

The paths of $\widetilde{X}_n = (\widetilde{X}_n(t), t \in \mathbb{R}_+)$ are in the Skorohod space $\mathcal{D}_d = \mathcal{D}(\mathbb{R}_+; \mathbb{R}^d)$ of right-continuous functions with left limits, endowed with the Skorohod metric (see e.g. [8, §3.5]).

Theorem 1.2. *Let* (A0)–(A6) *below hold for the random walk* X*. Let* $\mathcal{X}$ *be the unique (weak) solution of* (1.1) *with* $\mathcal{X}_0 = \mathbf{0}$*. Then, as* $n \uparrow \infty$*, the weak convergence* $\widetilde{X}_n \Rightarrow \mathcal{X}$ *on* $\mathcal{D}_d$ *holds.*

The class of random walks satisfying (A0)–(A6) consists of $\mathbb{R}^d$-valued Markov chains with an asymptotically stable increment covariance structure: see Section 2 for some examples. Thus Theorem 1.2 may be viewed as a multi-dimensional generalization of the classical invariance principle of Lamperti [19] for $\mathbb{R}_+$-valued Markov chains with asymptotically constant variance of the increments. The proof of Theorem 1.2 hinges on the radial invariance principle in [10], where it was shown that the process of norms of the walk converges weakly to a Bessel process of dimension V/U, and a d-dimensional invariance principle for martingale diffusions with discontinuous coefficients given in Theorem 4.1 below. Invariance principles with continuous coefficients, such as [8, Thm 7.4.1, p. 354], do not apply in our setting (both formally and) because, by Corollary 3.24 below, the process $\mathcal{X}$ may hit the discontinuity point $\mathbf{0}$ infinitely many times. In order to deal with the point-recurrence of $\mathcal{X}$, it is necessary to control the amount of time $\mathcal{X}$ spends near $\mathbf{0}$. This is achieved via the occupation times formula and the analysis of the local time of the radial component of $\mathcal{X}$ (see proof of Lemma 4.10 below). Note that neither the specific form of the law of the radial component nor the fact that $\mathcal{X}$ has no drift are crucial for the validity of Theorem 4.1. Some consequences of Theorem 1.2 for random walks are in Section 1.2 below. Its proof is in Section 4 below.

1.1 The diffusion limit

As described in [9], the recurrence/transience of our non-homogeneous random walks is determined by the interplay between the radial and transverse components of the variance of the increments. It is thus natural to assume some stability for these components of $\sigma^2 : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ (see (A4) below): namely, we require constant total $\operatorname{tr} \sigma^2(\mathbf{u}) = V$ and radial $\mathbf{u}^\top \sigma^2(\mathbf{u}) \mathbf{u} = U$ instantaneous variances for all $\mathbf{u} \in \mathbb{S}^{d-1}$ and some positive reals $U < V$. This ensures that the radial component of the process has a Bessel limit [10]. Further assumptions on σ^2 in Theorem 1.1 are smoothness (A5), which ensures that the angular part of the limit can be described in terms of a diffusion on the sphere, and a structural condition $\sigma^2(\mathbf{u})\mathbf{u} = U\mathbf{u}$ for all $\mathbf{u} \in \mathbb{S}^{d-1}$ ((A6) below), which ensures the existence of a skew-product decomposition of the excursions of $\mathcal{X}$.

$\mathcal{X}$ is a self-similar Markov process on $\mathbb{R}^d$ (with Brownian scaling)

The process $\|\mathcal{X}\|/\sqrt{U}$ is Bessel of dimension $V/U > 1$ (see Lemma 3.2 below). Hence, if $V/U \in (1, 2]$ (resp. $V/U > 2$), then $\liminf_{t\to\infty} \|\mathcal{X}_t\| = 0$ (resp. $\lim_{t\to\infty} \|\mathcal{X}_t\| = \infty$) and the origin $\mathbf{0}$ is recurrent for $\mathcal{X}$ if and only if $V/U < 2$. (The Foster–Lyapunov criteria [22, Thm 6.2.1] do not apply, even if Theorem 1.1 has been established, since $\mathbf{x} \mapsto \sigma^2(\hat{\mathbf{x}})$ is discontinuous.) Let $\mathbb{P}_{\mathbf{x}_0}$ be the law of $\mathcal{X}$ started at $\mathcal{X}_0 = \mathbf{x}_0 \in \mathbb{R}^d$. Define $\mathcal{Y} = (\mathcal{Y}_t, t \ge 0)$ by $\mathcal{Y}_t := c\mathcal{X}_{c^{-2}t}$, for some constant $c > 0$. Then the scale invariance of $\mathbf{x} \mapsto \sigma(\hat{\mathbf{x}})$ and W in (1.1) imply that $\mathcal{Y}$ solves SDE (1.1) with $\mathcal{Y}_0 = c\mathbf{x}_0$. By Theorem 1.1, the law of $\mathcal{Y}$ equals $\mathbb{P}_{c\mathbf{x}_0}$, making $\mathcal{X}$ a globally defined self-similar Markov process on $\mathbb{R}^d$, which may hit $\mathbf{0}$ infinitely many times.

A stationary diffusion ψ on $\mathbb{S}^{d-1}$

Consider the following Stratonovich SDE on $\mathbb{S}^{d-1}$,

$$\mathrm{d}\phi_t = (\sigma_{\mathrm{sy}}(\phi_t) - \phi_t\phi_t^\top) \circ \mathrm{d}W_t - (I - \phi_t\phi_t^\top)A_0(\phi_t)\mathrm{d}t, \tag{1.3}$$

where I is the d by d identity matrix, W is a standard BM on $\mathbb{R}^d$, σ_{sy} is the unique symmetric square root of σ^2, which is hence smooth by Lemma 3.1 below, and the vector field A_0 is a linear combination of the derivatives of the columns of σ_{sy} defined in Section 3.4 below. By Lemma 3.6 below, SDE (1.3) has a unique strong solution over $t \in \mathbb{R}_+$ for any starting point on $\mathbb{S}^{d-1}$. We reserve ϕ for such a solution of (1.3) indexed by $\mathbb{R}_+$. We also describe how to construct a *stationary* solution ψ of (1.3) indexed by $\mathbb{R}$.

In the case $\sigma^2 = \sigma_{\mathrm{sy}} = I$, SDE (1.3) clearly reduces to Stroock's representation of the BM on $\mathbb{S}^{d-1}$ with the Riemannian metric induced by the ambient Euclidean space [11, p. 83] ($\mathcal{X}$ in this case is a BM on $\mathbb{R}^d$).

The key ingredient of the excursion measure of $\mathcal{X}$ is the stationary distribution μ on $\mathbb{S}^{d-1}$ of the solution ϕ of (1.3). In order to analyse ϕ and characterize μ, it turns out to be essential to modify the geometry on $\mathbb{S}^{d-1}$ via the Riemannian metric $g_{\mathbf{x}}(v_1, v_2) := \langle \sigma^{-2}(\mathbf{x})v_1, v_2\rangle$, where $\mathbf{x} \in \mathbb{S}^{d-1}$, $v_1, v_2 \in \mathbb{R}^d$ are in the tangent space of $\mathbb{S}^{d-1}$ at $\mathbf{x}$ and $\langle\,\cdot\,,\cdot\,\rangle$ is the inner product on $\mathbb{R}^d$. On the Riemannian manifold $(\mathbb{S}^{d-1}, g)$, by Lemma 3.6, ϕ is a BM with drift, generated by $\mathcal{G} = (1/2)\Delta_g + V_0$, where Δ_g is the Laplace–Beltrami operator and V_0 is a tangential vector field on $\mathbb{S}^{d-1}$, explicit in σ^2 and its derivatives of order one. Prop. 3.7 states that the stationary measure μ is unique. We can use the stationary measure μ to define a stationary solution ψ of (1.3), indexed by $\mathbb{R}$, with law $\mathbb{P}_\Psi$ (see Prop. 3.7 below), namely, ψ_t has law μ for any $t \in \mathbb{R}$ and the evolution of $(\psi_s, s \geq t)$ is determined by (1.3).

The proof of Prop. 3.7 shows that in fact $\mu(\mathrm{d}\mathbf{x}) = p(\mathbf{x})\mathrm{d}_g\mathbf{x}$, where $p : \mathbb{S}^{d-1} \to \mathbb{R}$ is a strictly positive density with respect to the Riemannian volume element $\mathrm{d}_g\mathbf{x}$ on $(\mathbb{S}^{d-1}, g)$ (see e.g. [12, p. 291] for a definition), uniquely determined by the PDE $\mathcal{G}^*p = 0$ with $\mathcal{G}^*$ denoting the adjoint of $\mathcal{G}$ on $L^2(\mathbb{S}^{d-1}; \mathrm{d}_g\mathbf{x})$. Recall that for any vector field V on $\mathbb{S}^{d-1}$, $\operatorname{div} V$ is the trace of the endomorphism of the tangent space given by the directional derivatives of V via the Levi-Civita connection and, for any smooth f on $\mathbb{S}^{d-1}$, we have $\Delta_g f = \operatorname{div}(\operatorname{grad}(f))$ (see Section 3.3 below). Integration by parts implies that p is the unique positive solution of the PDE

$$\frac{1}{2}\Delta_g p - \operatorname{div}(pV_0) = 0, \qquad \text{satisfying} \int_{\mathbb{S}^{d-1}} p(\mathbf{x})\mathrm{d}_g\mathbf{x} = 1. \tag{1.4}$$

In the case that $V_0 = \operatorname{grad} F_0$ for a smooth $F_0 : \mathbb{S}^{d-1} \to \mathbb{R}$, the definition of $\operatorname{grad} F_0$ on $(\mathbb{S}^{d-1}, g)$ in Section 3.3 below implies that $p := \exp(2F_0)/\int_{\mathbb{S}^{d-1}} \exp(2F_0(\mathbf{x}))\mathrm{d}_g\mathbf{x}$ is the unique solution of (1.4). Moreover, by [15, Thms 4.2 & 6.1], SDE (1.3) is time reversible: for any random time $T \in \mathbb{R}$, independent of ψ, the process $(\psi_{T-t}, t \in \mathbb{R}_+)$ solves (1.3) started according to the law μ. In particular, if $F_0 \equiv 0$, then ψ is the stationary spherical BM on $(\mathbb{S}^{d-1}, g)$ and the measure μ is uniform with respect to the volume element.

Point-transient case: skew-product decomposition of $\mathcal{X}$

Suppose that $2 \leq V/U$. If $\mathcal{X}_0 \neq \mathbf{0}$, a Bessel process $r/\sqrt{U}$ of dimension V/U (with $r_0 = \|\mathcal{X}_0\|$) is strictly positive and we may define $\rho_s(t) = \int_s^t r_u^{-2}\mathrm{d}u$ for $t, s \geq 0$. Then the process $(r_t\phi_{\rho_0(t)}, t \in \mathbb{R}_+)$, where the solution ϕ of SDE (1.3), started at $\phi_0 = \hat{\mathcal{X}}_0$, and r are independent, has the same law as $\mathcal{X}$ (see Section 3.5).

The relevant case for Theorem 1.2 is $\mathcal{X}_0 = \mathbf{0}$. As $\mathcal{X}$ starts from $\mathbf{0}$ and never returns, a natural description of its law is via a family of entrance laws at positive times s and the subsequent evolution. The latter is given in terms of a Bessel process and a time-changed angular process solving (1.3) as above: $(r_t\phi_{\rho_s(t)}, t \geq s)$ with $\phi_0 := \hat{\mathcal{X}}_s$. The random vector $\hat{\mathcal{X}}_s$ is forced to be *independent* of r_s and distributed according to the *stationary* law μ of ϕ, due to the *rapid spinning* of the process $\mathcal{X}$ as it leaves $\mathbf{0}$: $\rho_s(t) \to \infty$ as $s \downarrow 0$ for fixed $t > 0$ (see Lemma 3.12 below). As $\rho_s(t) = \rho_s(1) + \rho_1(t)$ for any $s, t > 0$, the processes $(r_t\psi_{\rho_1(t)}, t > 0)$ and $(\mathcal{X}_t, t > 0)$ are equal in law, where ψ (the stationary solution to (1.3)) and r are independent. The analogy with the classical case of the skew product of BM on $\mathbb{R}^d$ in both cases $\mathcal{X}_0 \neq \mathbf{0}$ and $\mathcal{X}_0 = \mathbf{0}$ (see [25, §IV.35, p. 73] and [13, p. 276]) is clear.

Point-recurrent case: skew-product decomposition of excursions of $\mathcal{X}$

Assume $V/U \in (1, 2)$ and $\mathcal{X}_0 = \mathbf{0}$. The process $\mathcal{X}$ returns to $\mathbf{0}$ infinitely often since $\|\mathcal{X}\|/\sqrt{U}$ is Bessel of dimension V/U. As the excursions of $\mathcal{X}$ turn out to exhibit the rapid spinning behaviour at each end, its excursion measure may be constructed as follows. Mark each Bessel excursion by an independent draw ψ from the law $\mathbb{P}_\Psi$ on $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$, the stationary solution to (1.3) given in Prop. 3.7 below. Since, due to rapid spinning at the beginning of each excursion of $\mathcal{X}$, the angular component of the excursion is distributed according to the stationary measure μ of SDE (1.3) at all times, we need to map the marked Bessel excursion by time-changing the mark ψ via an additive functional of the Bessel excursion, see Section 3.6.1 below for details. Note that the mapping has to be defined for Bessel excursions lasting longer than a (for any fixed $a > 0$), since the time-change can only be "anchored" at a pre-specified time during the life time of the excursion. Although this causes some technical difficulties, the mapped Poisson point processes can be interpreted consistently (for all $a > 0$). Its excursion measure turns out to be that of $\mathcal{X}$.

We stress that this construction of the excursion measure depends only on σ^2, which specifies the dimension of the Bessel process and hence its excursion measure and determines the marks via SDE (1.3) (the mapping uses only the information contained in the Bessel excursion). Moreover, the local time at $\mathbf{0}$ of $\mathcal{X}$ can be defined as that of $\|\mathcal{X}\|$ at 0, without a reference to the strong Markov property of $\mathcal{X}$. Hence, once the excursion measure has been constructed (Section 3.6.1 below), the key step in the proof of Theorem 1.1 consists of establishing that (without the strong Markov property) the point process of excursions of $\mathcal{X}$ is the Poisson point process with the excursion measure described above. The details are in Section 3.6.2 below.

In the case $\mathcal{X}_0 \neq \mathbf{0}$, up to the first hitting time of $\mathbf{0}$, the skew product of excursions coincides with the generalized Lamperti representation for self-similar Markov processes on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ [1], where the Lévy process is a scalar BM with drift and the angular component equals the diffusion on $\mathbb{S}^{d-1}$ in (1.3) started at $\hat{\mathcal{X}}_0$. Note also that there is a literature (see e.g. [28] and the reference therein) on the extensions of strong Markov processes on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ with skew-product decomposition beyond the first hitting time of the origin, of which $\mathcal{X}$ is an example.

Splitting excursions at the maximum: a generalized Pitman–Yor representation

If the vector field V_0 in (1.4) has a potential, the excursions of $\mathcal{X}$ provide a multi-dimensional generalization of the famous Pitman–Yor [23] representation of the Bessel excursions with dimension $\delta = V/U \in (1, 2)$. Let $U = 1$ and recall from [23] that the unique maximum M of the Bessel excursion e^r is drawn from the σ-finite density $m \mapsto m^{\delta-3}$ on the interval $(0, \infty)$. Then, conditional on M, the excursion e^r is obtained by joining back to back two independent Bessel processes β and β' of dimension $4 - \delta$, both started at 0 and run until the first times (T_M and T'_M respectively) they hit M: $e^r(t) =$

$\mathbf{1}\{t \in (0, T_M]\}\beta_t + \mathbf{1}\{t \in (T_M, T_M + T'_M)\}\beta'_{T_M+T'_M-t}$. A trivial (but crucial) observation is that when the maximum is reached, the process is neither at the beginning nor the end of the excursion. Hence, due to rapid spinning, the angular component $\hat{e}^{\mathcal{X}}(T_M)$ of the corresponding excursion $e^{\mathcal{X}}$ of $\mathcal{X}$ at T_M must follow the stationary law μ of SDE (1.3). As SDE (1.3) is time-reversible (see the paragraph after (1.4) above), the excursion $e^{\mathcal{X}}$ equals

$$e^{\mathcal{X}}(t) = \mathbf{1}\{t \in (0, T_M]\}\beta_t \phi_{\rho(T_M-t)} + \mathbf{1}\{t \in (T_M, T_M + T'_M)\}\beta'_{T_M+T'_M-t}\phi'_{\rho'(t-T_M)}, \qquad (1.5)$$

where ϕ, ϕ' are solutions of SDE (1.3) with the same initial condition $\phi_0 = \phi'_0$, distributed according to μ, and driven by independent BMs. The time-changes $\rho(t) = \int_0^t \beta_{T_M-s}^{-2}\mathrm{d}s$, $t \in (0, T_M]$, and $\rho'(t) = \int_0^t \beta'^{-2}_{T'_M-s}\mathrm{d}s$, $t \in [0, T'_M)$, satisfy $\lim_{t\downarrow 0}\rho(T_M - t) = \lim_{t\uparrow T'_M}\rho'(t) = \infty$.

In the limit as $U \uparrow V$, which is excluded from our results, the angular motion degenerates to a constant as the trace of σ^2 equals the radial eigenvalue. The radial part becomes the modulus of the scalar BM, while rapid spinning and (1.5) suggest that the singular diffusion in the limit changes the ray it lives on every time it hits the origin according to a law on $\mathbb{S}^{d-1}$, which is the limit of the stationary measures of SDE (1.3) as $V/U \downarrow 1$. It hence appears that the limiting singular diffusion is a generalization of the Walsh BM (or Brownian spider) [2] to $\mathbb{R}^d$.

Smooth square roots and pathwise uniqueness: the Stroock–Yor phenomenon

SDE (1.1) need not (but clearly could) possess pathwise uniqueness even if σ^2 is the identity (consider $\sigma(\mathbf{u}) = \operatorname{diag}(\operatorname{sgn}(u_1), \ldots, \operatorname{sgn}(u_d))$ and recall the scalar Tanaka SDE [24, §IX.1, Ex. (1.19)]). This behaviour persists even for smooth square roots σ. Below we give a generalization of the SDE for complex Brownian motion in [27, Thm 3.12], with the property that the failure of pathwise uniqueness occurs precisely when the solution starts from (or visits) $\mathbf{0}$.

Note first that a simple application of the occupation times formula and the fact that $\mathcal{X}_t = \mathbf{0}$ if and only if $\|\mathcal{X}_t\| = 0$ imply that if $\mathcal{X}$ solves SDE (1.1) for a given choice of $\hat{\mathbf{0}}$, then it also solves the SDE for any other choice $\hat{\mathbf{0}} \in \mathbb{S}^{d-1}$. If a square root σ satisfies (I) $P\sigma(\mathbf{u}) = \sigma(P\mathbf{u})$ for all $\mathbf{u} \in \mathbb{S}^{d-1}$, where $P \in SO(d) \setminus \{I\}$ [1], then Itô's formula and the remark above imply that for any solution $(\mathcal{X}, W)$ of (1.1) started from $\mathbf{0}$, the process $(\mathcal{Y}, W)$, where $\mathcal{Y} := P\mathcal{X}$, is also a solution. By Theorem 1.1, $\mathcal{X}$ and $\mathcal{Y}$ have the same law but are clearly not equal. If, in addition, σ satisfies (II) $\mathbf{u} = \sigma(\mathbf{u})\mathbf{c}$ for all $\mathbf{u} \in \mathbb{S}^{d-1}$ and some $\mathbf{c} \in \mathbb{S}^{d-1}$, the BM driving the process $\|\mathcal{X}\|$ equals $\mathbf{c}^\top W$ (Lemma 3.2 below), making $\|\mathcal{X}\|$ adapted to W. Moreover, assuming $\mathcal{X}$ never visits $\mathbf{0}$, the BM driving the angular component via SDE (1.3) is a time-change of $\int_0^\cdot \|\mathcal{X}_s\|^{-1}\mathrm{d}W_s$ (see (3.15) and Prop. 3.11 below). Hence the skew product $\|\mathcal{X}_t\|\phi_{\rho_0(t)}$, $t \in \mathbb{R}_+$, where $\rho_0(t) = \int_0^t \|\mathcal{X}_u\|^{-2}\mathrm{d}u$, makes $\mathcal{X}$ a strong solution of (1.1).

It remains to exhibit a smooth σ satisfying (I) and (II) above. Note first that (I) may only hold in even dimensions. We rely on the Lie group structure of the spheres in dimensions $d \in \{2, 4\}$ for our examples. Pick a positive-definite $A \in \mathbb{R}^d \otimes \mathbb{R}^d$ and let $\sigma(\mathbf{u}) = R(\mathbf{u})A$, where $R : \mathbb{S}^{d-1} \to SO(d)$ is smooth. For $d = 4$, view $\mathbb{S}^3$ as unit quaternions and define R by $R(\mathbf{u})\mathbf{v} := \mathbf{u} \bullet \mathbf{v}$, where $\mathbf{u} \bullet \mathbf{v}$ denotes the multiplication of quaternions $\mathbf{v} \in \mathbb{R}^4$ and $\mathbf{u}$ (see e.g. [25, p. 229]). It is easy to check that $R(\mathbf{u}) \in SO(4)$ and $R(\mathbf{u})\mathbf{e}_1 = \mathbf{u}$ for all $\mathbf{u} \in \mathbb{S}^3$, where $\mathbf{e}_1$ is the first standard basis element of $\mathbb{R}^4$, i.e. the real unit quaternion. If $A\mathbf{e}_1 = \mathbf{e}_1$ (as is the case if (A6) holds), then (II) holds. Moreover, $\sigma(\mathbf{u})$ is a smooth square root of $\sigma^2(\mathbf{u}) = R(\mathbf{u})A^2R(\mathbf{u})^{-1}$. Pick a unit quaternion $\mathbf{p} \in \mathbb{S}^3 \setminus \{\mathbf{e}_1\}$

[1] $SO(d)$ is the group of orientation-preserving orthogonal matrices in $\mathbb{R}^d \otimes \mathbb{R}^d$ and I is the identity matrix.

and define $P := R(\mathbf{p}) \in SO(4)$. The associativity of the product $\bullet$ yields the matrix identity $PR(\mathbf{u}) = R(P\mathbf{u})$ for $\mathbf{u} \in \mathbb{S}^3$, implying (I). Hence pathwise uniqueness fails when $\mathcal{X}_0 = \mathbf{0}$. Since $\sigma^2(\mathbf{u})\mathbf{u} = \mathbf{u}$, the process $\mathcal{X}$ hits $\mathbf{0}$ if and only if $\operatorname{tr}(\sigma^2(\mathbf{u})) = \operatorname{tr}(A^2) \in (1,2)$ and we may choose independently a different rotation P for each excursion, exhibiting uncountably many solutions of (1.1) for a fixed BM W. The complex case is analogous: a BM in [27, Thm 3.12] solves (1.1) with $\sigma(\mathbf{u}) = R(\mathbf{u})$ a multiplication by $\mathbf{u} \in \mathbb{S}^1$.

1.2 Angular convergence and the first exit from large balls of the random walk

We now describe the behaviour of the angular component of the random walk X and its asymptotic law at $\tau_a^n := \inf\{m \in \mathbb{Z}_+ : \|X_m\| \geq a\sqrt{n}\}$ its first exit out of the ball centred at $\mathbf{0}$ with radius $a\sqrt{n}$ (for some $a > 0$). Both statements are easy consequences of Theorem 1.2.

Let r be a Bessel process of dimension $\delta > 1$, $r_0 = 0$, and $\tau_a := \inf\{t \in \mathbb{R}_+ : r_t = a\}$ (thus $\tau_a < \infty$ a.s). Recall that $\mathbb{P}[r_1 \leq x] = \int_0^{x^2/2} z^{\alpha-1}\mathrm{e}^{-z}\mathrm{d}z/\Gamma(\delta/2)$ for all $x \in \mathbb{R}_+$ [24, Cor. XI.1.4], where Γ denotes the gamma function, and $\mathbb{E}[\exp(-\lambda\tau_a)] = (a\sqrt{2\lambda})^\nu/(2^\nu\Gamma(\nu+1)I_\nu(a\sqrt{2\lambda}))$, for any $\lambda > 0$, where I_ν denotes the modified Bessel function of the first kind of order $\nu := (\delta-2)/2$ (see [16] for a series expansion of the density of τ_a in terms of the zeros of Bessel functions).

Corollary 1.3. *Let the random walk X satisfy the assumptions of Theorem 1.2 with $U = 1$ and define $\delta := V$. Let the random vector θ with the law μ on $\mathbb{S}^{d-1}$, whose density satisfies* (1.4)*, be independent of r. Then, as $n \to \infty$, the following weak limits hold:*

$$n^{-1/2}X_n \Rightarrow r_1\theta \quad \textit{(and hence } \hat{X}_n \Rightarrow \theta\textit{)} \qquad \textit{and} \qquad (\tau_a^n/n, n^{-1/2}X_{\tau_a^n}) \Rightarrow (\tau_a, a\theta).$$

For a continuous $f : \mathbb{S}^{d-1} \to \mathbb{R}$, Cor. 1.3 and [4, Thm 2.1] imply $\lim_{n\uparrow\infty} \mathbb{E}[f(\hat{X}_n)] = \int_{\mathbb{S}^{d-1}} f\mathrm{d}\mu$. However, the ergodic average $\frac{1}{n}\sum_{k=0}^{n-1} f(\hat{X}_k)$ cannot in general converge in probability to the constant $\int_{\mathbb{S}^{d-1}} f\mathrm{d}\mu$, since by Theorem 1.2, an analogous argument to the one in the proof of Lemma 4.10 below and (1.2), the average converges weakly to a non-degenerate limit (for a non-constant function f): $\frac{1}{n}\sum_{k=0}^{n-1} f(\hat{X}_k) \Rightarrow \int_0^1 f(\hat{\mathcal{X}}_t)\mathrm{d}t$.

Proof of Corollary 1.3. By (1.2) and Theorem 1.2, $n^{-1/2}X_n = \widetilde{X}_n(1) \Rightarrow \mathcal{X}_1$. Since $\mathcal{X}_0 = \mathbf{0}$, the skew product structure (Lem. 3.12 (polar case) and Prop. 3.21 (point-recurrent case)) yields the first limit. The mapping theorem [4, Thm 5.1] implies the second ($\mathbf{x} \mapsto \hat{\mathbf{x}}$ is continuous on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ and $\mathbb{P}[\mathcal{X}_1 = \mathbf{0}] = 0$). Note that $\tau_a^n = \tau^a(\widetilde{X}_n)$ and $\tau_a = \tau^a(r)$, where $\tau^a(x)$, $x \in \mathcal{D}_d$, is defined in (4.9). As r reaches new maxima immediately after τ_a, $\lim_{b\to a}\tau^b(r) = \tau^a(r)$ holds a.s. By Lemma 4.7, Remark (a) just after it, Theorem 1.2 and [4, Thm 5.1] the final limit holds. □

2 Assumptions and examples

Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$ be the standard orthonormal basis in $\mathbb{R}^d$ ($d \geq 2$) with respect to the Euclidean inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^d$, and $\mathbb{S}^{d-1} := \{\mathbf{u} \in \mathbb{R}^d : \|\mathbf{u}\| = 1\}$ the unit sphere in $\mathbb{R}^d$, where $\|\cdot\|$ is the Euclidean norm. For $\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and the origin $\mathbf{0}$, let $\hat{\mathbf{x}} := \mathbf{x}/\|\mathbf{x}\|$ and $\hat{\mathbf{0}} := \mathbf{e}_1$, respectively.

Let $X = (X_n, n \in \mathbb{Z}_+)$ be a discrete-time, time-homogeneous Markov process on an unbounded Borel subset $\mathbb{X}$ of $\mathbb{R}^d$. Suppose X_0 is a non-random point in $\mathbb{X}$. Denote the increments of X by $\Delta_n := X_{n+1} - X_n$. Since the law of Δ_n depends only on X_n, we often take $n = 0$ and write Δ for Δ_0. Let $\mathbb{P}_\mathbf{x}[\,\cdot\,] = \mathbb{P}[\,\cdot \mid X_0 = \mathbf{x}]$ and $\mathbb{E}_\mathbf{x}[\,\cdot\,] = \mathbb{E}[\,\cdot \mid X_0 = \mathbf{x}]$ denote the probabilities and expectations when the walk is started from $\mathbf{x} \in \mathbb{X}$. We make the following assumptions.

(A0) Suppose that $\sup_{\mathbf{x}\in\mathbb{X}} \mathbb{E}_\mathbf{x}[\|\Delta\|^4] < \infty$.

By (A0), the mean $\mu(\mathbf{x}) := \mathbb{E}_{\mathbf{x}}[\Delta]$ and the covariance matrix $M(\mathbf{x}) := \mathbb{E}_{\mathbf{x}}[\Delta\Delta^\top]$ exist $\forall \mathbf{x} \in \mathbb{X}$.

(A1) Suppose that $\mu(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{X}$.

The next assumption ensures that Δ is *uniformly non-degenerate*.

(A2) There exists $v > 0$ such that $\operatorname{tr} M(\mathbf{x}) = \mathbb{E}_{\mathbf{x}}[\|\Delta\|^2] \geq v$ for all $\mathbf{x} \in \mathbb{X}$.

For a matrix $M \in \mathbb{R}^d \otimes \mathbb{R}^d$ define the norm $\|M\| := \sup_{\mathbf{u} \in \mathbb{S}^{d-1}} \|M\mathbf{u}\|$. Throughout the paper, let $\sigma^2(\mathbf{u})$ be a positive-definite matrix for all $\mathbf{u} \in \mathbb{S}^{d-1}$.

(A3) Suppose that, as $r \to \infty$, we have $\varepsilon(r) := \sup_{\mathbf{x} \in \mathbb{X}: \|\mathbf{x}\| \geq r} \|M(\mathbf{x}) - \sigma^2(\hat{\mathbf{x}})\| \to 0$.

(A4) Suppose that there exist constants U, V with $0 < U < V < \infty$ such that, for all $\mathbf{u} \in \mathbb{S}^{d-1}$, $\langle \mathbf{u}, \sigma^2(\mathbf{u})\mathbf{u} \rangle = U$ and $\operatorname{tr} \sigma^2(\mathbf{u}) = V$. In the case $2U = V$, suppose in addition that $\varepsilon(r)$ as defined in (A3) satisfies $\varepsilon(r) = O(r^{-\delta})$ for some $\delta > 0$.

Under assumptions (A0)–(A4), it was proved in [9] that the walk is transient if and only if $2U < V$, while [10] gives an invariance principle for the radial component $\|X\|$. The full invariance principle of the present paper requires additional structure on the limiting covariance matrix σ^2 to ensure that the angular part is a suitably well-behaved process on the sphere.

(A5) Suppose that $\sigma^2 : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ is a $\mathcal{C}^\infty$-function.

Controlling the dependence between the radial and angular components requires the following.

(A6) Suppose that $\mathbf{u}$ is an eigenvector of $\sigma^2(\mathbf{u})$ for all $\mathbf{u} \in \mathbb{S}^{d-1}$.

Following [9, §3], we describe a family of examples satisfying (A0)–(A6) in which the increment distribution is supported on an ellipsoid having one distinguished axis aligned in the radial direction. The model is specified by positive constants a, b. Let $Q_{\hat{\mathbf{x}}}$ be an orthogonal matrix representing a transformation of $\mathbb{R}^d$ that maps $\mathbf{e}_1$ to $\hat{\mathbf{x}}$, and write $D = \sqrt{d} \operatorname{diag}(a, b, \dots, b)$. Given X_0, the law of X_1 is generated by taking ζ uniform on $\mathbb{S}^{d-1}$; if $X_0 = \mathbf{0}$ set $X_1 = \zeta$, otherwise set $X_1 - X_0 = Q_{\hat{\mathbf{x}}} D \zeta$. In words, from $X_0 \neq \mathbf{0}$ the position X_1 is generated by taking a uniform point on the unit sphere centred at X_0, stretched differentially in the radial and transverse directions to give a point on an ellipsoid. The special case $a = b$ is a Pearson–Rayleigh random walk. A calculation [9, p. 104] shows that

$$\sigma^2(\mathbf{u}) = a^2 \mathbf{u}\mathbf{u}^\top + b^2(I - \mathbf{u}\mathbf{u}^\top).$$

In particular, $\operatorname{tr} \sigma^2(\mathbf{u}) = a^2 + (d-1)b^2$, $\sigma^2(\mathbf{u})\mathbf{u} = a^2\mathbf{u}$, and $\langle \mathbf{u}, \sigma^2(\mathbf{u})\mathbf{u} \rangle = a^2$, while $M(\mathbf{x}) = \sigma^2(\hat{\mathbf{x}})$ for $\mathbf{x} \neq \mathbf{0}$. Thus (A0)–(A6) hold. Without loss of generality, we may take $U = a = 1$. Then $V = 1 + (d-1)b^2$, and $\sigma_{\mathrm{sy}}(\mathbf{u}) = \mathbf{u}\mathbf{u}^\top + b(I - \mathbf{u}\mathbf{u}^\top)$, so that the spherical part of $\mathcal{X}$ is driven by the SDE (3.8), which reduces in this case to $\mathrm{d}X_t = b(I - \hat{X}_t\hat{X}_t^\top)\mathrm{d}W_t - \frac{(d-1)b^2}{2} \frac{\hat{X}_t}{\|X_t\|}\mathrm{d}t$, which corresponds to a BM on $\mathbb{S}^{d-1}$ sped up by a factor of b. The diffusion limits generated by this family of random walks thus include the classical skew-product description of BM as a special case, but also include examples where $\mathbf{0}$ is recurrent.

3 The diffusion limit

3.1 Overview

Let $\sigma_{\rm sy} : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ be the unique positive-definite matrix-valued function satisfying $\sigma_{\rm sy}\sigma_{\rm sy}^\top = \sigma^2$, i.e. $\sigma_{\rm sy}$ is the unique *symmetric square root* of σ^2. Pick any measurable square root $\sigma : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ of σ^2 and note that, since σ^2 and $\sigma_{\rm sy}$ commute, the matrix $\sigma_{\rm sy}^{-1}(\mathbf{u})\sigma(\mathbf{u})$ is orthogonal for all $\mathbf{u} \in \mathbb{S}^{d-1}$. By Lévy's characterisation of Brownian motion, it is hence sufficient to prove Theorem 1.1 for the SDE

$$\mathrm{d}\mathcal{X}_t = \sigma_{\rm sy}(\hat{\mathcal{X}}_t)\mathrm{d}W_t, \qquad \mathcal{X}_0 = \mathbf{x}_0 \in \mathbb{R}^d. \tag{3.1}$$

The next step is to establish weak existence for SDE (3.1). We start with a simple lemma.

Lemma 3.1. *Under* (A4) *and* (A5)*, $\sigma_{\rm sy}$ is uniformly elliptic in the following sense: there exists a constant $\lambda > 0$ such that $\langle \mathbf{v}, \sigma_{\rm sy}(\mathbf{u})\mathbf{v}\rangle \geq \lambda$ for all $\mathbf{u}, \mathbf{v} \in \mathbb{S}^{d-1}$.*

Proof. Since σ^2 is positive-definite, by (A5) and the compactness of $\mathbb{S}^{d-1}$ there exists $\varepsilon > 0$ such that $\det(\sigma^2) > \varepsilon$ on $\mathbb{S}^{d-1}$. By (A4) we have $\operatorname{tr}\sigma^2(\mathbf{u}) = V$. Hence the smallest eigenvalue $\lambda_{\min}(\mathbf{u})$ of $\sigma^2(\mathbf{u})$ satisfies $\varepsilon < \lambda_{\min}(\mathbf{u})V^{d-1}$ for all $\mathbf{u} \in \mathbb{S}^{d-1}$. Since $\sigma_{\rm sy}$ is symmetric and non-degenerate, its eigenvalues are positive and the smallest one is equal to $\sqrt{\lambda_{\min}(\mathbf{u})}$. Hence the inequality in the lemma holds for the constant $\lambda := (\varepsilon/V^{d-1})^{1/2}$. $\square$

Since the function $\mathbf{x} \mapsto \sigma_{\rm sy}(\hat{\mathbf{x}})$ is bounded and uniformly elliptic by Lemma 3.1, [18, §2.6, Thm 1] implies that weak existence holds for SDE (3.1). Once uniqueness in law for SDE (3.1) is established, the strong Markov property (and hence Theorem 1.1) follows by [26, Thm 6.2.2].

The proof of uniqueness in law proceeds as follows. Throughout this section, assume $U = 1$ in (A4). In Section 3.2 we prove that the radial component of any solution of (3.1) is Bessel of dimension $V > 1$. Section 3.3 introduces the Riemannian structure on the sphere, needed in Section 3.4 to characterize the law of a stationary diffusion on $\mathbb{S}^{d-1}$ indexed by $\mathbb{R}$. This process is a key ingredient in the description of the projection of the path of the solution $\mathcal{X}$ of SDE (3.1) (away from $\mathbf{0}$) onto $\mathbb{S}^{d-1}$. In Section 3.5 we analyse the case when 0 is polar for the radial process ($V \geq 2$). We prove that any solution has a skew-product decomposition constructed using the components from Sections 3.2 and 3.4 that are unique in law. In Section 3.6 we consider the recurrent case ($1 < V < 2$). We develop the excursion theory (away from $\mathbf{0}$) of the solution $\mathcal{X}$ of (3.1) without reference to the strong Markov property of $\mathcal{X}$. We characterize the excursion measure in terms of the excursion measure of the radial part, given in [23], and the law of the diffusion on $\mathbb{S}^{d-1}$ from Section 3.4. This implies the uniqueness in law for SDE (3.1).

3.2 The radial process

Let $r := \|\mathcal{X}\|$ be the radial part of a solution $\mathcal{X}$ of SDE (3.1).

Lemma 3.2. *Let* (A4) *hold and $\sigma^2 : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$ be measurable. For any solution $(\mathcal{X}, W)$ of SDE* (3.1)*, adapted to a filtration $(\mathcal{F}_t, t \geq 0)$, the process $y = (y_t, t \geq 0)$, $y_t := \|\mathcal{X}_t\|^2$, is the unique strong solution of SDE*

$$y_t = \|\mathcal{X}_0\|^2 + 2\int_0^t \sqrt{y_s}\mathrm{d}Z_s + Vt, \quad t \geq 0, \tag{3.2}$$

where $(Z_t, t \geq 0)$ is an $(\mathcal{F}_t)$ Brownian motion given by $Z_t := \int_0^t \hat{\mathcal{X}}_s^\top \sigma_{\rm sy}(\hat{\mathcal{X}}_s)\mathrm{d}W_s$. In particular, the law of $r = \sqrt{y}$ is $\mathrm{BES}^V(\|\mathcal{X}_0\|)$.

Remark 3.3. A solution $\mathcal{X}$ of SDE (3.1) is continuous and hence predictable (see [24, §IV.5]). Since $\mathbf{x} \mapsto \sigma_{\mathrm{sy}}(\hat{\mathbf{x}})\hat{\mathbf{x}}$ is measurable on $\mathbb{R}^d$ (recall that we defined $\hat{\mathbf{0}} := \mathbf{e}_1$), the integrand in the definition of Z is a bounded predictable process. Hence the stochastic integral Z is well defined, even though (due to rapid spinning, see Section 3.6 below) its integrand is far from continuous. Moreover, the integrand does not in general have paths in $\mathcal{D}_d$ (defined in Section 4.1 below).

Remark 3.4. Assuming (A6), the Brownian motion Z in Lemma 3.2 can be expressed as

$$Z_t = \int_0^t \hat{\mathcal{X}}_u^{\top} \mathrm{d}W_u. \tag{3.3}$$

Proof of Lemma 3.2. For any solution $(\mathcal{X}, W)$ of (3.1), the processes y and Z defined in the lemma are $(\mathcal{F}_t)$-adapted. Itô's formula and the assumption (A4) imply that equation (3.2) holds. The process Z is a Brownian motion by Lévy's characterisation, (A4) and assumption $U = 1$. Since SDE (3.2) has weak existence and pathwise uniqueness, the law of y is $\mathrm{BESQ}^V(\|\mathcal{X}_0\|^2)$. □

3.3 A Riemannian structure on $\mathbb{S}^{d-1}$

This section introduces a Riemannian metric g on $\mathbb{S}^{d-1}$, gives an explicit description of its inverse tensor in local coordinates and relates it to the Laplace–Beltrami operator corresponding to g (see [14] as reference on Riemannian geometry).

Identify the tangent space $T_{\mathbf{x}}\mathbb{S}^{d-1}$ at $\mathbf{x} \in \mathbb{S}^{d-1}$ with the $(d-1)$-dimensional linear subspace $\{v \in \mathbb{R}^d : \langle v, \mathbf{x}\rangle = 0\}$ of $\mathbb{R}^d$ and let the cotangent space $T^*_{\mathbf{x}}\mathbb{S}^{d-1}$ be the vector space dual of $T_{\mathbf{x}}\mathbb{S}^{d-1}$. Denote by $T\mathbb{S}^{d-1}$ and $T^*\mathbb{S}^{d-1}$ the tangent and cotangent [14, Def 2.1.9] bundles over $\mathbb{S}^{d-1}$, respectively. Any smooth section of the vector bundle $T^*\mathbb{S}^{d-1} \otimes T^*\mathbb{S}^{d-1}$, defined in [14, Def 2.1.10], is known as a $(0,2)$-tensor field. Let

$$g_{\mathbf{x}}(v_1, v_2) := \langle \sigma^{-2}(\mathbf{x}) v_1, v_2\rangle \qquad \text{for any } \mathbf{x} \in \mathbb{S}^{d-1} \text{ and } v_1, v_2 \in T_{\mathbf{x}}\mathbb{S}^{d-1}. \tag{3.4}$$

By (A5), g is a symmetric positive-definite $(0,2)$-tensor field, i.e., a Riemmanian metric on the smooth manifold $\mathbb{S}^{d-1}$. The metric g provides a *canonical* way of identifying tangent and cotangent vectors: the map $\tilde{g} : T\mathbb{S}^{d-1} \to T^*\mathbb{S}^{d-1}$ given by $\tilde{g}_{\mathbf{x}}(v) : T_{\mathbf{x}}\mathbb{S}^{d-1} \to \mathbb{R}$, where $\tilde{g}_{\mathbf{x}}(v)(u) := g_{\mathbf{x}}(v, u)$ for any $\mathbf{x} \in \mathbb{S}^{d-1}$, $v, u \in T_{\mathbf{x}}\mathbb{S}^{d-1}$, is a bundle isomorphism [14, Def. 2.1.6]. For any $f \in \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R})$, there exists a unique smooth section df of the cotangent bundle $T^*\mathbb{S}^{d-1}$, representing the action of the derivative of f on each tangent space [14, §1.2]. A vector field on the sphere is an element in the module $\Gamma(T\mathbb{S}^{d-1})$ (over the ring $\mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R})$) of smooth sections of $T\mathbb{S}^{d-1}$ [14, Def 2.1.3]. Let the *gradient* of f be $\operatorname{grad} f := \tilde{g}^{-1}(df)$. Hence $\operatorname{grad} f$ is the unique vector field satisfying the identity $g(\operatorname{grad} f, X) = df X$ for all $X \in \Gamma(T\mathbb{S}^{d-1})$. Moreover, the operator $\operatorname{grad} : \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R}) \to \Gamma(T\mathbb{S}^{d-1})$ is defined in a coordinate free fashion.

There exists a unique connection (the Levi-Civita connection) [14, Def 4.1.1] $\nabla : T\mathbb{S}^{d-1} \times \Gamma(T\mathbb{S}^{d-1}) \to T\mathbb{S}^{d-1}$ on $(\mathbb{S}^{d-1}, g)$, which is metric and torsion-free [14, Thm 4.3.1]. In short, the connection ∇ allows us to compare tangent vectors in nearby tangent spaces in a way that is compatible with the geometry induced by the metric g, cf. [14, §§4.1 & 4.2]. In particular, a vector field $X \in \Gamma(T\mathbb{S}^{d-1})$ gives rise to a linear endomorphism $(\nabla X)_{\mathbf{x}} : T_{\mathbf{x}}\mathbb{S}^{d-1} \to T_{\mathbf{x}}\mathbb{S}^{d-1}$ for any $\mathbf{x} \in \mathbb{S}^{d-1}$ [14, Def. 4.1.1]. Put differently, $\nabla_v X$ is the derivative of the vector field X at $\mathbf{x}$ in the direction $v \in T_{\mathbf{x}}\mathbb{S}^{d-1}$. Define the *divergence* of the vector field X to be the trace of this linear endomorphism, $(\operatorname{div} X)(\mathbf{x}) := \operatorname{tr}(\nabla X)_{\mathbf{x}}$. This yields a coordinate-free definition of the divergence operator $\operatorname{div} : \Gamma(T\mathbb{S}^{d-1}) \to \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R})$. The *Laplace–Beltrami operator* $\Delta_g : \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R}) \to \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R})$ on the Riemannian manifold $(\mathbb{S}^{d-1}, g)$ can now also be defined in a coordinate-free way as $\Delta_g f := \operatorname{div}(\operatorname{grad} f)$ for any $f \in \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R})$.

We now introduce local coordinates on $\mathbb{S}^{d-1}$ in order to identify the bundle isomorphism $\tilde{g}^{-1} : T^*\mathbb{S}^{d-1} \to T\mathbb{S}^{d-1}$. For each $q \in \{1,\dots,d\}$, define $[q] := \{1,\dots,d\} \setminus \{q\}$ and, throughout this section, identify $\mathbb{R}^{d-1}$ with the linear subspace of $\mathbb{R}^d$ spanned by $\{\mathbf{e}_i; i \in [q]\}$. Consider an atlas of charts $\mathbf{z}_q : H_q^\pm \to B^{d-1}$ on $\mathbb{S}^{d-1}$, where $\pm$ is either $+$ or $-$, $H_q^\pm := \{\mathbf{x} = (x_1,\dots,x_d)^\top \in \mathbb{S}^{d-1} : \pm x_q > 0\}$ is a hemisphere, B^{d-1} is the open unit ball in $\mathbb{R}^{d-1}$ and $\mathbf{z}_q(\mathbf{x}) := \sum_{i\in[q]} x_i \mathbf{e}_i$. The derivative of the smooth inverse $\mathbf{z}_q^{-1} : B^{d-1} \to H_q^\pm$ induces a linear isomorphism $d\mathbf{z}_q^{-1}(z) : T_z B^{d-1} \to T_{\mathbf{z}_q^{-1}(z)} H_q^\pm$ for each $z \in B^{d-1}$. Using the canonical identification $T_z B^{d-1} \equiv \mathbb{R}^{d-1}$ for all $z \in B^{d-1}$, at each $\mathbf{x} \in H_q^\pm$ we obtain the basis $\mathcal{B}_\mathbf{x} := \{E_i := d\mathbf{z}_q^{-1}(\mathbf{z}_q(\mathbf{x}))\mathbf{e}_i; i \in [q]\}$ of $T_\mathbf{x}\mathbb{S}^{d-1}$ and dual basis $\mathcal{B}^*_\mathbf{x} := \{E_i^*; i \in [q]\}$ of $T^*_\mathbf{x}\mathbb{S}^{d-1}$, defined by $E_i^*(E_j) = \delta_{ij}$ for $i,j \in [q]$, where δ_{ij} is the Kronecker delta. We interpret the tangent vector E_i as a linear map $E_i : \mathcal{C}^\infty(H_q^\pm, \mathbb{R}) \to \mathcal{C}^\infty(H_q^\pm, \mathbb{R})$ satisfying the Leibniz rule, $E_i(f) : \mathbf{x} \mapsto \partial_i (f \circ \mathbf{z}_q^{-1})(\mathbf{z}_q(\mathbf{x}))$, where ∂_i is the partial derivative in the i-th component [12, p. 247].

Lemma 3.5. *Assume* (A4)–(A6). *For $\mathbf{x} \in H_q^\pm$, the matrix $(g^{ij}(\mathbf{x}))_{i,j\in[q]}$ corresponding to the linear isomorphism $\tilde{g}_\mathbf{x}^{-1} : T^*_\mathbf{x}\mathbb{S}^{d-1} \to T_\mathbf{x}\mathbb{S}^{d-1}$ in terms of the bases $\mathcal{B}^*_\mathbf{x}$ and $\mathcal{B}_\mathbf{x}$, equals $g^{ij}(\mathbf{x}) = \sigma^2_{ij}(\mathbf{x}) - x_i x_j$ for any $i,j \in [q]$. The inverse matrix $(g_{ij}(\mathbf{x}))_{i,j\in[q]}$, corresponding to the isomorphism $\tilde{g}_\mathbf{x} : T_\mathbf{x}\mathbb{S}^{d-1} \to T^*_\mathbf{x}\mathbb{S}^{d-1}$, is given by $g_{ij}(\mathbf{x}) = \sigma^{-2}_{ij}(\mathbf{x}) + \sigma^{-2}_{qq}(\mathbf{x}) x_i x_j / \langle \mathbf{x}, \mathbf{e}_q\rangle^2 - (\sigma^{-2}_{qi}(\mathbf{x}) x_j + \sigma^{-2}_{qj}(\mathbf{x}) x_i)/\langle \mathbf{x}, \mathbf{e}_q\rangle$, for any $i,j \in [q]$. Moreover, in the coordinates on $H_q^\pm$, Δ_g equals*

$$\Delta_g f = \sum_{i,j\in[q]} g^{ij}\big(E_i(E_j(f)) - \sum_{k\in[q]} \Gamma^k_{ij} E_k(f)\big), \qquad \text{for any } f \in \mathcal{C}^\infty(H_q^\pm, \mathbb{R}),$$

where $\Gamma^k_{ij} := \frac{1}{2}\sum_{\ell\in[q]} g^{k\ell}(E_i(g_{j\ell}) + E_j(g_{i\ell}) - E_\ell(g_{ij}))$ for $i,j,k \in [q]$.

Proof. Recall that $B^{d-1} \subset \mathbb{R}^{d-1} \equiv \mathrm{Lin}\{\mathbf{e}_i; i \in [q]\} \subset \mathbb{R}^d$. For any point $z \in B^{d-1}$ and tangent vector $u \in \mathbb{R}^{d-1}$ we have $d\mathbf{z}_q^{-1}(z)u = u - \mathbf{e}_q \langle z, u\rangle / \langle \mathbf{z}_q^{-1}(z), \mathbf{e}_q\rangle$. Since $g_{ij}(\mathbf{x}) = g_\mathbf{x}(d\mathbf{z}_q^{-1}(\mathbf{z}_q(\mathbf{x}))\mathbf{e}_i, d\mathbf{z}_q^{-1}(\mathbf{z}_q(\mathbf{x}))\mathbf{e}_j)$ for any $i,j \in [q]$, the formula for $g_{ij}(\mathbf{x})$ follows by (3.4).

We now prove that $(g^{ij}(\mathbf{x}))_{i,j\in[q]}$, defined in the lemma, is the inverse of $(g_{ij}(\mathbf{x}))_{i,j\in[q]}$. Define $(d-1)$-dimensional square matrices S^- and S as follows: $S^-_{ij} := \sigma^{-2}_{ij}(\mathbf{x})$ and $S_{ij} := \sigma^2_{ij}(\mathbf{x})$ for any $i,j \in [q]$. Define $(d-1)$-dimensional vectors S^-_q, S_q by $S^-_{q,i} := \sigma^{-2}_{qi}(\mathbf{x})$ and $S_{q,i} := \sigma^2_{qi}(\mathbf{x})$ for $i \in [q]$. Let $s := \sigma^2_{qq}(\mathbf{x})$ and $s^- := \sigma^{-2}_{qq}(\mathbf{x})$. Since $\sigma^{-2}(\mathbf{x})\sigma^2(\mathbf{x})$ is the identity on $\mathbb{R}^d$, we have

$$S^- S + S^-_q S_q^\top = I, \qquad S^- S_q = -s S^-_q, \qquad S S^-_q = -s^- S_q, \tag{3.5}$$

where I denotes the identity matrix on $\mathbb{R}^{d-1}$. Denote $z := \mathbf{z}_q(\mathbf{x})$, and $D := \pm\sqrt{1 - \|z\|^2}$. Since $\mathbf{x} = z + D\mathbf{e}_q \in \mathbb{S}^{d-1}$, the assumption in (A6) implies $\sigma^{-2}(\mathbf{x})(z + D\mathbf{e}_q) = z + D\mathbf{e}_q$ (recall $U = 1$). Hence the following identities hold,

$$S^- z_q = z_q - D S^-_q, \qquad z_q^\top S^-_q = (1 - s^-)D, \qquad S z_q = z_q - D S_q, \tag{3.6}$$

where z_q denotes the $(d-1)$-tuple of coordinates of z expressed in the basis $\{\mathbf{e}_i; i \in [q]\}$ of $\mathbb{R}^{d-1}$. Define $(d-1)$-dimensional square matrices G, G^- as follows:

$$G^- := S - z_q z_q^\top, \qquad G := S^- + s^- z_q z_q^\top / D^2 - (z_q S_q^{-\top} + S^-_q z_q^\top)/D.$$

A direct calculation, using identities in (3.5)–(3.6) and the fact that $S = S^\top$ and $S^- = S^{-\top}$, yields $GG^- = I$. It remains to note that $G^-_{ij} = g^{ij}(\mathbf{x})$ and $G_{ij} = g_{ij}(\mathbf{x})$ for all $i,j \in [q]$.

The expression for the Laplace–Beltrami operator $\Delta_g = \operatorname{div}\operatorname{grad} = \operatorname{tr}\operatorname{Hess}_g$ in local coordinates in terms of the Christoffel symbols Γ^k_{ij} is well-known, cf. [12, Ch V, Eqs (4.19) & (4.32)]. □

3.4 A stationary diffusion on $\mathbb{S}^{d-1}$

Define $A : \mathbb{R}^d \setminus \{\mathbf{0}\} \to \mathbb{R}^d \otimes \mathbb{R}^d$ by $A(\mathbf{y}) := \sigma_{\mathrm{sy}}(\hat{\mathbf{y}})$, $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, and note that it is an extension of $\sigma_{\mathrm{sy}} : \mathbb{S}^{d-1} \to \mathbb{R}^d \otimes \mathbb{R}^d$. For any $j \in \{1, \dots, d\}$, define $A_j : \mathbb{R}^d \setminus \{\mathbf{0}\} \to \mathbb{R}^d$ by $A_j(\mathbf{y}) = A(\mathbf{y})\mathbf{e}_j$ and note that its derivative $DA_j(\mathbf{y})$ at $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ (i.e. a linear endomorphism of $\mathbb{R}^d$ satisfying $(A_j(\mathbf{y} + \mathbf{h}) - A_j(\mathbf{y}) - DA_j(\mathbf{y})\mathbf{h})/\|\mathbf{h}\| \to \mathbf{0}$ as $\|\mathbf{h}\| \to 0$) exists since, by Lemma 3.1, σ_{sy} can be expressed as an absolutely convergent power series in σ^2, which is smooth by (A5). Let $A_0 : \mathbb{R}^d \setminus \{\mathbf{0}\} \to \mathbb{R}^d$ be given by $A_0(\mathbf{y}) := \frac{1}{2}\sum_{j=1}^d DA_j(\mathbf{y})A_j(\mathbf{y})$ for any $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$.

Let $S_0, S_j : \mathbb{S}^{d-1} \to \mathbb{R}^d$ be $S_0(\mathbf{x}) := -(I - \mathbf{x}\mathbf{x}^\top)A_0(\mathbf{x})$ and $S_j(\mathbf{x}) := (\sigma_{\mathrm{sy}}(\mathbf{x}) - \mathbf{x}\mathbf{x}^\top)\mathbf{e}_j$ for any $\mathbf{x} \in \mathbb{S}^{d-1}$ and $j \in \{1, \dots, d\}$. Let $\mathcal{C}(\mathbb{R}_+, \mathbb{S}^{d-1})$ be equipped with the Borel σ-algebra generated by the compact-open topology [6, §XII.1], which coincides with the σ-algebra generated by the projections at any time $t \in \mathbb{R}$, cf. [4, p. 57].

Lemma 3.6. *Assume* (A4)–(A6). *Then the following statements hold.*

(a) $S_0(\mathbf{x}), \dots, S_d(\mathbf{x}) \in T_{\mathbf{x}}\mathbb{S}^{d-1}$ *for all* $\mathbf{x} \in \mathbb{S}^{d-1}$ *and the vector fields* $S_0, \dots, S_d$ *are in* $\Gamma(T\mathbb{S}^{d-1})$.

(b) Let W *be a standard Brownian motion on* $\mathbb{R}^d$. *The Stratonovich SDE on* $\mathbb{S}^{d-1}$, *given by*

$$\mathrm{d}X_t = S_0(X_t)\mathrm{d}t + \sum_{j=1}^{d} S_j(X_t) \circ \mathrm{d}W_t^j, \qquad X_0 = \mathbf{x} \in \mathbb{S}^{d-1}, \tag{3.7}$$

has a unique strong solution in the sense of [12, Ch V, Def 1.1 & Thm 1.1].

(c) Let $\mathbb{P}_{\mathbf{x}}$ *denote the law of the solution of* (3.7) *on* $\mathcal{C}(\mathbb{R}_+, \mathbb{S}^{d-1})$. *Then* $\{\mathbb{P}_{\mathbf{x}}, \mathbf{x} \in \mathbb{S}^{d-1}\}$ *is a strongly Markovian system [12, p. 204], determined uniquely by its generator* $\mathcal{G}$,

$$\mathcal{G}f := S_0(f) + \frac{1}{2}\sum_{i=1}^{d} S_i(S_i(f)) \qquad \text{for any } f \in \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R}),$$

where the vector fields S_i, $i \in \{0, \dots, d\}$, *are viewed as linear (over* $\mathbb{R}$*) maps* $\mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R}) \to \mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R})$ *satisfying the Leibniz rule.*

(d) $V_0 := \mathcal{G} - \frac{1}{2}\Delta_g$ *is a vector field in* $\Gamma(T\mathbb{S}^{d-1})$, *making the solution of* (3.7) *a Brownian motion with drift on the Riemannian manifold* $(\mathbb{S}^{d-1}, g)$ *with generator* $\frac{1}{2}\Delta_g + V_0$.

(e) Any solution (X, W) *of the Itô SDE*

$$\mathrm{d}X_t = (\sigma_{\mathrm{sy}}(\hat{X}_t) - \hat{X}_t\hat{X}_t^\top)\mathrm{d}W_t - \frac{V-1}{2}\frac{\hat{X}_t}{\|X_t\|}\mathrm{d}t, \quad X_0 = \mathbf{x} \in \mathbb{S}^{d-1} \tag{3.8}$$

satisfies $\|X_t\| = 1$ *for all* $t \in \mathbb{R}_+$ *and is a solution of SDE* (3.7).

Proof. The vector fields S_j, $j \in \{0, \dots, d\}$, are tangential to $\mathbb{S}^{d-1}$ by (A6) and smooth by (A5). Hence (a) holds. Moreover, we may interpret S_j as a linear map on $\mathcal{C}^\infty(\mathbb{S}^{d-1}, \mathbb{R})$ satisfying the Leibniz rule [12, p. 248] (see e.g. (3.9) below). Hence part (b) of the lemma follows from [12, Ch V, Thm 1.1]. The family of laws $\{\mathbb{P}_{\mathbf{x}}, \mathbf{x} \in \mathbb{S}^{d-1}\}$ is a strongly Markovian system generated by the second order differential operator $\mathcal{G}$ by [12, Ch V, Thm 1.2], which establishes part (c).

To establish part (d), consider a chart $\mathbf{z}_q : H_q^\pm \to B^{d-1}$ (for some $q \in \{1, \dots, d\}$) and the corresponding frame field $\{E_i, i \in [q]\}$, defined in the paragraph preceding Lemma 3.5. Then we can express the vector field S_j on $H_q^\pm$ as a linear mapping from

$\mathcal{C}^\infty(H_q^\pm,\mathbb{R}) \to \mathcal{C}^\infty(H_q^\pm,\mathbb{R})$, satisfying the Leibniz rule, as follows: for any $\mathbf{x} \in H_q^\pm$ and $j \in [q]$ we have

$$S_j(f)(\mathbf{x}) = (D\mathbf{z}_q(\mathbf{x})S_j(\mathbf{x}))^\top \sum_{i\in[q]} E_i(f)(\mathbf{x})\mathbf{e}_i = \sum_{i\in[q]} S_j^i(\mathbf{x})E_i(f)(\mathbf{x}), \tag{3.9}$$

where the second equality holds by $D\mathbf{z}_q = \mathbf{z}_q$, and where $S_j^i(\mathbf{x}) = \langle S_j(\mathbf{x}), \mathbf{e}_i\rangle$. This implies $S_j(S_j(f)) = \sum_{i,k\in[q]} S_j^i S_j^k E_i(E_k(f)) + \sum_{k\in[q]} \bar{V}_{k,j} E_k(f)$ for some functions $\bar{V}_{k,j} \in \mathcal{C}^\infty(H_q^\pm,\mathbb{R})$, $k,j \in [q]$, and all $f \in \mathcal{C}^\infty(H_q^\pm,\mathbb{R})$. The definition of S_j above, (A4), (A6) and Lemma 3.5 imply $\sum_{j=1}^d S_j^i(\mathbf{x})S_j^k(\mathbf{x}) = g^{ik}(\mathbf{x})$ for all $\mathbf{x} \in H_q^\pm$ and $i,k \in [q]$. Hence, by the definition of $\mathcal{G}$ in the lemma and the expression for Δ_g in the local coordinates on $H_q^\pm$ in Lemma 3.5, the equality $V_0(f) = \sum_{i\in[q]} V_{0,i}E_i(f)$ holds for some functions $V_{0,i} \in \mathcal{C}^\infty(H_q^\pm,\mathbb{R})$, $i \in [q]$. Since such an equality holds for every $q \in \{1,\dots,d\}$ and choice of $\pm$ (i.e. for every chart in our atlas), V_0 satisfies the Leibniz rule and is hence an element of $\Gamma(T\mathbb{S}^{d-1})$, implying (d).

Extend the vector fields $S_0, S_1, \dots, S_d$ to $\mathbb{R}^d \setminus \{\mathbf{0}\}$ by defining $\bar{S}_0(\mathbf{y}) := -(I - \hat{\mathbf{y}}\hat{\mathbf{y}}^\top)A_0(\mathbf{y})$ and $\bar{S}_j(\mathbf{y}) := (A(\mathbf{y}) - \hat{\mathbf{y}}\hat{\mathbf{y}}^\top)\mathbf{e}_j$, $j \in \{1,\dots,d\}$, for any $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$. Define a function $R : \mathbb{R}^d \setminus \{\mathbf{0}\} \to \mathbb{R}^d$ by $R(\mathbf{y}) := \frac{1}{2}\sum_{j=1}^d D\bar{S}_j(\mathbf{y})\bar{S}_j(\mathbf{y})$. To prove (e), we establish the following formula

$$R(\mathbf{y}) = (I - \hat{\mathbf{y}}\hat{\mathbf{y}}^\top)A_0(\mathbf{y}) - \frac{V-1}{2}\frac{\hat{\mathbf{y}}}{\|\mathbf{y}\|} \qquad \text{for all } \mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}. \tag{3.10}$$

Let $G(\mathbf{y}) := \hat{\mathbf{y}}$ for any $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and note that $A = A \circ G$ and $DG(\mathbf{y}) = (I - \hat{\mathbf{y}}\hat{\mathbf{y}}^\top)/\|\mathbf{y}\|$, implying $DG(\mathbf{y})\mathbf{y} = \mathbf{0}$, $DG(\mathbf{y})^\top = DG(\mathbf{y})$ and $DA_j(\mathbf{y})\mathbf{y} = DA_j(\hat{\mathbf{y}})DG(\mathbf{y})\mathbf{y} = \mathbf{0}$ for all $j \in \{1,\dots,d\}$. Since $\bar{S}_j(\mathbf{y}) = A_j(\mathbf{y}) - \hat{\mathbf{y}}\langle\hat{\mathbf{y}},\mathbf{e}_j\rangle$, we get $D\bar{S}_j(\mathbf{y}) = DA_j(\mathbf{y}) - (\hat{\mathbf{y}}^\top\mathbf{e}_j I + \hat{\mathbf{y}}\mathbf{e}_j^\top)DG(\mathbf{y})$ by the product rule, where I is the identity matrix on $\mathbb{R}^d$. Hence, using the fact that $A(\mathbf{y})\mathbf{y} = \mathbf{y}$, we get $D\bar{S}_j(\mathbf{y})\bar{S}_j(\mathbf{y}) = DA_j(\mathbf{y})A_j(\mathbf{y}) - (\hat{\mathbf{y}}^\top\mathbf{e}_j I + \hat{\mathbf{y}}\mathbf{e}_j^\top)(A(\mathbf{y}) - \hat{\mathbf{y}}\hat{\mathbf{y}}^\top)\mathbf{e}_j/\|\mathbf{y}\|$. Summing over $j \in \{1,\dots,d\}$ yields the identity $2R(\mathbf{y}) = 2A_0(\mathbf{y}) - \mathrm{tr}(A(\mathbf{y}) - \hat{\mathbf{y}}\hat{\mathbf{y}}^\top)\hat{\mathbf{y}}/\|\mathbf{y}\|$. Differentiating the identity $A(\mathbf{y})\mathbf{y} = \mathbf{y}$ (in $\mathbf{y}$) yields $I = A(\mathbf{y}) + \sum_{j=1}^d \langle\mathbf{y},\mathbf{e}_j\rangle DA_j(\mathbf{y})$, and hence $A(\mathbf{y}) = A^2(\mathbf{y}) + \sum_{j=1}^d \langle\mathbf{y},\mathbf{e}_j\rangle DA_j(\mathbf{y})A(\mathbf{y})$, for all $\mathbf{y} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$. Since A is symmetric we have $DA_j(\mathbf{y})^\top\mathbf{e}_i = DA_i(\mathbf{y})^\top\mathbf{e}_j$ for all $i,j \in \{1,\dots,d\}$. Hence we have $2\langle A_0(\mathbf{y}),\mathbf{e}_j\rangle = \sum_{i=1}^d \langle A_i(\mathbf{y}), DA_i(\mathbf{y})^\top\mathbf{e}_j\rangle = \mathrm{tr}(DA_j(\mathbf{y})A(\mathbf{y}))$. Together with (A4), this implies $\mathrm{tr}\,A(\mathbf{y}) = V + 2\langle A_0(\mathbf{y}),\mathbf{y}\rangle$ and (3.10) follows.

Let (X,W) be a solution of (3.8). A simple application of Itô's formula yields $\mathrm{d}\|X_t\|^2 = 0$, implying the first statement in (e). By (3.10) it follows that X in fact satisfies the SDE $\mathrm{d}X_t = (\bar{S}_0(X_t) + R(X_t))\mathrm{d}t + \sum_{j=1}^d \bar{S}_j(X_t)\mathrm{d}W_t^j$, where $\bar{S}_j$, $j \in \{1,\dots,d\}$, are defined above (3.10). By the definition of the Stratonovich integral on $\mathbb{R}^d$ [12, Ch III, §1, Eq (1.10)], it follows that $\mathrm{d}X_t = \bar{S}_0(X_t)\mathrm{d}t + \sum_{j=1}^d \bar{S}_j(X_t) \circ \mathrm{d}W_t^j$. Since $S_j = \bar{S}_j$, $j \in \{0,\dots,d\}$, on $\mathbb{S}^{d-1}$ and X stays on the sphere for all time, SDE (3.7) holds for X (see [12, Ch V, Rem 1.1]). □

By Lemma 3.6(c), the map $\mathbf{x} \mapsto \mathbb{P}_\mathbf{x}[A]$ on $\mathbb{S}^{d-1}$ is Borel measurable for any Borel measurable set A in $\mathcal{C}(\mathbb{R}_+,\mathbb{S}^{d-1})$. We can hence define a transition function on $\mathbb{S}^{d-1}$, $P_t(\mathbf{x},\cdot\,) := \mathbb{P}_\mathbf{x}[\phi_t \in \cdot\,]$, where $(t,\mathbf{x}) \in \mathbb{R}_+ \times \mathbb{S}^{d-1}$ and $(\phi_u, u \in \mathbb{R}_+)$ is the coordinate process on $\mathcal{C}(\mathbb{R}_+,\mathbb{S}^{d-1})$. In particular, the law $\mathbb{P}$ of the solution of (3.7), started according to a probability measure ν on $\mathbb{S}^{d-1}$, equals $\mathbb{P}[\,\cdot\,] = \int_{\mathbb{S}^{d-1}} \nu(\mathrm{d}\mathbf{x})\mathbb{P}_\mathbf{x}[\,\cdot\,]$.

Proposition 3.7. *Let* (A4)–(A6) *hold. There exists a unique probability measure μ on $\mathbb{S}^{d-1}$ with full support, such that $\mu(\cdot) = \int_{\mathbb{S}^{d-1}} \mu(\mathrm{d}\mathbf{x})P_t(\mathbf{x},\cdot)$ for all $t \in \mathbb{R}_+$ and the transition function $P_t(\mathbf{x},\cdot)$ converges to its stationary measure μ in the following sense:*[2]

$$\lim_{t\to\infty} \sup_{\mathbf{x}\in\mathbb{S}^{d-1}} \|P_t(\mathbf{x},\,\cdot\,) - \mu(\,\cdot\,)\|_{\mathrm{TV}} = 0. \tag{3.11}$$

[2]Recall that $\|\nu_1(\cdot) - \nu_2(\cdot)\|_{\mathrm{TV}} := \sup_{\mathfrak{A}\subset\mathbb{S}^{d-1}} |\nu_1(\mathfrak{A}) - \nu_2(\mathfrak{A})|$ for probability measures ν_1 and ν_2 on $\mathbb{S}^{d-1}$.

Furthermore, there exists a unique law $\mathbb{P}_\Psi[\,\cdot\,]$ *on the Borel sets of* $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$ *with compact-open topology, satisfying* $\mathbb{P}_\Psi[\psi_s \in \cdot\,] = \mu(\,\cdot\,)$ *and* $\mathbb{P}_\Psi[\psi_{s+t} \in \cdot \mid \psi_s] = P_t(\psi_s, \,\cdot\,)$ *for all* $(s,t) \in \mathbb{R} \times \mathbb{R}_+$*, where* $(\psi_u, u \in \mathbb{R})$ *denotes the coordinate process on* $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$*.*

Remark 3.8. (a) The unique stationary measure μ exists and has full support essentially because the vector fields $S_1, \ldots, S_d$ in Lemma 3.6(a) span $T_{\mathbf{x}}\mathbb{S}^{d-1}$ at every $\mathbf{x} \in \mathbb{S}^{d-1}$. The proof uses the representation in Lemma 3.6(d) of the process as a Brownian motion with drift and applies the well-known results for the stability of elliptic diffusions on compact Riemannian manifolds [22].
(b) The geometry introduced in Section 3.3 allows us to characterize the time-reversibility of the diffusion X satisfying SDE (3.7). This leads to an explicit description, given in (1.5) of Section 1.1 above, of the excursions of the process $\mathcal{X}$ appearing in Theorem 1.1.
(c) Kolmogorov's extension theorem [24, Thm III.1.5] and the first statement in Prop. 3.7 imply that $\mathbb{P}_\Psi[\,\cdot\,]$ exists and is unique: for $t_1 < \cdots < t_k$ in $\mathbb{R}$ the finite-dimensional distribution is $\int_{\mathfrak{A}_1} \mu(\mathrm{d}\mathbf{x}_1) \int_{\mathfrak{A}_2} P_{t_2 - t_1}(\mathbf{x}_1, \mathrm{d}\mathbf{x}_2) \cdots \int_{\mathfrak{A}_k} P_{t_k - t_{k-1}}(\mathbf{x}_{k-1}, \mathrm{d}\mathbf{x}_k)$ for measurable sets $\mathfrak{A}_i \subset \mathbb{S}^{d-1}$, $i = 1, \ldots, k$ (cf. [24, §XII.4]).

Proof of Proposition 3.7. By Lemma 3.6(d), the generator of the strong Markov process satisfying SDE (3.7) takes the form $\mathcal{G} = \frac{1}{2}\Delta_g + V_0$. The volume element $\mathrm{d}_g\mathbf{x}$ on the Riemannian manifold $(\mathbb{S}^{d-1}, g)$ is a $(d-1)$-dimensional form, given in local coordinates on $H_q^\pm$ by $\sqrt{\det G} \prod_{i \in [q]} \mathrm{d}x_i$, where $G = (g_{ij}(\mathbf{x}))_{i,j \in [q]}$ (see [12, p. 291] and Lemma 3.5 above). Let $\mathcal{G}^\star$ be the adjoint of $\mathcal{G}$ with respect to the measure $\mathrm{d}_g\mathbf{x}$. Assumptions of [22, Ch 4, Thm 11.1] are satisfied for the generator $\mathcal{G}$ since its second order term is the Laplace–Beltrami operator and the vector field V_0 is smooth by (A5). Hence by [22, Ch 4, Thm 11.1], all harmonic functions for $\mathcal{G}$ are constant and there exists a unique positive function $h \in \mathcal{C}^2(\mathbb{S}^{d-1}, \mathbb{R})$ satisfying $\mathcal{G}^\star h = 0$ and $\int_{\mathbb{S}^{d-1}} h(\mathbf{x})\mathrm{d}_g\mathbf{x} = 1$. Moreover, by [22, Ch 4, Thm 11.1(ix)], the assumptions of [22, Ch 4, Thm 8.6] for the Riemannian manifold $(\mathbb{S}^{d-1}, g)$ and the operator $\mathcal{G}$ are satisfied, implying that $\mu(\mathrm{d}\mathbf{x}) = h(\mathbf{x})\mathrm{d}_g\mathbf{x}$ is the unique stationary probability measure for the transition function $P_t(\mathbf{x}, \mathrm{d}\mathbf{y})$. Again, by [22, Ch 4, Thm 11.1(ix)], the assumptions of [22, Ch 4, Thm 9.9] for $(\mathbb{S}^{d-1}, g)$ and $\mathcal{G}$ are satisfied. Hence, as $\mathbb{S}^{d-1}$ is compact, [22, Ch 4, Thm 9.9] implies the convergence in total variation in (3.11). □

3.5 Proof of Theorem 1.1 when 0 is polar for the radial process

Assume throughout this section that $V \geq 2$ (and $U = 1$) and let $(\mathcal{X}, W)$ be any solution to (3.1), adapted to $(\mathcal{F}_t, t \geq 0)$, on a probability space that supports a one-dimensional $(\mathcal{F}_t)$ Brownian motion, independent of $(\mathcal{X}, W)$. By Lemma 3.2, 0 is polar for $r = \|\mathcal{X}\|$.

Lemma 3.9. *Let* (A4) *hold. If either (i)* $s > 0$*; or (ii)* $\mathcal{X}_0 \neq \mathbf{0}$ *and* $s = 0$*, define*

$$\rho_s(t) := \int_s^t r_u^{-2} \mathrm{d}u, \quad t \geq s. \tag{3.12}$$

Then, almost surely, $\rho_s : [s, \infty) \to \mathbb{R}_+$ *is continuously increasing and* $\lim_{t \uparrow \infty} \rho_s(t) = \infty$*. Its continuous inverse* $c_s : \mathbb{R}_+ \to [s, \infty)$ *is* $c_s(t) := \inf\{u \geq s : \rho_s(u) = t\}$*. In particular,* $c_s(0) = s$*.*

Lemma 3.9 is a direct consequence of the next lemma.

Lemma 3.10. *Pick* $x, m \in \mathbb{R}_+$ *and* $\delta \geq 2$*. Let* $\beta = (\beta_t, t \geq 0)$ *be* $\mathrm{BES}^\delta(x)$*,* $\tau_m := \inf\{t \geq 0 : \beta_t = m\}$ *(with* $\inf \emptyset = \infty$*) and* $f_m(y) := (m - y)^{-2}$*. If* $m > x$ *or* $x > 0 = m$*, then* $\int_0^{\tau_m} f_m(\beta_u)\mathrm{d}u = \infty$ *a.s. If* $x = m = 0$*, then for any* $t > 0$ *it holds that* $\int_0^t f_0(\beta_u)\mathrm{d}u = \infty$ *a.s.*

Proof. If $x < m$, then $\tau_m \in (0, \infty)$ a.s. for any $\delta \geq 2$, and $y \mapsto |y - m| f_m(y)$ is not integrable at m, so [5, Thm 2.2, Eq (2.5)] shows $\int_0^{\tau_m} f_m(\beta_u)\mathrm{d}u = \infty$ a.s. If $x > 0 = m$, then $\tau_0 = \infty$ a.s. for any $\delta \geq 2$, and the same result follows from [5, Thm 2.3(ii)] (when $\delta > 2$) and [5, Thm 2.4] ($\delta = 2$). Assume $x = m = 0$ and time-reverse β killed at τ_a (for some large $a > 0$) at the last time the process visits some $b \in (0, a)$ (this is a co-optional time, see [24, Ch VII.4] for details on time reversals). The time reversal is a diffusion on $(0, a)$ with the same volatility function as β and the scale function given by $\bar{s} = 1/(s(a) - s) : (0, a) \to \mathbb{R}$, where $s(y) = -y^{2-\delta}$ (resp. $\log(y)$) if $\delta > 2$ (resp. $\delta = 2$). Note that $\lim_{y\downarrow 0} \bar{s}(y) = 0$, $\lim_{y\uparrow a} \bar{s}(y) = \infty$ and $\bar{s} f_0/\bar{s}' = (s(a) - s) f_0/s'$ is not integrable at 0. Hence the lemma follows by [21, Thm 2.11(ii)]. $\square$

Proposition 3.11. *Suppose that* (A4)*,* (A5) *and* (A6) *hold. Assume either (i)* $s > 0$*; or (ii)* $\mathcal{X}_0 \neq \mathbf{0}$ *and* $s = 0$ *hold. Let a standard one-dimensional Brownian motion* Z *be given by* (3.3) *and let* c_s *be as in Lemma 3.9. The process* $\varphi = (\varphi_t, t \geq 0)$ *on* $\mathbb{S}^{d-1}$*, defined by* $\varphi_t := \hat{\mathcal{X}}_{c_s(t)}$*, is a strong solution of SDE* (3.8) *started at* $\varphi_0 = \hat{\mathcal{X}}_s$ *and driven by a* d*-dimensional Brownian motion* $(B_t, t \geq 0)$ *adapted to the filtration* $(\mathcal{F}_{c_s(t)}, t \geq 0)$*, independent of* $(Z_t, t \geq 0)$*.*

Proof. By assumption we have $r_s > 0$ a.s. Since 0 is polar for $\mathrm{BESQ}^V(r_s^2)$, $(r_t^{-2}; t \geq s)$ is a continuous semimartingale. Hence $\mathrm{d}(r_t^{-1}) = -r_t^{-2}\mathrm{d}Z_t - (V-3)/(2r_t^3)\mathrm{d}t$ by Itô's formula and (3.2). By (A6), the covariation equals $\mathrm{d}[\mathcal{X}, r^{-1}]_t = \sigma_{\mathrm{sy}}(\hat{\mathcal{X}}_t)\mathrm{d}[W, -W^\top]_t \sigma_{\mathrm{sy}}(\hat{\mathcal{X}}_t)\hat{\mathcal{X}}_t/r_t^2 = -\hat{\mathcal{X}}_t/r_t^2\mathrm{d}t$, and Itô's product rule implies

$$\mathrm{d}\hat{\mathcal{X}}_t = f(\hat{\mathcal{X}}_t) r_t^{-2}\mathrm{d}t + g(\hat{\mathcal{X}}_t) r_t^{-1}\mathrm{d}W_t, \quad t \geq s, \tag{3.13}$$

where we have used the notation

$$f(x) := -\frac{V-1}{2}\frac{\hat{x}}{\|x\|} \quad \text{and} \quad g(x) := \sigma_{\mathrm{sy}}(\hat{x}) - \hat{x}\hat{x}^\top, \qquad \text{for any } x \in \mathbb{R}^d. \tag{3.14}$$

Define continuous local martingales $A = (A_t; t \geq 0)$ and $\zeta = (\zeta_t; t \geq 0)$ by

$$A_t := \int_s^{c_s(t)} r_u^{-1}\mathrm{d}W_u \qquad \text{and} \qquad \zeta_t := \int_s^{c_s(t)} r_u^{-1}\mathrm{d}Z_u, \tag{3.15}$$

where Z is given in (3.3). Both A and ζ are adapted to $(\mathcal{F}_{c_s(t)}, t \geq 0)$. By [24, Prop. V.1.4–5] and Lemma 3.9 it holds that $[A, A^\top]_t = I\int_s^{c_s(t)} \frac{\mathrm{d}u}{r_u^2} = It$, where I is the identity matrix on $\mathbb{R}^d$, and $[\zeta, \zeta]_t = t$. Hence, by Lévy's characterisation theorem, both A and ζ are $(\mathcal{F}_{c_s(t)})$ Brownian motions. Furthermore, by (3.3) and [24, Prop. V.1.4–5], we have that $\zeta_t = \int_s^{c_s(t)} \hat{\mathcal{X}}_u^\top r_u^{-1}\mathrm{d}W_u = \int_0^t \varphi_u^\top \mathrm{d}A_u$ for all $t \geq 0$. Let $(\gamma_t', t \geq 0)$ be a one-dimensional $(\mathcal{F}_t)$ Brownian motion, independent of $(\mathcal{X}, W)$. Define $(\mathcal{F}_{c_s(t)})$ Brownian motion $\gamma = (\gamma_t, t \geq 0)$ by $\gamma_t := \int_s^{c_s(t)} r_u^{-1}\mathrm{d}\gamma_u'$ and note that $[\zeta, \gamma] \equiv 0$. Define $B = (B_t, t \geq 0)$ by $B_t := A_t - \int_0^t \varphi_u \mathrm{d}\zeta_u + \int_0^t \varphi_u \mathrm{d}\gamma_u$ and observe $\mathrm{d}[B, B^\top]_t = (I - \varphi_t\varphi_t^\top)^2\,\mathrm{d}t + \varphi_t\varphi_t^\top\mathrm{d}t = I\mathrm{d}t$ and $\mathrm{d}[B, \zeta]_t = (I - \varphi_t\varphi_t^\top)\mathrm{d}[A, A^\top]_t\varphi_t + \varphi_t\mathrm{d}[\gamma, \zeta]_t = 0$. In particular, B is a d-dimensional $(\mathcal{F}_{c_s(t)})$ Brownian motion, independent of ζ.

We now show B is independent of Z. By the Markov property, B_t depends on $\mathcal{F}_s = \mathcal{F}_{c_s(0)}$ only via $B_0 = \mathbf{0}$, so B is independent of $\mathcal{F}_s$. Hence B is independent of $(Z_t, t \in [0, s])$. It remains to prove that B is independent of $(Z_t - Z_s, t \geq s)$. Note that by (3.15) and Lemma 3.9 it holds that $Z_{c_s(t)} - Z_s = \int_s^{c_s(t)} r_u r_u^{-1}\mathrm{d}Z_u = \int_0^t r_{c_s(v)}\mathrm{d}\zeta_v$ for all $t \geq 0$. Hence the covariation of $\mathcal{F}_{c_s(t)}$-local martingales $M := Z_{c_s(\cdot)} - Z_s$ and B is identically equal to zero. Since the inverse of the quadratic variation $[M]_u = c_s(u) - s$ equals $\rho_s(s+u)$, by Knight's theorem [24, Thm V.1.9], the processes $M_{\rho_s(s+\cdot)}$ and B are independent Brownian motions. It only remains to note that $M_{\rho_s(s+u)} = Z_{s+u} - Z_s$ for any $u \geq 0$.

By definition we have $\varphi_t = \hat{\mathcal{X}}_s + \int_s^{c_s(t)} \mathrm{d}\hat{\mathcal{X}}_u$. Hence the change of variable formulas for Stieltjes [24, Prop. 0.4.1] and stochastic [24, Prop. V.1.4] integrals and (3.13) imply

$$\varphi_t = \varphi_0 + \int_0^t (\sigma_{\mathrm{sy}}(\varphi_u) - \varphi_u\varphi_u^\top)\mathrm{d}A_u - \frac{V-1}{2}\varphi_u\mathrm{d}u, \qquad t \geq 0. \tag{3.16}$$

Since $(\sigma_{\mathrm{sy}}(\varphi_t) - \varphi_t\varphi_t^\top)\mathrm{d}B_t = (\sigma_{\mathrm{sy}}(\varphi_t) - \varphi_t\varphi_t^\top)\big((I - \varphi_t\varphi_t^\top)\mathrm{d}A_t + \varphi_t\mathrm{d}\gamma_t\big) = (\sigma_{\mathrm{sy}}(\varphi_t) - \varphi_t\varphi_t^\top)\mathrm{d}A_t$, the process φ satisfies SDE (3.8) driven by $(B_t, t \geq 0)$ as required. ☐

Proof of Theorem 1.1 in the transient case with $\mathcal{X}_0 \neq \mathbf{0}$. By Prop. 3.11 (enlarge the probability space if needed), the law of any solution $\mathcal{X}$ of SDE (3.1), satisfying $\mathcal{X}_0 \neq \mathbf{0}$, is equal to that of $(r_t\varphi_{\rho_0(t)}, t \geq 0)$, where $r \sim \mathrm{BES}^V(\|\mathcal{X}_0\|)$, $\rho_0(\,\cdot\,)$ is given in (3.12) and φ is the unique solution of (3.8) with $\varphi_0 = \hat{\mathcal{X}}_0$, independent of r. ☐

In order to characterize the law of $\mathcal{X}$ in the case $V \geq 2$ with $\mathcal{X}_0 = \mathbf{0}$, we need to understand the law of the $\hat{\mathcal{X}}_s$ (for any fixed $s > 0$) and its dependence on the path of the radial process r. Define $\mathcal{F}^r_\infty := \sigma(r_t, t \geq 0)$. Since $r \sim \mathrm{BES}^V(0)$ is non-negative and r^2 is a strong solution of SDE (3.2), we have $\mathcal{F}^r_\infty = \sigma(r_t^2, t \geq 0) = \sigma(Z_t, t \geq 0)$. Recall that by Prop. 3.7, the process φ defined in Prop. 3.11 has a unique stationary measure μ.

Lemma 3.12. *Suppose that* (A4), (A5) *and* (A6) *hold. Then for any $t > 0$, $\hat{\mathcal{X}}_t$ has the law μ and is independent of $\mathcal{F}^r_\infty$. Put differently, the conditional law takes the form*

$$\mathbb{P}[\hat{\mathcal{X}}_t \in \cdot \mid \mathcal{F}^r_\infty] = \mu(\,\cdot\,), \text{ a.s., for any } t > 0.$$

Proof. Fix $t > 0$ and let $s \in (0, t)$. By Prop. 3.11 and Lemma 3.9 we have $\hat{\mathcal{X}}_t = \varphi_{\rho_s(t)}$, where φ satisfies SDE (3.8). By (e), (b) and (c) of Lemma 3.6 and Prop. 3.7, φ is strong Markov with the transition function $P_u(\mathbf{x}, \cdot)$ that does not depend on s. Hence, for $\mathfrak{A} \subseteq \mathbb{S}^{d-1}$, we find

$$\mathbb{P}[\hat{\mathcal{X}}_t \in \mathfrak{A} \mid \mathcal{F}^r_\infty] = \mathbb{E}[\mathbb{P}[\hat{\mathcal{X}}_t \in \mathfrak{A} \mid \sigma(\hat{\mathcal{X}}_s) \vee \mathcal{F}^r_\infty] \mid \mathcal{F}^r_\infty] = \mathbb{E}[P_{\rho_s(t)}(\hat{\mathcal{X}}_s, \mathfrak{A}) \mid \mathcal{F}^r_\infty], \tag{3.17}$$

as $\varphi_{\rho_s(t)}$ depends on $\mathcal{F}^r_\infty$ only through $\rho_s(t)$ and $\varphi_0 = \hat{\mathcal{X}}_s$. Crucially, (3.17) holds for any fixed time $s \in (0, t)$, and also for any random time $s = S \in (0, t)$ if S is $\mathcal{F}^r_\infty$-measurable.

By Lemma 3.10 we have $\lim_{s\downarrow 0} \rho_s(t) = \infty$. Hence, for sufficiently small s, an arbitrarily large time interval separates $\varphi_0 = \hat{\mathcal{X}}_s$ and $\varphi_{\rho_s(t)}$, and so stationarity must be attained at the latter, regardless of $\hat{\mathcal{X}}_s$. Formally, we apply the uniform ergodicity of φ in (3.11). Lemmas 3.9 and 3.10 imply that for any $u > 0$, there is an $\mathcal{F}^r_\infty$-measurable random variable $S = S(t, u)$ with $S \in (0, t)$ a.s. such that $\rho_S(t) \geq u$. By (3.11), for any $\varepsilon > 0$ there exists $u > 0$ such that $|P_{\rho_S(t)}(\varphi_0, \mathfrak{A}) - \mu(\mathfrak{A})| \leq \varepsilon,$ a.s. Hence, by (3.17) applied at the random time S, we have $|\mathbb{P}[\hat{\mathcal{X}}_t \in \mathfrak{A} \mid \mathcal{F}^r_\infty] - \mu(\mathfrak{A})| \leq \varepsilon,$ a.s. Since $\varepsilon > 0$ was arbitrary, the result follows. ☐

Proof of Theorem 1.1 in the transient case with $\mathcal{X}_0 = \mathbf{0}$. For any $k \in \mathbb{N}$ and open set $U \subset \mathbb{R}^k$, define the measurable function $F_U : (0, \infty)^k \to [0, 1]$ by $F_U(t_1, \ldots, t_k) := \mathbb{P}_\Psi[(\psi_{t_1}, \ldots, \psi_{t_k}) \in U]$, where the law $\mathbb{P}_\Psi[\cdot]$ is defined in Prop. 3.7. By Lemma 3.9, Prop. 3.11 and Lemma 3.12 we have $\mathbb{P}[(\hat{\mathcal{X}}_{t_1}, \ldots, \hat{\mathcal{X}}_{t_k}) \in U \mid \mathcal{F}^r_\infty] = F_U(\rho_s(t_1), \ldots, \rho_s(t_k))$ a.s. for $0 < s < t_1 < \cdots < t_k$. Hence $\mathbb{P}[(\hat{\mathcal{X}}_{t_1}, \ldots, \hat{\mathcal{X}}_{t_k}) \in U] = \mathbb{E}\, F_U(\rho_s(t_1), \ldots, \rho_s(t_k))$. Therefore the finite-dimensional distributions of $(\hat{\mathcal{X}}_t, t > 0)$ are uniquely determined by $\mathbb{P}_\Psi[\,\cdot\,]$ and the law of r. Moreover, by Lemma 3.2, the law of $(\|\mathcal{X}\|, \hat{\mathcal{X}})$, and hence of $\mathcal{X}$, is uniquely determined by $\mathrm{BES}^V(0)$ and $\mathbb{P}_\Psi[\,\cdot\,]$. The uniqueness in law of (3.1) implies that $\mathcal{X}$ is strong Markov and Theorem 1.1 follows in the transient case. ☐

3.6 Proof of Theorem 1.1 in the recurrent case: rapid spinning of $\hat{\mathcal{X}}$

In this section we assume $V \in (1,2)$ and $U = 1$. Hence, by Lemma 3.2, $r = \|\mathcal{X}\|$ is $\mathrm{BES}^V(0)$ where $\mathcal{X}$ is a solution of SDE (3.1). We recall briefly the necessary elements of excursion theory (see [23, Ch XII], [3, Ch IV] as a general reference). Since 0 is regular and instantaneous for r, there exists Markov local time $L = (L_t, t \geq 0)$ at 0. By [24, Prop. XI.1.11], up to a constant factor, L is a time-change of the local time at 0 of a Brownian motion, where the time-change is a constant multiple of $(\int_0^t r_u^{-2(V-1)}\mathrm{d}u; t \geq 0)$. Hence, by [5, Thm 2.4], $\lim_{t\uparrow\infty} L_t = \infty$ $\mathbb{P}$-a.s. Let $L_\lambda^{-1} := \inf\{t \geq 0 : L_t > \lambda\}$ (for $\lambda \geq 0$) be the right-continuous inverse of L and $L_{\lambda^-}^{-1} := \lim_{\kappa\uparrow\lambda} L_\kappa^{-1}$ (for $\lambda > 0$), $L_{0^-}^{-1} := 0$. The process $(L_\lambda^{-1}, \lambda \geq 0)$ is a subordinator (i.e. a Lévy process with non-decreasing paths). Furthermore, as L tends to infinity, L^{-1} is not killed: $\mathbb{P}[L_\lambda^{-1} \in \mathbb{R}_+ \forall \lambda \in \mathbb{R}_+] = 1$. Define the (countable) set of jump times by $\Lambda^r := \{\lambda \geq 0 : L_{\lambda^-}^{-1} < L_\lambda^{-1}\}$, set $\tau_\lambda^r := L_\lambda^{-1} - L_{\lambda^-}^{-1}$ and note that both L_λ^{-1} and $L_{\lambda^-}^{-1}$ are stopping times for any $\lambda \in \mathbb{R}_+$. For any $w \in \mathcal{C}_d = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$, let $\tau_{\mathbf{0}}(w) := \inf\{t > 0 : w(t) = \mathbf{0}\}$ ($\inf\emptyset = \infty$) and define $\mathcal{E}_d := \{w \in \mathcal{C}_d : 0 < \tau_{\mathbf{0}}(w) < \infty \text{ and } w(t) = \mathbf{0} \text{ for all } t \notin (0, \tau_{\mathbf{0}}(w))\}$ with the topology induced by the compact-open topology [6, §XII.1] on $\mathcal{C}_d$. Let δ_d be the zero function in $\mathcal{C}_d$. Since 0 is recurrent for the strong Markov process r, by [3, Ch IV, Thm 10(i)], the point process $e^r = (e_\lambda^r, \lambda \geq 0)$ with values in $\mathcal{E}_1 \cup \{\delta_1\}$, defined by $e_\lambda^r(t) := r_{L_{\lambda^-}^{-1}+t}\mathbf{1}\{t \leq \tau_\lambda^r\}$ (resp. $e_\lambda^r = \delta_1$) if $\lambda \in \Lambda^r$ (resp. $\lambda \notin \Lambda^r$), is a Poisson point process (PPP) with excursion measure μ_r on $\mathcal{E}_1$.

3.6.1 Marked Bessel excursions

Pick $a \in (0,\infty)$ and let $t \wedge a := \min(t,a), t \vee a := \max(t,a)$ for any $t \in \mathbb{R}$. For any $w \in \mathcal{E}_1$ satisfying $\tau_0(w) > a$, define $\varrho_w^a : (0, \tau_0(w)) \to \mathbb{R}$ by the formula

$$\varrho_w^a(t) := \operatorname{sgn}(t-a)\int_{t\wedge a}^{t\vee a} w(u)^{-2}\mathrm{d}u, \qquad t \in (0, \tau_0(w)). \tag{3.18}$$

Let $\mathcal{E}_1^{(a)} := \{w \in \mathcal{E}_1 : w \geq 0, \tau_0(w) > a \text{ and } \lim_{t\uparrow\tau_0(w)} \varrho_w^a(t) = -\lim_{t\downarrow 0}\varrho_w^a(t) = \infty\}$ and, for $d \in \mathbb{N}\setminus\{1\}$, define the set $\mathcal{E}_d^{(a)} := \{w \in \mathcal{E}_d : \|w\| \in \mathcal{E}_1^{(a)}\}$ and the map $\Phi_a : \mathcal{E}_1^{(a)} \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1}) \to \mathcal{E}_d^{(a)}$,

$$\Phi_a(w,\theta)(t) := \begin{cases} w(t)\cdot\theta\circ\varrho_w^a(t) & t \in (0,\tau_0(w)), \\ \mathbf{0} & t \in \mathbb{R}_+\setminus(0,\tau_0(w)). \end{cases}$$

The topology on $\mathcal{E}_d^{(a)}$ is induced by the compact-open topology on $\mathcal{C}_d$ [6, §XII.1]. Hence the Borel σ-algebra on $\mathcal{E}_d^{(a)}$ is generated by $\pi_t : \mathcal{E}_d^{(a)} \to \mathbb{R}^d$, $\pi_t(w) := w(t)$, for any $t \in \mathbb{R}_+$ [4, p. 57].

Lemma 3.13. *The following statements hold for any fixed $a \in (0,\infty)$.*

(i) For $w \in \mathcal{E}_1^{(a)}$, $\varrho_w^a : (0,\tau_0(w)) \to \mathbb{R}$ is continuous, increasing and $c_w^a : \mathbb{R} \to (0,\tau_0(w))$, given by $c_w^a(u) := \inf\{t \in (0,\tau_0(w)) : \varrho_w^a(t) \geq u\}$, is continuous, increasing and $c_w^a(0) = a$.

(ii) Pick $b \in (0,a)$, $w \in \mathcal{E}_1^{(a)}$ and let $I_b^a(w) := \varrho_w^b(t) - \varrho_w^a(t)$, $t \in (0,\tau_0(w))$. Then $I_b^a(w) > 0$ does not depend on t, satisfies $c_w^a(u) = c_w^b(u + I_b^a(w))$ for all $u \in \mathbb{R}$ and $\lim_{b\to 0} I_b^a(w) = \infty$.

(iii) $\Phi_a : \mathcal{E}_1^{(a)} \times \mathcal{C}(\mathbb{R},\mathbb{S}^{d-1}) \to \mathcal{E}_d^{(a)}$ is a Borel isomorphism, i.e. Φ_a is a bijection with inverse given by $\Phi_a^{-1}(w) = (\|w\|, w\circ c_{\|w\|}^a/\|w\circ c_{\|w\|}^a\|)$, $w \in \mathcal{E}_d^{(a)}$, and both Φ_a and Φ_a^{-1}

are Borel measurable. Moreover, for any $s \in \mathbb{R}$, the map $\mathcal{E}_d^{(a)} \to \mathbb{R}_+$, $w \mapsto c^a_{\|w\|}(s)$, is continuous.

(iv) Define the set $\Upsilon_d^{(a)} := \{(b, w) \in (a, \infty) \times \mathcal{E}_d^{(a)} : w \in \mathcal{E}_d^{(b)}\}$ for any $d \in \mathbb{N}$. Then the map $Q_a : \Upsilon_1^{(a)} \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1}) \to \mathcal{E}_1^{(b)} \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$, $Q_a(b, w, \theta) := (w, \theta(\,\cdot\, + I^b_a(w)))$, is continuous and the equality $\Phi_b^{-1}(w) = Q_a(b, \Phi_a^{-1}(w))$ holds for any $(b, w) \in \Upsilon_d^{(a)}$.

(v) The map $\{(b, b', w) \in (0, \infty)^2 \times \mathcal{E}_1 : w \in \mathcal{E}_1^{(b \vee b')}\} \to \mathbb{R}$, $(b, b', w) \mapsto \varrho^{b'}_w(b)$, is continuous.

Remark 3.14. (a) The maps Φ_a and Φ_a^{-1} in Lemma 3.13(iii) are homeomorphisms. The proof of this fact is more complicated than that of Lemma 3.13(iii) and is omitted as it is not used.
(b) The topology on $\Upsilon_d^{(a)}$ is induced by $(a, \infty) \times \mathcal{E}_d^{(a)}$. Parts (iii) and (iv) of Lemma 3.13 imply that the map $(b, w) \mapsto \Phi_b^{-1}(w)$, defined on $\Upsilon_d^{(a)}$, is measurable. The map in (v) is measurable.

Proof of Lemma 3.13. Since $w(u) > 0$ for all $u \in (0, \tau_0(w))$, (i) holds. Note that $\mathcal{E}_1^{(a)} \subset \mathcal{E}_1^{(b)}$ and $I^a_b(w) = \int_b^a 1/w(u)^2 \mathrm{d}u$. Part (ii) follows by the representation of c^a_w from (i) and the definition of $\mathcal{E}_1^{(a)}$.

For part (iii), note that $\tau_0(w) = \tau_{\mathbf{0}}(\Phi_a(w, \theta))$ for all $w \in \mathcal{E}_1^{(a)}$ and $\theta \in \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$. Since θ is bounded and w is continuous and equals 0 on $\mathbb{R}_+ \setminus (0, \tau_0(w))$, both Φ_a and its inverse are well-defined. Since the σ-algebra on $\mathcal{E}_d^{(a)}$ is generated by the projections, the map Φ_a is Borel measurable if and only if $\pi_t \circ \Phi_a$ is a measurable map into $\mathbb{R}^d$ for every $t \in \mathbb{R}_+$. Since, for any measurable set A in $\mathbb{R}^d$, $(\pi_0 \circ \Phi_a)^{-1}(A)$ is either empty or the whole space we may assume $t > 0$. Then, $(\pi_t \circ \Phi_a)^{-1}(\{\mathbf{0}\}) = (\mathcal{E}_1^{(a)} \setminus \{w \in \mathcal{E}_1^{(a)} : w(t) > 0\}) \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$ is clearly measurable. It is therefore sufficient to prove that $(\pi_t \circ \Phi_a)^{-1}(B)$ is open for any ball B centred at $b \in \mathbb{R}^d$ of radius $\varepsilon' \in (0, \|b\|)$. Pick $(w, \theta) \in (\pi_t \circ \Phi_a)^{-1}(B)$ and set $\varepsilon := (\varepsilon' - \|\Phi_a(w, \theta)(t) - b\|)/2 > 0$. Then $I_w := \inf_{s \in [t \wedge a, t \vee a]} w(s) > 0$. In particular, $[t \wedge a, t \vee a] \subset (0, \tau_0(w))$. Define $S_w := \sup_{s \in [t \wedge a, t \vee a]} w(s)$. There exists $\delta_0 \in (0, 1)$ such that if $|\varrho^a_w(t) - s| < \delta_0$ then $\|\theta(\varrho^a_w(t)) - \theta(s)\| < \varepsilon/(3S_w + 3)$. Assume now that $t \neq a$ and pick $\delta \in (0, 1)$ smaller than $\min\{\varepsilon/3, I_w/2, \delta_0 I_w^4 (4(2S_w + 1)|a - t|)^{-1}\}$. Define the compact $K_1 := [t \wedge a, t \vee a] \subset \mathbb{R}_+$ (resp. $K_2 := [\varrho^a_w(t) - 1, \varrho^a_w(t) + 1] \subset \mathbb{R}$), $\varepsilon_1 := \delta$ (resp. $\varepsilon_2 := \varepsilon/(3S_w + 3)$) and the neighbourhood $N_{\varepsilon_1}(K_1) := \{u \in \mathcal{E}_1^{(a)} : \sup_{s \in K_1} |w(s) - u(s)| < \varepsilon_1\}$ (resp. $N_{\varepsilon_2}(K_2) := \{\phi \in \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1}) : \sup_{s \in K_2} \|\theta(s) - \phi(s)\| < \varepsilon_2\}$) of w (resp. θ) in $\mathcal{E}_1^{(a)}$ (resp. $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$). Pick $(u, \phi) \in N_{\varepsilon_1}(K_1) \times N_{\varepsilon_2}(K_2)$ and note that $u(s) > I_w - \delta > I_w/2$ for all $s \in K_1$. Hence, by (3.18), we have $|\varrho^a_w(t) - \varrho^a_u(t)| \leq 4(2S_w + 1)|a - t| I_w^{-4} \delta < \delta_0 < 1$, implying $u(t)\|\theta(\varrho^a_w(t)) - \theta(\varrho^a_u(t))\| < \varepsilon/3$ and $\varrho^a_u(t) \in K_2$. Hence $u(t)\|\theta(\varrho^a_u(t)) - \phi(\varrho^a_u(t))\| < \varepsilon/3$ and the following inequalities hold

$$
\begin{aligned}
\|\Phi_a(w, \theta)(t) - \Phi_a(u, \phi)(t)\| &\leq |w(t) - u(t)| \\
&\quad + u(t)(\|\theta(\varrho^a_w(t)) - \theta(\varrho^a_u(t))\| + \|\theta(\varrho^a_u(t)) - \phi(\varrho^a_u(t))\|) \\
&< \varepsilon.
\end{aligned}
$$

Thus $\|\Phi_a(u, \phi)(t) - b\| \leq \varepsilon + \|\Phi_a(w, \theta)(t) - b\| < \varepsilon'$, implying $N_{\varepsilon_1}(K_1) \times N_{\varepsilon_2}(K_2) \subset (\pi_t \circ \Phi_a)^{-1}(B)$ and hence that $\pi_t \circ \Phi_a$ is measurable for $t \neq a$. If $t = a$, we have $\varrho^a_u(t) = 0$ for all $u \in \mathcal{E}_1^{(a)}$. Hence $(u, \phi) \in \mathcal{E}_1^{(a)} \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$, such that $|w(t) - u(t)| < (w(t) \wedge \varepsilon)/2$ and $\|\theta(0) - \phi(0)\| < 2\varepsilon/w(t)$, satisfies $\Phi_a(u, \phi)(t) \in B$ (where $(w, \theta), B, \varepsilon$ are as above) and the measurability of $\pi_t \circ \Phi_a$ follows.

Due to the product structure of the image, the map Φ_a^{-1} is measurable if $\mathcal{E}_d^{(a)} \to \mathcal{C}(\mathbb{R}, \mathbb{R}^d \setminus \{\mathbf{0}\})$, $w \mapsto w \circ c^a_{\|w\|}$, is measurable, which is equivalent to $g_s : \mathcal{E}_d^{(a)} \to \mathbb{R}^d \setminus \{\mathbf{0}\}$, $g_s(w) := w(c^a_{\|w\|}(s))$, being measurable for every $s \in \mathbb{R}$. The map g_s is in fact continuous.

If $s = 0$, then $g_s(w) = w(a)$ is an evaluation at a, which is continuous in the compact-open topology. If $s \neq 0$, let B denote an open ball centred at $b \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ of radius $\varepsilon' \in (0, \|b\|)$, pick $w \in g_s^{-1}(B)$ and let $\varepsilon := (\varepsilon' - \|g_s(w) - b\|)/2$. Define $t := c^a_{\|w\|}(s) \neq a$ and let $S_{\|w\|} := \sup_{p\in[t\wedge a, t\vee a]} \|w(p)\|$, $I_{\|w\|} := \inf_{p\in[t\wedge a,t\vee a]} \|w(p)\|$, $K_1 := [0, \tau_{\mathbf{0}}(w)]$ and $\bar{S}_{\|w\|} := \sup_{p\in K_1} \|w(p)\|$. There exists $\delta_0 \in (0,1)$ such that $[t-\delta_0, t+\delta_0] \subset (0, \tau_{\mathbf{0}}(w))$ and $\forall x \in [t-\delta_0, t+\delta_0]$ we have $\|w(x) - w(t)\| < \varepsilon/2$. Choose $\delta \in (0,1)$ smaller than $\min\{\varepsilon/2, I_{\|w\|}/2, \delta_0 I^4_{\|w\|}(4(2S_{\|w\|}+1)|a-t|(\bar{S}_{\|w\|}+1)^2)^{-1}\}$, and pick arbitrary u in $N_\delta(K_1) := \{u \in \mathcal{E}^{(a)}_d : \sup_{p\in K_1} \|w(p) - u(p)\| < \delta\}$. Then $|\varrho^a_{\|w\|}(t) - \varrho^a_{\|u\|}(t)| < \delta_0/(\bar{S}_{\|w\|}+1)^2$ and hence $\varrho^a_{\|u\|}(t) \in K_2 := [\varrho^a_{\|w\|}(t)-1, \varrho^a_{\|w\|}(t)+1]$. As $s = \varrho^a_{\|w\|}(t)$, $c^a_{\|w\|}(s) = c^a_{\|u\|}(\varrho^a_{\|u\|}(t))$ and $\sup\{\|u(c^a_{\|u\|}(q))\|^2 : q \in K_2\} \leq (\bar{S}_{\|w\|}+1)^2$, we have

$$|c^a_{\|w\|}(s) - c^a_{\|u\|}(s)| \leq |\varrho^a_{\|w\|}(t) - \varrho^a_{\|u\|}(t)|(\bar{S}_{\|w\|}+1)^2 < \delta_0. \tag{3.19}$$

Hence, $\|g_s(w)-g_s(u)\| \leq \|w(c^a_{\|w\|}(s))-w(c^a_{\|u\|}(s))\|+\|w(c^a_{\|u\|}(s))-u(c^a_{\|u\|}(s))\| \leq \varepsilon/2+\varepsilon/2 = \varepsilon$ and the inclusion $N_\delta(K_1) \subset g_s^{-1}(B)$, implying the continuity of g_s, follows. Since δ_0 could be arbitrarily small, the bound in (3.19) also implies the continuity of $w \mapsto c^a_{\|w\|}(s)$.

The equality in part (iv) follows from (ii) and (iii). What remains to be proved is that $(b, w, \theta) \mapsto \theta(\,\cdot\, + I^b_a(w))$ is continuous at an arbitrary point $(b_0, w_0, \theta_0) \in \Upsilon^{(a)}_1 \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$. Since for any $t \in \mathbb{R}$ we have $\|\theta_0(t + I^{b_0}_a(w_0)) - \theta(t + I^b_a(w))\| \leq \|\theta_0(t + I^{b_0}_a(w_0)) - \theta_0(t + I^b_a(w_0))\| + \|\theta_0(t + I^b_a(w_0)) - \theta_0(t + I^b_a(w))\| + \|\theta_0(t + I^b_a(w)) - \theta(t + I^b_a(w))\|$, the uniform continuity of θ_0 on any compact, together with the proximity of (b_0, w_0) and (b, w), yields a uniform control on compacts of the first two terms. The third term is controlled by the proximity of θ_0 and θ in $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$. The estimates, analogous to the ones in the proof of (iii), are omitted.

Pick (b_0, b_0', w_0) in the domain of the map in (v) and let (b, b', w) be an arbitrary element close to it. If $b_0 = b_0'$, then $\varrho^{b_0'}_w(b_0) = 0$ and $w_0(b_0) > 0$. Then b and b' must be very close to b_0 (and hence each other) and w must be positive in the neighbourhood of b_0. Hence the continuity of the map in (v) follows. If $b_0 < b_0'$, then $-\varrho^{b_0'}_w(b_0) = \int_{b_0}^{b_0'} \mathrm{d}u/w_0^2(u)$ and w_0 is bounded away from zero on compact interval $K \supset [b_0, b_0']$. Moreover, we may assume that $b < b'$, $K \supset [b, b']$ and that w is uniformly close to w_0 on K. Hence $|\varrho^{b_0'}_w(b_0) - \varrho^{b'}_w(b)|$ is arbitrarily small and the continuity follows. The remaining case $b_0' < b_0$ is analogous. □

Remark 3.15. The continuity of the functions g_s, $s \in \mathbb{R}$, in the proof of Lemma 3.13(iii) above does not imply the continuity of the map Φ_a^{-1}.

Define $\mathcal{E}^+_d := \cup_{a>0} \mathcal{E}^{(a)}_d \subset \mathcal{E}_d$ (for $d \in \mathbb{N}$) with the topology induced by that of $\mathcal{C}_d$.

Proposition 3.16. *The excursion measure of r satisfies $\mu_r(\mathcal{E}_1 \setminus \mathcal{E}^+_1) = 0$. Let $\mathbb{P}_\Psi$ be the law on $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$ from Prop. 3.7. Then there exists a unique σ-finite atomless Borel measure ν on $\mathcal{E}^+_d$, satisfying $\nu(A \cap \mathcal{E}^{(a)}_d) = \mu_r \otimes \mathbb{P}_\Psi[\Phi_a^{-1}(A \cap \mathcal{E}^{(a)}_d)]$ for all $a > 0$ and Borel measurable $A \subseteq \mathcal{E}^+_d$.*

Remark 3.17. By Prop. 3.16, e^r is a PPP on $\mathcal{E}^+_1 \cup \{\delta_1\}$ and ν induces a PPP on $\mathcal{E}^+_d \cup \{\delta_d\}$.

Proof of Proposition 3.16. In order to establish $\mu_r(\mathcal{E}_1 \setminus \mathcal{E}^+_1) = 0$, note that by [23], the excursion measure μ_r has the following representation: any excursion e^r_λ has a finite maximum and this maximum is attained at a unique time. Furthermore, conditional on the maximum being at some level $M > 0$, the excursion has the same law as the path formed by taking two independent $\mathrm{BES}^{4-\delta}(0)$ processes, both run up until their first hitting time of the level M, and placing them end-to-end. Since $2 < 4 - \delta < 3$, by Lemma 3.10, any excursion in the support of μ_r is in $\mathcal{E}^+_1$.

Let $\Psi = (\Psi^\lambda, \lambda \geq 0)$ be a family of independent stationary diffusions $\Psi^\lambda = (\Psi^\lambda_t, t \in \mathbb{R})$ with the law $\mathbb{P}_\Psi$ from Prop. 3.7. Assume that r is independent of Ψ. By the Marking and Mapping theorems of [17] (the latter applies since Φ_a is measurable and bijective by Lemma 3.13(iii)), the point process $e^{r,\Psi,a} = (e^{r,\Psi,a}_\lambda, \lambda \geq 0)$, defined by $e^{r,\Psi,a}_\lambda := \delta_d$, if $\tau^r_\lambda \leq a$, and $e^{r,\Psi,a}_\lambda := \Phi_a(e^r_\lambda, \Psi^\lambda)$, if $\tau^r_\lambda > a$, is a PPP in $\mathcal{E}^{(a)}_d \cup \{\delta_d\}$ with excursion measure $\mu_r \otimes \mathbb{P}_\Psi[\Phi_a^{-1}(\,\cdot\,)]$ on $\mathcal{E}^{(a)}_d$ of finite total mass $\mu_r \otimes \mathbb{P}_\Psi[\Phi_a^{-1}(\mathcal{E}^{(a)}_d)] = \mu_r(\mathcal{E}^{(a)}_1) < \infty$. Moreover, by [17, p. 13], $\mu_r \otimes \mathbb{P}_\Psi[\Phi_a^{-1}(\,\cdot\,)]$ is atomless. Hence any measure ν satisfying the identity in the proposition for all $a \in (0,\infty)$ is also atomless, σ-finite and unique.

The proposition now follows from the claim that $\mu_r \otimes \mathbb{P}_\Psi[\Phi_a^{-1}(A)] = \mu_r \otimes \mathbb{P}_\Psi[\Phi_b^{-1}(A)]$ for any $0 < b < a$ and measurable $A \subseteq \mathcal{E}^{(a)}_d$.

It remains to establish this claim. Consider $Q : \mathcal{E}^{(a)}_1 \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1}) \to \mathcal{E}^{(a)}_1 \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$, $Q(w,\theta) := Q_b(a,w,\theta)$, where Q_b is defined in Lemma 3.13(iv). Hence $Q = \Phi_a^{-1} \circ \Phi_b|_{\mathcal{E}^{(a)}_1 \times \mathcal{C}(\mathbb{R},\mathbb{S}^{d-1})}$ is a Borel isomorphism. It suffices to show that Q is measure preserving, i.e. $\mu_r \otimes \mathbb{P}_\Psi[B] = \mu_r \otimes \mathbb{P}_\Psi[Q(B)]$ for any measurable $B \subseteq \mathcal{E}^{(a)}_1 \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$. The measure $(\mu_r/\mu_r(\mathcal{E}^{(b)}_1)) \otimes \mathbb{P}_\Psi$, restricted to $\mathcal{E}^{(b)}_1 \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$, is the probability law of the random element $(X,Y) := (e^r_{\lambda_b}, \Psi^{\lambda_b})$, where λ_b is the time of the first jump of size greater than b of the subordinator L^{-1}. In particular, we need to show $\mathbb{P}[(X,Y) \in B] = \mathbb{P}[Q^{-1}(X,Y) \in B]$. Since $Q^{-1}(w,\theta) = (w, \theta(\,\cdot\, - I^a_b(w)))$, $I^a_b(w)$ depends only on w by Lemma 3.13(ii) and, by Prop. 3.7, the process Y is stationary, it holds that $\mathbb{P}[(X,Y) \in B \mid \sigma(X)] = \mathbb{P}[Q^{-1}(X,Y) \in B \mid \sigma(X)]$, implying the claim. □

3.6.2 Proof of Theorem 1.1 in the recurrent case

Let $(\mathcal{X}, W)$ be a solution of SDE (3.1) with $\mathcal{X}_0 = \mathbf{0}$, adapted to $(\mathcal{F}_t, t \geq 0)$. Since we are only interested in the law of the solution, we may assume that we are in the canonical setting, i.e. the probability space is $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^n)$ (for some $n \in \mathbb{N}$) and the filtration satisfies the usual conditions with respect to the probability measure $\mathbb{P}$ on Ω. Define the point process $e^{\mathcal{X}} = (e^{\mathcal{X}}_\ell, \ell \geq 0)$ of excursions of $\mathcal{X}$ away from $\mathbf{0}$ by $e^{\mathcal{X}}_\ell := \delta_d$ if $\ell \in \mathbb{R}_+ \setminus \Lambda^r$, and $e^{\mathcal{X}}_\ell : \mathbb{R}_+ \to \mathbb{R}^d$, where

$$e^{\mathcal{X}}_\ell(u) := \begin{cases} \mathcal{X}_{L^{-1}_{\ell-}+u} & u \in (0, \tau^r_\ell), \\ \mathbf{0} & u \in \mathbb{R}_+ \setminus (0, \tau^r_\ell), \end{cases} \tag{3.20}$$

if $\ell \in \Lambda^r$ (the notation introduced earlier in Section 3.6 will be used throughout Section 3.6.2). The point process $\|e^{\mathcal{X}}\| = (\|e^{\mathcal{X}}_\ell\|, \ell \geq 0)$ with excursions $\|e^{\mathcal{X}}_\ell(u)\| = r_{L^{-1}_{\ell-}+u} \mathbf{1}\{u \leq \tau^r_\ell\}$, $u \in \mathbb{R}_+$, for any $\ell \in \Lambda^r$, is clearly equal to the PPP e^r defined above. Since $\mathcal{X}_t = \mathbf{0}$ if and only if $r_t = 0$, $e^{\mathcal{X}}$ takes values in $\mathcal{E}^+_d \cup \{\delta_d\}$. The key step in the proof of Theorem 1.1 is to characterize $e^{\mathcal{X}}$: this will establish uniqueness in law of $\mathcal{X}$ (see Corollary 3.23), and, at the same time, show that $e^{\mathcal{X}}$ is a PPP with excursion measure from Prop. 3.16 (Corollary 3.24).

For the rest of the section, fix an arbitrary $(\mathcal{F}_t)$-stopping time τ with $\mathbb{P}[\tau < \infty] = 1$. Then $L^{-1}_{L_\tau}$ is an $(\mathcal{F}_t)$-stopping time. Define $\tilde{r} = (\tilde{r}_u, u \geq 0)$ by $\tilde{r}_u := r_{L^{-1}_{L_\tau}+u}$. By the strong Markov property of r, the process $\tilde{r}$ is strong Markov with respect to the filtration $(\mathcal{F}_{L^{-1}_{L_\tau}+u}, u \geq 0)$, has the same law as r and is independent of $\mathcal{F}_{L^{-1}_{L_\tau}}$. The (Markov) local time $(\tilde{L}_u, u \geq 0)$ of $\tilde{r}$ at 0 satisfies $\tilde{L}_u = L_{L^{-1}_{L_\tau}+u} - L_\tau$. The inverse local time $\tilde{L}^{-1} = (\tilde{L}^{-1}_\mu, \mu \geq 0)$ is a subordinator satisfying $\tilde{L}^{-1}_\mu = L^{-1}_{L_\tau+\mu} - L^{-1}_{L_\tau}$, independent of $\mathcal{F}_{L^{-1}_{L_\tau}}$. Pick $a > 0$ and define recursively the stopping times: $\mu^0_a := 0$ and $\mu^n_a := \inf\{t > \mu^{n-1}_a : \tau^r_{t+L_\tau} > a\}$ for any $n \in \mathbb{N}$. Here $\tau^r_{t+L_\tau} = \tau^{\tilde{r}}_t := \tilde{L}^{-1}_t - \tilde{L}^{-1}_{t-}$ is the jump of the subordinator $\tilde{L}^{-1}$ and μ^n_a is the epoch of local time corresponding to the n-th excursion of $\tilde{r}$, lasting longer than a. For any $u \in \mathbb{R}_+$, the equality $e^r_{L_{u+L^{-1}_{L_\tau}}} = e^{\tilde{r}}_{\tilde{L}_u}$ holds, where

$(e^{\tilde{r}}_\mu, \mu \geq 0)$ is given by $e^{\tilde{r}}_\mu := \tilde{r}_{\tilde{L}^{-1}_{\mu-}+u}\mathbf{1}\{u \leq \tau^{\tilde{r}}_\mu\}$, $u \in \mathbb{R}_+$. Finally, for any $b \in (0, a)$, let $N_b(t) := \sup\{m \in \mathbb{N} : \tilde{L}^{-1}_{\mu^m_b -} < t\}$ (with convention $\sup\emptyset := 0$) be the number of excursions of $\tilde{r}$ started before time $t \in \mathbb{R}_+$ with length at least b. Note that all the random elements defined in this paragraph depend on the choice of the stopping time τ. The next result is the basis of our characterization of $e^{\mathcal{X}}$.

Theorem 3.18. *Suppose that* (A4)*,* (A5) *and* (A6) *hold, with $U = 1$ and $V \in (1, 2)$. For any $a > 0$, $n \in \mathbb{N}$ and finite $(\mathcal{F}_t)$-stopping time τ, the regular conditional distribution of the random element $e^{\mathcal{X}}_{L_\tau+\mu^n_a}$ (defined in* (3.20) *with $\ell = L_\tau + \mu^n_a$) in $\mathcal{E}^{(a)}_d$, given $\mathcal{F}_{L^{-1}_{L_\tau}}$, takes the form*

$$\mathbb{P}[e^{\mathcal{X}}_{L_\tau+\mu^n_a} \in \cdot \mid \mathcal{F}_{L^{-1}_{L_\tau}}] = \mu_r \otimes \mathbb{P}_\Psi[\Phi^{-1}_a(\,\cdot\,)]/\mu_r(\mathcal{E}^{(a)}_1) \qquad \textit{a.s.}$$

Here the law $\mathbb{P}_\Psi$ on $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$ is defined in Prop. 3.7 and μ_r is the excursion measure of the PPP e^r. In particular, the excursion $e^{\mathcal{X}}_{L_\tau+\mu^n_a}$ is independent of $\mathcal{F}_{L^{-1}_{L_\tau}}$ and its law on $\mathcal{E}^{(a)}_d$, $\mu_r \otimes \mathbb{P}_\Psi[\Phi^{-1}_a(\,\cdot\,)]/\mu_r(\mathcal{E}^{(a)}_1)$, depends neither on $n \in \mathbb{N}$ nor on the stopping time τ.

Remark 3.19. Theorem 3.18 would follow trivially if we knew that $\mathcal{X}$ was strong Markov. However, this cannot be assumed *a priori*. Once the uniqueness in law of SDE (3.1) has been established, the strong Markov property of $\mathcal{X}$ follows.

As $e^{\mathcal{X}}_{L_\tau+\mu^n_a} \in \mathcal{E}^{(a)}_d$, we can define the process $\theta^{a,n}$ by $(e^r_{L_\tau+\mu^n_a}, \theta^{a,n}) := \Phi^{-1}_a(e^{\mathcal{X}}_{L_\tau+\mu^n_a})$; then $\theta^{a,n}$ has paths in $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$. The key step in the proof of Theorem 3.18 is given by the following lemma.

Lemma 3.20. *Under assumptions (and notation) of Theorem 3.18, the regular conditional distribution of $\theta^{a,n}$ takes the form $\mathbb{P}[\theta^{a,n} \in \cdot \mid \mathcal{F}_{L^{-1}_{L_\tau}} \vee \mathcal{F}^r_\infty] = \mathbb{P}_\Psi[\,\cdot\,]$ a.s. (recall $\mathcal{F}^r_\infty = \sigma(r_t, t \geq 0)$).*

Proof. Since $\mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$ is Polish, the regular conditional distribution $\mathbb{P}[\theta^{a,n} \in \cdot \mid \mathcal{F}_{L^{-1}_{L_\tau}} \vee \mathcal{F}^r_\infty]$ exists. Moreover, as every trajectory of $\theta^{a,n}$ is continuous, it is sufficient to prove that $\mathbb{P}$-a.s. the finite-dimensional distributions at rational times coincide with those of $\mathbb{P}_\Psi$. Since the set of all finite subsets of the rationals is countable and the Borel σ-algebra on $\mathbb{S}^{d-1}$ is generated by a countable family of open balls, by a diagonalization argument it suffices to prove that the finite-dimensional distributions at a given set of (rational) times (evaluated on the products of the finite intersections of generating sets) coincide $\mathbb{P}$-a.s. We establish this in two steps. First, we show that the process $(\theta^{a,n}_t, t \geq 0)$ solves SDE (3.8), started at $\theta^{a,n}_0 = \hat{\mathcal{X}}_{a+L^{-1}_{(L_\tau+\mu^n_a)-}}$ and driven by a Brownian motion B independent of $\mathcal{F}^r_\infty$. Second, we use this to prove the equality of the finite-dimensional marginals of the two measures.

Since, for $s \in \mathbb{R}_+$, the map $w \mapsto c^a_w(s)$ on $\mathcal{E}^{(a)}_d$ is continuous (and hence measurable) by Lemma 3.13(iii), we may define a non-negative random variable $\eta_a(s) := c^a_{e^r_{L_\tau+\mu^n_a}}(s) + L^{-1}_{(L_\tau+\mu^n_a)-}$. Since $\eta_a(0) - L^{-1}_{L_\tau}$ is the first time an excursion of $\tilde{r}$ lasts longer than a, after $n-1$ such excursions have occurred, $\eta_a(0)$ is a finite $(\mathcal{F}_t)$-stopping time. The definition of c^a_w implies that $\eta_a(s) = \eta_a(0) + \inf\{t \in (0, \infty) : \int_{\eta_a(0)}^{\eta_a(0)+t} r_u^{-2}\mathrm{d}u \geq s\}$ is also an $(\mathcal{F}_t)$-stopping time for any $s > 0$. In fact for $0 \leq s \leq u$ it holds that $\eta_a(s) \leq \eta_a(u) < L^{-1}_{L_\tau+\mu^n_a}$. Put differently, $(\eta_a(s), s \geq 0)$ is a stochastic time-change and we can define the filtration $(\mathcal{G}_s, s \geq 0)$ by $\mathcal{G}_s := \mathcal{F}_{\eta_a(s)}$.

Since, on the stochastic interval $(0, L^{-1}_{L_\tau+\mu^n_a} - \eta_a(0))$, the process $r^{-1}_{\eta_a(0)+\,\cdot}$ is continuous and $(\mathcal{F}_{\eta_a(0)+t})$-adapted, we can define continuous local martingales $A = (A_s; s \geq 0)$ and

$\zeta = (\zeta_s; s \geq 0)$ by

$$A_s := \int_{\eta_a(0)}^{\eta_a(s)} r_u^{-1} \mathrm{d}W_u \qquad \text{and} \qquad \zeta_s := \int_{\eta_a(0)}^{\eta_a(s)} r_u^{-1} \mathrm{d}Z_u,$$

where Z is given in (3.3). Both A and ζ are adapted to $(\mathcal{G}_s, s \geq 0)$. As in the proof of Prop. 3.11, it follows that A and ζ are $(\mathcal{G}_s)$ Brownian motions. We may then apply [24, Prop. V.1.4] and (3.3) to ζ to obtain $\zeta_s = \int_0^s (\hat{\mathcal{X}}_{\eta_a(u)})^\top r_{\eta_a(u)}^{-1} \mathrm{d}W_{\eta_a(u)}$. Similarly we get $A_s = \int_0^s r_{\eta_a(u)}^{-1} \mathrm{d}W_{\eta_a(u)}$. Since by definition $\hat{\mathcal{X}}_{\eta_a(u)} = \theta_u^{a,n}$ for all $u \in \mathbb{R}_+$, we find $\zeta_s = \int_0^s (\theta_u^{a,n})^\top \mathrm{d}A_u$ for all $s \geq 0$. Without loss of generality there exists a one-dimensional $(\mathcal{F}_t)$ Brownian motion, $\bar{\gamma} = (\bar{\gamma}_t, t \geq 0)$, independent of $(\mathcal{X}, W)$. Define a $(\mathcal{G}_s)$ Brownian motion $\gamma = (\gamma_t, t \geq 0)$ by $\gamma_s := \int_{\eta_a(0)}^{\eta_a(s)} r_u^{-1} \mathrm{d}\bar{\gamma}_u$. Then, as in the proof of Prop. 3.11, the process $B = (B_t, t \geq 0)$, $B_s := A_s - \int_0^s \theta_u^{a,n} \mathrm{d}\zeta_u + \int_0^s \theta_u^{a,n} \mathrm{d}\gamma_u$, is a d-dimensional $(\mathcal{G}_s)$ Brownian motion, independent of ζ.

We next show that B is independent of Z. Recall that $\eta_a(0)$ and $L^{-1}_{L_\tau + \mu_a^n}$ are $(\mathcal{F}_t)$-stopping times. Since $B_0 = \mathbf{0}$, B is independent of $\mathcal{G}_0 = \mathcal{F}_{\eta_a(0)}$ and hence of $(Z_s, 0 \leq s \leq \eta_a(0))$. B is measurable with respect to $\bigvee_{s \in \mathbb{R}_+} \mathcal{G}_s \subseteq \mathcal{F}_{L^{-1}_{L_\tau + \mu_a^n}}$ and hence independent of the Brownian motion $(Z_{u + L^{-1}_{L_\tau + \mu_a^n}} - Z_{L^{-1}_{L_\tau + \mu_a^n}}, u \geq 0)$. We now prove that B is independent of the stopped Brownian motion $(\bar{Z}_s, s \geq 0)$, $\bar{Z}_s := Z_{(s + \eta_a(0)) \wedge L^{-1}_{L_\tau + \mu_a^n}} - Z_{\eta_a(0)}$. Define the $\mathcal{G}_s$-local martingale $M = (M_u, u \geq 0)$, $M_u := Z_{\eta_a(u)} - Z_{\eta_a(0)}$, and note that $M_u = \int_0^u r_{\eta_a(v)} (\theta_v^{a,n})^\top \mathrm{d}A_v = \int_0^u r_{\eta_a(v)} \mathrm{d}\zeta_v$. Hence the covariation of M and B is identically equal to zero. Furthermore, the quadratic variation $[M]_u = c^a_{e^r_{L_\tau + \mu_a^n}}(u) - a$ of M converges, i.e., $[M]_\infty := \lim_{u \uparrow \infty} [M]_u = L^{-1}_{L_\tau + \mu_a^n} - \eta_a(0)$, with inverse given by $v \mapsto \varrho^a_{e^r_{L_\tau + \mu_a^n}}(a + v)$, $v \in [0, [M]_\infty)$. Since the limit $M_\infty := \lim_{u \uparrow \infty} M_u = Z_{L^{-1}_{L_\tau + \mu_a^n}} - Z_{\eta_a(0)}$ exists, we can define the processes $(M_{\varrho^a_{e^r_{L_\tau + \mu_a^n}}(a+t)}, 0 \leq t \leq [M]_\infty)$, which is independent of B by [24, Thm V.1.9]. Then noting that $M_{\varrho^a_{e^r_{L_\tau + \mu_a^n}}(a+t)} = \bar{Z}_t$ for any $t \in [0, [M]_\infty]$, we verify that B is independent of Z, and hence (by Lemma 3.2) of r.

By Lemma 3.2, the process $r^{-2}_{\eta_a(0)+\cdot}$ is a continuous semimartingale on the stochastic interval $(0, \tau^r_{L_\tau + \mu_a^n} - a)$. In particular, an analogous calculation to the one that established (3.13) implies

$$\hat{\mathcal{X}}_{\eta_a(0)+t} = \hat{\mathcal{X}}_{\eta_a(0)} + \int_{\eta_a(0)}^{\eta_a(0)+t} f(\hat{\mathcal{X}}_u) r_u^{-2} \mathrm{d}u + \int_{\eta_a(0)}^{\eta_a(0)+t} g(\hat{\mathcal{X}}_u) r_u^{-1} \mathrm{d}W_u, \quad t \in (0, \tau^r_{L_\tau + \mu_a^n} - a),$$

with f, g in (3.14). Applying the stochastic time-change $(c^a_{e^r_{L_\tau + \mu_a^n}}(u) - a, u \geq 0)$ with [24, Prop. V.1.4] and noting that $\eta_a(u) = \eta_a(0) + c^a_{e^r_{L_\tau + \mu_a^n}}(u) - a$ and $\hat{\mathcal{X}}_{\eta_a(u)} = \theta_u^{a,n}$ for all $u \in \mathbb{R}_+$, implies that $(\theta_u^{a,n}, u \geq 0)$ satisfies the SDE in (3.8), started at $\theta_0^{a,n} = \hat{\mathcal{X}}_{a + L^{-1}_{(L_\tau + \mu_a^n)-}}$ driven by the Brownian motion A defined above. It is easy to see from the definition of the Brownian motion B above that $\int_0^t (\sigma_{\mathrm{sy}}(\theta_u^{a,n}) - \theta_u^{a,n} (\theta_u^{a,n})^\top) \mathrm{d}B_u = \int_0^t (\sigma_{\mathrm{sy}}(\theta_u^{a,n}) - \theta_u^{a,n} (\theta_u^{a,n})^\top) \mathrm{d}A_u$ for all $t \geq 0$. Hence $(\theta_u^{a,n}, u \geq 0)$ satisfies SDE (3.8) driven by B independent of $\mathcal{F}^r_\infty$.

The second step in the proof of the lemma analyses the conditional law of $\theta^{a,n}$. The number of excursions longer than b started before the start of the n-th excursion of $\tilde{r}$ of length at least a, i.e. $N_b(\tilde{L}^{-1}_{\mu_a^n -})$, is $\mathcal{F}^r_\infty$ measurable. Fix $t \in \mathbb{R}$ and note that by Lemma 3.13(ii) we have $\lim_{b \downarrow 0} t + I_b^a(e^r_{L_\tau + \mu_a^n}) = \infty$. On the event $\{N_b(\tilde{L}^{-1}_{\mu_a^n -}) = k - 1\}$, by Lemma 3.13(ii)–(iii), it holds that $\theta_t^{a,n} = \theta^{b,k}_{t + I_b^a(e^r_{L_\tau + \mu_a^n})}$. Pick an arbitrary measurable subset $\mathfrak{A} \subseteq \mathbb{S}^{d-1}$. Then it holds that

$$\mathbb{P}[\theta_t^{a,n} \in \mathfrak{A} \mid \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r] = \sum_{k\in\mathbb{N}} \mathbf{1}\{N_b(\tilde{L}^{-1}_{\mu_a^n -}) = k-1\}\mathbb{P}[\theta^{b,k}_{t+I_b^a(e^r_{L_\tau+\mu_a^n})} \in \mathfrak{A} \mid \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r].$$

For all $b \in (0,a)$ such that $I_b^a(e^r_{L_\tau+\mu_a^n}) > -t$, the first step of the proof implies

$$|\mathbb{P}[\theta_t^{a,n} \in \mathfrak{A} \mid \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r] - \mu(\mathfrak{A})| \le \int_{\mathbb{S}^{d-1}} |P_{t+I_b^a(e^r_{L_\tau+\mu_a^n})}(x,\mathfrak{A}) - \mu(\mathfrak{A})|\mathbb{P}_b[\mathrm{d}x], \tag{3.21}$$

where $\mathbb{P}_b[\mathrm{d}x] := \sum_{k\in\mathbb{N}} \mathbf{1}\{N_b(\tilde{L}^{-1}_{\mu_a^n -}) = k-1\}\mathbb{P}[\theta_0^{b,k} \in \mathrm{d}x \mid \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r]$ is a probability measure on $\mathbb{S}^{d-1}$, P is the transition function from Prop. 3.7 and μ denotes its stationary measure. By (3.11) in Prop. 3.7, Lemma 3.13(ii) and (3.21), for any $\varepsilon > 0$ there exists $b \in (0,a)$ such that $|\mathbb{P}[\theta_t^{a,n} \in \mathfrak{A} \mid \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r] - \mu(\mathfrak{A})| \le \varepsilon$. Hence we must have $\mathbb{P}[\theta_t^{a,n} \in \mathfrak{A} \mid \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r] = \mu(\mathfrak{A}) = \mathbb{P}_\Psi[\{f \in \mathcal{C}(\mathbb{R},\mathbb{S}^{d-1}) : f(t) \in \mathfrak{A}\}]$. An analogous argument shows that finite-dimensional distributions of $\mathbb{P}_\Psi[\,\cdot\,]$ and $\mathbb{P}[\theta_t^{a,n} \in \cdot \mid \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r]$ coincide. This proves the lemma. □

Proof of Theorem 3.18. Pick an arbitrary measurable set B in $\mathcal{E}_d^{(a)}$ and define a subset $A := \Phi_a^{-1}(B)$ of $\mathcal{E}_1^{(a)} \times \mathcal{C}(\mathbb{R},\mathbb{S}^{d-1})$. A standard argument, based on the Monotone-Class Theorem, implies that the function $F_A : \mathcal{E}_1^{(a)} \to [0,1]$, given for $e \in \mathcal{E}^{(a)}$ by $F_A(e) := \int_{\mathcal{C}(\mathbb{R},\mathbb{S}^{d-1})} \mathbf{1}\{A\}(e,f)\mathbb{P}_\Psi[\mathrm{d}f]$, is measurable. Hence Lemma 3.20, the tower property and the definition of the map Φ_a^{-1} imply $\mathbb{P}[e^{\mathcal{X}}_{L_\tau+\mu_a^n} \in B \mid \mathcal{F}_{L_{L_\tau}^{-1}}] = \mathbb{P}[(e^r_{L_\tau+\mu_a^n}, \theta^{a,n}) \in A \mid \mathcal{F}_{L_{L_\tau}^{-1}}] = \mathbb{E}[F_A(e^r_{L_\tau+\mu_a^n}) \mid \mathcal{F}_{L_{L_\tau}^{-1}}]$. Since r is strong Markov, we get $\mathbb{P}[e^{\mathcal{X}}_{L_\tau+\mu_a^n} \in B \mid \mathcal{F}_{L_{L_\tau}^{-1}}] = \mathbb{E}[F_A(e^r_{L_\tau+\mu_a^n})]$. Since the law of the excursion $e^r_{L_\tau+\mu_a^n}$ is given by $\mu_r(\,\cdot\,)/\mu_r(\mathcal{E}_1^{(a)})$, the theorem follows. □

Pick $v \in (0,\infty)$ and a measurable $B \subseteq \mathbb{R}^d$. Let $B_v := \{\hat{\mathbf{y}} : \mathbf{y} \in B \setminus \{\mathbf{0}\}, \|\mathbf{y}\| = v\}$ be the intersection $B \cap (v\mathbb{S}^{d-1})$ projected onto the unit sphere. For any $b \in \mathbb{R}$, define the measurable set $\mathfrak{A}_v^b(B) := \{f \in \mathcal{C}(\mathbb{R},\mathbb{S}^{d-1}) : f(b) \in B_v\}$. The remaining step in our characterization of $e^{\mathcal{X}}$ is provided by the following result, which will enable us to describe finite-dimensional distributions.

Proposition 3.21. *Pick $k \in \mathbb{N}$ and indices $0 =: i_0 < i_1 < i_2 < \cdots < i_{k-1} < i_k$. Define $n := i_k$ and choose measurable sets $B_1, \ldots, B_n \subseteq \mathbb{R}^d$ and times $0 < u_1 < u_2 < \cdots < u_n$. For $0 \le i < j \le n$, let $F_{i,j} : (\mathbb{R}_+ \times (0,\infty))^{j-i} \to [0,1]$ be $F_{i,j}(b_p, v_p; i+1 \le p \le j) := \mathbb{P}_\Psi[\cap_{p=i+1}^j \mathfrak{A}_{v_p}^{b_p}(B_p)]$. Define $a_j := u_j - \tilde{L}^{-1}_{\tilde{L}_{u_j}-}$ for any $j \in \{1,\ldots,n\}$ (recall that $\tilde{L}$ depends on τ). Then, on the event $E_k := \{\tilde{L}_{u_{i_0+1}} = \tilde{L}_{u_{i_1}} < \tilde{L}_{u_{i_1+1}} = \tilde{L}_{u_{i_2}} < \tilde{L}_{u_{i_2+1}} = \tilde{L}_{u_{i_3}} < \cdots < \tilde{L}_{u_{i_{k-1}+1}} = \tilde{L}_{u_{i_k}}\}$, it holds that*

$$\begin{aligned}\mathbb{P}\left[e^{\mathcal{X}}_{L_\tau+\tilde{L}_{u_j}}(a_j) \in B_j \quad \text{for } j \in \{1,\ldots,n\} \,\middle|\, \mathcal{F}_{L_{L_\tau}^{-1}} \vee \mathcal{F}_\infty^r\right] \\ = \prod_{l=0}^{k-1} F_{i_l,i_{l+1}}\left(\varrho^{a_{i_l+1}}_{e^{\tilde{r}}_{\tilde{L}_{u_p}}}(a_p), e^{\tilde{r}}_{\tilde{L}_{u_p}}(a_p); i_l+1 \le p \le i_{l+1}\right).\end{aligned} \tag{3.22}$$

Remark 3.22. In (3.22), for any $p \in \{i_l+1,\ldots,i_{l+1}\}$, it holds that $\tilde{L}_{u_p} = \tilde{L}_{u_{i_l+1}}$ and hence $e^{\tilde{r}}_{\tilde{L}_{u_p}}$ refers to a single excursion. Note also that E_k depends on the sequence $i_1 < \cdots < i_k$ and not just on the index k. This information is suppressed from the notation for brevity.

Proof of Proposition 3.21. A moment's reflection reveals that $F_{i,j}$, defined in the proposition, is measurable and $E_k \in \mathcal{F}^r_\infty$. Note that a_j is $\mathcal{F}^r_\infty$-measurable and $a_j > 0$ $\mathbb{P}$-a.s. for any $j \in \{1, \dots, n\}$. Moreover, on E_k, by Remark 3.22 the triplet $(a_{i_l+1}, a_p, e^{\tilde r}_{\tilde L_{u_p}})$ is in the domain of the map in Lemma 3.13(v) for all $l \in \{0, \dots, k-1\}$ and $p \in \{i_l+1, \dots, i_{l+1}\}$. Hence we may define $\mathcal{F}^r_\infty$-measurable random variables $t^p_l := \varrho^{a_{i_l+1}}_{e^{\tilde r}_{\tilde L_{u_p}}}(a_p)$ and $v^p_l := e^{\tilde r}_{\tilde L_{u_p}}(a_p)$. In fact, on E_k, $v^p_l > 0$ and $t^p_l \geq 0$ $\mathbb{P}$-a.s. Hence the right-hand side of (3.22) is well-defined on E_k and $\mathcal{F}^r_\infty$-measurable.

Assume first that $k = 1$, i.e. $i_1 = n$, $E_1 = \{\tilde L_{u_1} = \tilde L_{u_n}\}$ and $a_j = u_j - \tilde L^{-1}_{\tilde L_{u_1}-}$ for $j \in \{1, \dots, n\}$. Pick $b > 0$ and let $E^b_1 := E_1 \cap \{a_1 > b\}$. By (iii) and (iv) of Lemma 3.13, the map $Q_b : \Upsilon^{(b)}_1 \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1}) \to \mathcal{E}^{(b)}_1 \times \mathcal{C}(\mathbb{R}, \mathbb{S}^{d-1})$ is measurable. Hence, on E^b_1, we may define a random element $Q_b(a_1, \Phi^{-1}_b(e^{\mathcal{X}}_{L_{u_j + L^{-1}_{L_\tau}}})) = \Phi^{-1}_{a_1}(e^{\mathcal{X}}_{L_{u_j + L^{-1}_{L_\tau}}})$. Recall that $N_{a_1}(\tilde L^{-1}_{u_1})$ is the number of excursions or $\tilde r$ that started prior to $\tilde L^{-1}_{u_1}$ with length of at least a_1. Clearly, $N_{a_1}(\tilde L^{-1}_{u_1})$ is $\mathcal{F}^r_\infty$-measurable. Hence, conditional on $\mathcal{F}_{L^{-1}_{L_\tau}} \vee \mathcal{F}^r_\infty$, the law of $\theta^{a_1, N_{a_1}(\tilde L^{-1}_{u_1})}$ equals $\mathbb{P}_\Psi[\,\cdot\,]$ by Lemma 3.20, where $\Phi^{-1}_{a_1}(e^{\mathcal{X}}_{L_{u_j + L^{-1}_{L_\tau}}}) = (e^{\tilde r}_{\tilde L_{u_j}}, \theta^{a_1, N_{a_1}(\tilde L^{-1}_{u_1})})$. On E^b_1, the left-hand side of (3.22) is

$$\mathbb{P}\left[\theta^{a_1, N_{a_1}(\tilde L^{-1}_{u_1})} \in \mathfrak{A}^{t^j_0}_{v^j_0}(B_j) \text{ for } j \in \{1, \dots, n\} \,\middle|\, \mathcal{F}_{L^{-1}_{L_\tau}} \vee \mathcal{F}^r_\infty\right] \;=\; F_{0,n}(t^p_0, v^p_0; 1 \;\leq\; p \;\leq\; n).$$

Since this identity is independent of b and $E^b_1 \nearrow E_1$ as $b \downarrow 0$, the proposition holds for $k = 1$ and any $i_1 = n \in \mathbb{N}$.

We proceed by induction: assume that (3.22) holds for some $k \in \mathbb{N}$ and any increasing sequence of indices of length at most k. Pick an event E_{k+1}. Put differently, choose a sequence of indices $0 = i_0 < i_1 < \cdots < i_k < i_{k+1} = n$. The $(\mathcal{F}_t)$-stopping time $\rho := L^{-1}_{L_\tau} + u_{i_k}$ satisfies $L^{-1}_{L_\tau} < \rho \leq L^{-1}_{L_\rho}$. Since $L^{-1}_{L_\rho}$ is an $(\mathcal{F}_t)$-stopping time, the σ-algebra $\mathcal{F}_{L^{-1}_{L_\rho}}$ is well-defined and contains $\mathcal{F}_{L^{-1}_{L_\tau}}$. For the sequence $0 < i_1 < \cdots < i_k$, define the event E_k as in the statement of the proposition. Note that $E_{k+1} = E_k \cap E'_{k+1}$, where $E'_{k+1} := \{\tilde L_{u_{i_k}} < \tilde L_{u_{i_{k+1}}} = \tilde L_{u_{i_{k+1}}}\}$, and $E_{k+1}, E_k, E'_{k+1} \in \mathcal{F}^r_\infty$. Define a $\mathrm{BES}^V(0)$ process $r' = (r'_u, u \geq 0)$ by $r'_u := r_{L^{-1}_{L_\rho} + u}$. Then its Markov (resp. inverse) local time $L' = (L'_u, u \geq 0)$ (resp. $L'^{-1} = (L'^{-1}_\mu, \mu \geq 0)$) equals $L'_u = L_{L^{-1}_{L_\rho} + u} - L_\rho$ (resp. $L'^{-1}_\mu = L^{-1}_{L_\rho + \mu} - L^{-1}_{L_\rho}$) and L'^{-1} is a subordinator independent of $\mathcal{F}_{L^{-1}_{L_\rho}}$.

Pick $j \in \{i_k + 1, \dots, i_{k+1}\}$. On E'_{k+1} the inequality $u_j + L^{-1}_{L_\tau} > L^{-1}_{L_\rho}$ holds. Hence we can define positive times $u'_j := u_j + L^{-1}_{L_\tau} - L^{-1}_{L_\rho}$ that clearly satisfy $r'_{u'_j} = \tilde r_{u_j}$. Furthermore, we have

$$L'_{u'_j} = L_{u_j + L^{-1}_{L_\tau}} - L_\rho \quad \text{ and } \quad L'^{-1}_{L'_{u'_j}-} = L^{-1}_{(L_{u_j + L^{-1}_{L_\tau}})-} - L^{-1}_{L_\rho}.$$

Hence we find $a_j = u_j + L^{-1}_{L_\tau} - L^{-1}_{(L_{u_j + L^{-1}_{L_\tau}})-} = u'_j - L'^{-1}_{L'_{u'_j}-}$ for all $j \in \{i_k + 1, \dots, i_{k+1}\}$. Let $e^{r'} = (e^{r'}_\mu, \mu \geq 0)$ be the PPP given by $e^{r'}_\mu(u) := r'_{L'^{-1}_{\mu-} + u} \mathbf{1}\{u \leq \tau^{r'}_\mu\}$, $u \in \mathbb{R}_+$, where $\tau^{r'}_\mu := L'^{-1}_\mu - L'^{-1}_{\mu-}$ is the size of the jump of the subordinator L'^{-1} at the moment of local time μ. It holds that $e^{\tilde r}_{\tilde L_{u_j}} = e^r_{L_{u_j + L^{-1}_{L_\tau}}} = e^r_{L_{u'_j + L^{-1}_{L_\rho}}} = e^{r'}_{L'_{u'_j}}$, and hence $t^j_k = \varrho^{a_{i_k+1}}_{e^{r'}_{L'_{u'_j}}}(a_j)$, $v^j_k = e^{r'}_{L'_{u'_j}}(a_j)$, for all $j \in \{i_k + 1, \dots, i_{k+1}\}$. Trivially it holds that $e^{\mathcal{X}}_{L_{u_j + L^{-1}_{L_\tau}}} = e^{\mathcal{X}}_{L_{u'_j + L^{-1}_{L_\rho}}}$, so me may apply the basis of the induction (i.e. $k = 1$) to the stopping time ρ on the event

E'_{k+1} as follows:

$$\mathbb{P}\left[e^{\mathcal{X}}_{L_{u'_j+L^{-1}_{L_\rho}}}(a_j) \in B_j, \quad j \in \{i_k+1,\dots,i_{k+1}\} \,\middle|\, \mathcal{F}_{L^{-1}_{L_\rho}} \vee \mathcal{F}^r_\infty\right] = F_{i_k,i_{k+1}}\left(\varrho^{a_{i_k+1}}_{e^{r'}_{L'_{u'_j}}}(a_j), e^{r'}_{L'_{u'_j}}(a_j); i_k+1 \le j \le i_{k+1}\right).$$

Hence $\mathbb{P}[e^{\mathcal{X}}_{L_{u_j+L^{-1}_{L_\tau}}}(a_j) \in B_j,\ j \in \{i_k+1,\dots,i_{k+1}\} | \mathcal{F}_{L^{-1}_{L_\rho}} \vee \mathcal{F}^r_\infty] = F_{i_k,i_{k+1}}(t^j_k, v^j_k; i_k+1 \le j \le i_{k+1})$ on E'_{k+1}. Define the event $D_k := \cap_{j=1}^{i_k}\{e^{\mathcal{X}}_{L_{u_j+L^{-1}_{L_\tau}}}(a_j) \in B_j\} \cap E_k \in \mathcal{F}_{L^{-1}_{L_\rho}}$. On the event E_{k+1},

$$\mathbb{E}\left[\mathbf{1}\{D_k\}\mathbb{P}\left[e^{\mathcal{X}}_{L_{u_j+L^{-1}_{L_\tau}}}(a_j) \in B_j, \quad j \in \{i_k+1,\dots,i_{k+1}\} \,\middle|\, \mathcal{F}_{L^{-1}_{L_\rho}} \vee \mathcal{F}^r_\infty\right] \,\middle|\, \mathcal{F}_{L^{-1}_{L_\tau}} \vee \mathcal{F}^r_\infty\right] = \mathbb{P}\left[D_k \,\middle|\, \mathcal{F}_{L^{-1}_{L_\tau}} \vee \mathcal{F}^r_\infty\right] F_{i_k,i_{k+1}}(t^j_k, v^j_k; i_k+1 \le j \le i_{k+1}),$$

which equals the left-hand side in (3.22). The proposition follows by the induction hypothesis. □

Corollary 3.23. *Let $\mathcal{X}$ be a solution of SDE* (3.1) *started at* $\mathbf{0}$ *and adapted to* $(\mathcal{F}_t, t \ge 0)$.

(a) Let τ be a finite $(\mathcal{F}_t)$-stopping time. Then the process $\tilde{\mathcal{X}} = (\tilde{\mathcal{X}}_t, t \ge 0)$, defined by $\tilde{\mathcal{X}}_t := \mathcal{X}_{L^{-1}_{L_\tau}+t}$, is independent of $\mathcal{F}_{L^{-1}_{L_\tau}}$ and has the same law as $\mathcal{X}$.

(b) Let $\mathcal{Y}$ be a solution of SDE (3.1) *started at* $\mathbf{0}$. *Then the laws on $\mathcal{C}_d$ of $\mathcal{X}$ and $\mathcal{Y}$ coincide.*

Proof. (a) If we prove that for any $0 < u_1 < u_2 < \cdots < u_n$ and measurable sets $B_1,\dots,B_n \subseteq \mathbb{R}^d$, the equality $\mathbb{P}[\tilde{\mathcal{X}}_{u_1} \in B_1,\dots,\tilde{\mathcal{X}}_{u_n} \in B_n | \mathcal{F}_{L^{-1}_{L_\tau}}] = \mathbb{P}[\mathcal{X}_{u_1} \in B_1,\dots,\mathcal{X}_{u_n} \in B_n]$ holds $\mathbb{P}$-a.s., part (a) follows by a diagonalization argument (cf. first paragraph in the proof of Lemma 3.20), since $\tilde{\mathcal{X}}_0 = \mathcal{X}_0$ and all the trajectories of $\tilde{\mathcal{X}}$ are continuous. Recall that $L_{L^{-1}_{L_\tau}+u} = L_\tau + \tilde{L}_u$. Hence, for all $u \ge 0$, $\tilde{\mathcal{X}}_u = e^{\mathcal{X}}_{L_\tau+\tilde{L}_u}(u - \tilde{L}^{-1}_{\tilde{L}_u-})$ and in particular (take $\tau \equiv 0$) $\mathcal{X}_u = e^{\mathcal{X}}_{L_u}(u - L^{-1}_{L_u-})$. Note that the set E_k in Prop. 3.21 is determined by $k \in \{1,\dots,n\}$ and the indices $i_1 < \ldots < i_{k-1}$ (with $i_0 = 0$ and $i_k = n$) and should be denoted by $E^{i_1,\dots,i_{k-1}}_k$. Furthermore, $E^{i_1,\dots,i_{k-1}}_k \cap E^{i'_1,\dots,i'_{k'-1}}_{k'} \ne \emptyset$ if and only if $k = k', i_1 = i'_1, \dots, i_{k-1} = i'_{k'-1}$, in which case the two sets clearly coincide. Put differently, this finite family of sets is pairwise disjoint. Since the union of $E^{i_1,\dots,i_{k-1}}_k$ equals the entire probability space, we can define a path functional

$$F(\tilde{\mathcal{X}}) := \sum_{k,i_1<\cdots<i_{k-1}} \mathbf{1}\{E^{i_1,\dots,i_{k-1}}_k\} \prod_{l=0}^{k-1} F_{i_l,i_{l+1}}\left(\varrho^{a_{i_l+1}}_{e^{\tilde{r}}_{\tilde{L}_{u_p}}}(a_p), e^{\tilde{r}}_{\tilde{L}_{u_p}}(a_p); i_l+1 \le p \le i_{l+1}\right).$$

Note that F is defined $\mathbb{P}$-a.s. on Ω and is measurable. Furthermore, F is a function only of the radial component $\tilde{r} = \|\tilde{\mathcal{X}}\|$ of $\tilde{\mathcal{X}}$. By Prop. 3.21, we get $\mathbb{P}[\tilde{\mathcal{X}}_{u_1} \in B_1,\dots,\tilde{\mathcal{X}}_{u_n} \in B_n | \mathcal{F}_{L^{-1}_{L_\tau}} \vee \mathcal{F}^r_\infty] = F(\tilde{\mathcal{X}})$. An identical argument applied to $\mathcal{X}$ (with $\tau \equiv 0$) yields $\mathbb{P}[\mathcal{X}_{u_1} \in B_1,\dots,\mathcal{X}_{u_n} \in B_n | \mathcal{F}_0 \vee \mathcal{F}^r_\infty] = F(\mathcal{X})$. By the strong Markov property of r, the process $\tilde{r}$, and therefore $F(\tilde{\mathcal{X}})$, is independent of $\mathcal{F}_{L^{-1}_{L_\tau}}$. Hence $\mathbb{P}[\tilde{\mathcal{X}}_{u_1} \in B_1,\dots,\tilde{\mathcal{X}}_{u_n} \in B_n | \mathcal{F}_{L^{-1}_{L_\tau}}] = \mathbb{E}[F(\tilde{\mathcal{X}})]$ a.s. Since the laws of r and $\tilde{r}$ coincide, we have $\mathbb{E}[F(\tilde{\mathcal{X}})] = \mathbb{E}[F(\mathcal{X})] = \mathbb{P}[\mathcal{X}_{u_1} \in B_1,\dots,\mathcal{X}_{u_n} \in B_n]$. This concludes the proof of (a).

(b) As before it is sufficient to show $\mathbb{P}[\mathcal{X}_{u_1} \in B_1, \dots, \mathcal{X}_{u_n} \in B_n] = \mathbb{P}'[\mathcal{Y}_{u_1} \in B_1, \dots, \mathcal{Y}_{u_n} \in B_n]$ for any $0 < u_1 < u_2 < \cdots < u_n$ and measurable sets $B_1, \dots, B_n \subseteq \mathbb{R}^d$, where $\mathbb{P}'[\cdot]$ is the probability measure on the space where $\mathcal{Y}$ is defined. Prop. 3.21 implies this statement, using the same argument as in part (a) as the processes $\|\mathcal{X}\|$ and $\|\mathcal{Y}\|$ have the same law. □

Corollary 3.24. *Let $\mathcal{X}$ be a solution of SDE* (3.1) *started at* $\mathbf{0}$*. The point process $e^{\mathcal{X}}$ on $\mathcal{E}_d^+ \cup \{\delta_d\}$, defined in* (3.20)*, is a PPP with excursion measure characterized in Prop. 3.16.*

Proof. Let $\mathcal{X}$ be adapted to $(\mathcal{F}_t, t \geq 0)$. Pick $\lambda \in \mathbb{R}_+$ and recall that L_λ^{-1} is an $(\mathcal{F}_t)$-stopping time. Define $\tilde{\mathcal{X}} = (\tilde{\mathcal{X}}_t, t \geq 0)$ by $\tilde{\mathcal{X}}_t := \mathcal{X}_{L_\lambda^{-1}+t}$.

Claim 1. The process $\tilde{\mathcal{X}}$ is independent of $\mathcal{F}_{L_\lambda^{-1}}$ and its law is equal to that of $\mathcal{X}$.

Proof of Claim 1. Define an $(\mathcal{F}_t)$-stopping time $\tau := \inf\{t \geq 0 : L_t \geq \lambda\}$. Since the local time L is continuous and $\lim_{t\uparrow\infty} L_t = \infty$ a.s., it holds that $\mathbb{P}[L_\tau = \lambda] = \mathbb{P}[\tau < \infty] = 1$. In particular, $L_\lambda^{-1} = L_{L_\tau}^{-1}$ and, by Corollary 3.23(a), the claim follows.

Define the filtration $(\mathcal{G}_\lambda, \lambda \geq 0)$ by $\mathcal{G}_\lambda := \mathcal{F}_{L_\lambda^{-1}}$. Pick $a > 0$ and a measurable set $\mathfrak{A} \in \mathcal{E}_d^{(a)}$.

Claim 2. The counting process $N^{\mathfrak{A}} = (N_\lambda^{\mathfrak{A}}, \lambda \geq 0)$, where $N_\lambda^{\mathfrak{A}}$ equals the cardinality of the set $\{s \in (0, \lambda] : e_s^{\mathcal{X}} \in \mathfrak{A}\}$, is a $(\mathcal{G}_\lambda)$-Poisson process with intensity $\mu_r \otimes \mathbb{P}_\Psi[\Phi_a^{-1}(\mathfrak{A})]$.

Before proving the claim, note that it implies that $e^{\mathcal{X}}$ is a PPP with excursion measure ν from Prop. 3.16. Indeed, for disjoint sets $\mathfrak{A}_1, \dots, \mathfrak{A}_n$ in $\mathcal{E}_d^{(a)}$, the respective counting processes $N^{\mathfrak{A}_1}, \dots, N^{\mathfrak{A}_n}$ are, by Claim 2, $(\mathcal{G}_\lambda)$-Poisson processes that cannot jump simultaneously. Hence they must be independent. For any collection of disjoint sets $\mathfrak{A}_1 \times (s_1, t_1], \dots, \mathfrak{A}_n \times (s_n, t_n]$ in $\mathcal{E}_d^+ \times \mathbb{R}_+$ satisfying $0 < \nu(\mathfrak{A}_j) < \infty$ for all $j \in \{1, \dots, n\}$, by Prop. 3.16 there exists $a > 0$ such that all the sets are contained in $\mathcal{E}_d^{(a)} \times \mathbb{R}_+$. Furthermore, the numbers of points of $e^{\mathcal{X}}$ in each of the sets is given by n independent Poisson rvs $N_{t_j}^{\mathfrak{A}_j} - N_{s_j}^{\mathfrak{A}_j}$ with intensities $(t_j - s_j)\nu(\mathfrak{A}_j)$.

Proof of Claim 2. It is clear from the definition of $N^{\mathfrak{A}}$ that it is adapted to $(\mathcal{G}_\lambda, \lambda \geq 0)$. Pick $\lambda, \mu \in \mathbb{R}_+$. It is sufficient to prove that $N_{\mu+\lambda}^{\mathfrak{A}} - N_\lambda^{\mathfrak{A}}$ is independent of $\mathcal{G}_\lambda$ and has the same law as $N_\mu^{\mathfrak{A}}$. The number of excursions of $\mathcal{X}$ in $\mathfrak{A}$ completed during the time interval $(L_\lambda^{-1}, L_{\lambda+\mu}^{-1}]$ is by construction equal to the number $\tilde{N}_\mu^{\mathfrak{A}}$ of excursions in $\mathfrak{A}$ of $\tilde{\mathcal{X}}$ from Claim 1, completed in the time interval $(0, \tilde{L}_\mu^{-1}]$. Recall that $\tilde{L}_\mu^{-1} = L_{\lambda+\mu}^{-1} - L_\lambda^{-1}$ is the inverse local time at the origin of $\tilde{r} = \|\tilde{\mathcal{X}}\|$, and hence of $\tilde{\mathcal{X}}$. Since, by Claim 1, $\tilde{\mathcal{X}}$ is independent of $\mathcal{G}_\lambda$, so is $\tilde{N}_\mu^{\mathfrak{A}} = N_{\mu+\lambda}^{\mathfrak{A}} - N_\lambda^{\mathfrak{A}}$. Since, by Claim 1, the laws of $\mathcal{X}$ and $\tilde{\mathcal{X}}$ coincide, so do the laws of $N_\mu^{\mathfrak{A}}$ and $\tilde{N}_\mu^{\mathfrak{A}}$. This concludes the proof of Claim 2. □

Proof of Theorem 1.1 in the recurrent case. As mentioned in §3.1, weak existence for SDE (3.1) follows from [18, §2.6, Thm 1]. When $\mathbf{x}_0 = \mathbf{0}$, Corollary 3.23 shows uniqueness in law of solutions to (3.1), and, as mentioned in §3.1, the strong Markov property then follows by [26, Thm 6.2.2]. The case $\mathbf{x}_0 \neq \mathbf{0}$ is essentially the same, but one must deal separately with the initial partial excursion; since the case $\mathbf{x}_0 = \mathbf{0}$ is the one we need for Theorem 1.2, we omit the details of the (minor) adjustments required for other case. □

4 Invariance principle

4.1 Invariance principle with discontinuous coefficients

Recall that $\mathcal{D}_d = \mathcal{D}(\mathbb{R}_+; \mathbb{R}^d)$ is a space of functions $x : \mathbb{R}_+ \to \mathbb{R}^d$ that are right-continuous and have left limits (i.e. $x(t) = \lim_{s\downarrow t} x(s)$ for any $t \in \mathbb{R}_+$, $x(t-) := \lim_{s\uparrow t} x(s)$ exists in $\mathbb{R}^d$ for any $t > 0$ and, by convention, $x(0-) := x(0)$). We endow $\mathcal{D}_d$ with the Skorokhod metric (see e.g. [8, §3.5]). By [8, Prop 3.5.3, p. 119], the induced topology on the continuous functions $\mathcal{C}_d = \mathcal{C}(\mathbb{R}_+;\mathbb{R}^d)$ coincides with the compact-open topology. Theorem 4.1 may be viewed as an extension of [8, Thm 7.4.1, p. 354] to a setting with discontinuous coefficients. It is key in establishing Theorem 1.2.

Theorem 4.1. *Let $a = (a_{ij}) : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ be a bounded function that is continuous on $\mathbb{R}^d \setminus \{\mathbf{0}\}$, with image contained in the set of symmetric, non-negative definite matrices in $\mathbb{R}^d \otimes \mathbb{R}^d$. Suppose that the $\mathcal{C}_d$ martingale problem for (G, v) is well-posed, where $Gf := \frac{1}{2}\sum a_{ij}\partial_i\partial_j f$ (for a smooth $f : \mathbb{R}^d \to \mathbb{R}$ with compact support) and v is a distribution $\mathbb{R}^d$. For $n \in \mathbb{N}$, let Z_n be a process with sample paths in $\mathcal{D}_d$ and let $A_n = (A_n^{ij})$ be a symmetric $\mathbb{R}^d \otimes \mathbb{R}^d$-valued process started at zero, such that A_n^{ij} has sample paths in $\mathcal{D}_1$ and $A_n(t) - A_n(s)$ is non-negative definite for all $t > s \geq 0$. Set $\mathcal{F}_t^n := \sigma(Z_n(s), A_n(s), s \leq t)$. Suppose that Z_n^i and $Z_n^i Z_n^j - A_n^{ij}$ are $\mathcal{F}_t^n$-adapted local martingales for each $i,j \in \{1,\ldots,d\}$. Let $\tau_n^r := \inf\{t \geq 0 : \|Z_n(t)\| \geq r$ or $\|Z_n(t-)\| \geq r\}$ (with convention $\inf\emptyset := \infty$) and suppose that for every $r > 0$, $T > 0$, and $i,j \in \{1,\ldots,d\}$,*

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_n^r} \|Z_n(t) - Z_n(t-)\|^2\right] = 0; \tag{4.1}$$

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_n^r} \left|A_n^{ij}(t) - A_n^{ij}(t-)\right|\right] = 0; \tag{4.2}$$

and, as $n \to \infty$,

$$\sup_{0\leq t\leq T\wedge\tau_n^r} \left|A_n^{ij}(t) - \int_0^t a_{ij}(Z_n(s))\mathrm{d}s\right| \xrightarrow{\mathrm{P}} 0, \tag{4.3}$$

where $\xrightarrow{\mathrm{P}}$ denotes convergence in probability and $s \wedge t := \min\{s,t\}$ for $s,t \in [0,\infty]$. Assume $\sup_{n\in\mathbb{N}} \mathbb{E}\|Z_n(0)\|^2 < \infty$. Suppose that $Z_n(0)$ and $\|Z_n\|$ converge weakly to a probability law v on $\mathbb{R}^d$ and the law of a Bessel process of dimension greater than one, respectively. Then Z_n converges weakly to the solution of the martingale problem for (G,v).

The underlying idea for the proof of Theorem 4.1 is standard: show that every subsequence of $(Z_n)_{n\in\mathbb{N}}$ has a further subsequence converging weakly to the law given by the solution of the martingale problem (G,v) (cf. proof of [8, Thm 7.4.1, p. 354]). Since a in Theorem 4.1 is bounded, $a_i := \sup_{x\in\mathbb{R}^d} a_{ii}(x)$ is finite for each $i \in \{1,\ldots,d\}$. Since $A_n^{ii}(t) \geq A_n^{ii}(t-)$ for all $t \geq 0$ and $i \in \{1,\ldots,d\}$,

$$\eta_n := \inf\left\{t \geq 0 : \max_{1\leq i\leq d}\{A_n^{ii}(t) - a_i t\} \geq 1\right\}$$

is an $(\mathcal{F}_t^n)$-stopping time. Since $\eta_n \geq \inf\{t \geq 0 : \max_{1\leq i\leq d}|A_n^{ii}(t) - \int_0^t a_{ii}(Z_n(s))\mathrm{d}s| \geq 1\}$ and (4.3) holds for any $T, r > 0$, we have that

$$\eta_n \xrightarrow{\mathrm{P}} \infty \qquad \text{as } n \to \infty. \tag{4.4}$$

Define for given $r > 0$, $n \in \mathbb{N}$ and $i,j \in \{1,\ldots,d\}$ the processes $\tilde{Z}_n^r$ and $\tilde{A}_n^{ij}$ by

$$\tilde{Z}_n^r(t) := Z_n(t \wedge \eta_n \wedge \tau_n^r) \qquad \text{and} \qquad \tilde{A}_n^{ij}(t) := A_n^{ij}(t \wedge \eta_n \wedge \tau_n^r), \tag{4.5}$$

respectively ($\tilde{A}_n^{ij}$ depends on r but this is suppressed from the notation as it is clear from the context). Observe that for any $T > 0$ and $(\mathcal{F}_t^n)$-stopping time τ less than T, the modulus of any component of $\tilde{Z}_n^r(\tau) - \tilde{Z}_n^r(0)$ is bounded above by an integrable random variable:

$$\|\tilde{Z}_n^r(\tau) - \tilde{Z}_n^r(0)\| \leq 2r + \sup_{0\leq t\leq T\wedge\tau_n^r} \|Z_n(t) - Z_n(t-)\|. \tag{4.6}$$

Since $\tilde{Z}_n^r(0) = Z_n(0)$ is integrable by assumption, the local martingale $\tilde{Z}_n^r$ is of class (DL) and therefore a martingale [24, Prop. IV.1.7]. An analogous argument, relying on (4.1)–(4.2), the inequality $|\tilde{Z}_n^{r,i}\tilde{Z}_n^{r,j}| \le (\tilde{Z}_n^{r,i})^2 + (\tilde{Z}_n^{r,j})^2$ and the square integrability of $\|Z_n(0)\|$, shows that $\tilde{Z}_n^{r,i}\tilde{Z}_n^{r,j} - \tilde{A}_n^{ij}$ is also a martingale. Furthermore, since $A_n^{ii}(0) = 0$ for all indices $i \in \{1, \dots, d\}$, for any $t \ge 0$ we have

$$\tilde{A}_n^{ii}(t) \le a_i t + 1 + \sup_{0 \le s \le t \wedge \tau_n^r} \left(A_n^{ii}(s) - A_n^{ii}(s-) \right). \tag{4.7}$$

Lemma 4.2. *For each $r > 0$, the sequence of the laws of processes $(\tilde{Z}_n^r)_{n\in\mathbb{N}}$ on $\mathcal{D}_d$ is relatively compact in the metric space of all probability measures on $\mathcal{D}_d$ with the Prokhorov metric.*[3]

Proof. We prove the lemma by establishing the sufficient condition for the relative compactness of the sequence $(\tilde{Z}_n^r)_{n\in\mathbb{N}}$ given in [8, Thm 3.8.6, pp. 137–138]. Fix an arbitrary $T > 0$ and let B_K denote a closed ball of radius $K > 2r + 1$ in $\mathbb{R}^d$. Note that the bound in (4.6) and the Markov inequality imply

$$\begin{aligned} \mathbb{P}\left[\tilde{Z}_n^r(t) \in B_K \text{ for all } t \in [0, T]\right] &\ge \mathbb{P}\left[2r + \|Z_n(0)\| + \sup_{0 \le t \le T \wedge \tau_n^r} \|Z_n(t) - Z_n(t-)\| \le K\right] \\ &\ge 1 - \frac{C_0}{(K - 2r)^2} \qquad \text{for all } n \in \mathbb{N}, \end{aligned}$$

where $C_0 > 0$ depends on the quantities $\sup_{n\in\mathbb{N}} \mathbb{E}\left[\sup_{0\le t\le T\wedge\tau_n^r} \|Z_n(t) - Z_n(t-)\|^2\right]$ and $\sup_{n\in\mathbb{N}} \mathbb{E}\,\|Z_n(0)\|^2$, which are finite by assumption. As K is independent of n and can be arbitrarily large, the compact containment condition [8, Eq (7.9), p. 129] holds for $(\tilde{Z}_n^r)_{n\in\mathbb{N}}$. Hence condition (a) of [8, Thm 3.7.2], also assumed in [8, Thm 3.8.6, pp. 137–138], holds.

Since $\tilde{Z}_n^{r,i}$ and $(\tilde{Z}_n^{r,i})^2 - \tilde{A}_n^{ii}$ are martingales for all $i \in \{1, \dots, d\}$, it holds that

$$\mathbb{E}\left[\left\|\tilde{Z}_n^r(t+h) - \tilde{Z}_n^r(t)\right\|^2 \middle| \mathcal{F}_t^n\right] = \mathbb{E}\left[\sum_{i=1}^d \left(\tilde{A}_n^{ii}(t+h) - \tilde{A}_n^{ii}(t)\right) \middle| \mathcal{F}_t^n\right]$$

for any $t, h \ge 0$. With this in mind, define

$$\gamma_n(\delta) := \sup_{0\le t\le T\wedge\tau_n^r} \sum_{i=1}^d \left(\tilde{A}_n^{ii}(t+\delta) - \tilde{A}_n^{ii}(t)\right)$$

for any $\delta > 0$. In order to compare $\gamma_n(\delta)$ with the corresponding quantity for the limiting process, let

$$\Gamma_n(\delta) := \gamma_n(\delta) - \sup_{t\in[0,T\wedge\tau_n^r]} \sum_{i=1}^d \int_t^{t+\delta} a_{ii}(\tilde{Z}_n^r(s))\mathrm{d}s.$$

Now we have from (4.3) that

$$\sup_{0\le t\le T\wedge\tau_n^r} \left|\tilde{A}_n^{ii}(t+\delta) - \int_0^{t+\delta} a_{ii}(\tilde{Z}_n^r(s))\mathrm{d}s\right| \text{ and } \sup_{0\le t\le T\wedge\tau_n^r} \left|\tilde{A}_n^{ii}(t) - \int_0^t a_{ii}(\tilde{Z}_n^r(s))\mathrm{d}s\right|$$

[3] See [8, §3.1, p. 96] for the definition and properties of the Prokhorov metric on the set of probability measures defined on a Borel σ-algebra on a metric space. In this context we use the Skorokhod metric d on $\mathcal{D}_d$, cf. [8, §3.5, p. 116]. The induced topology is the one of weak convergence of probability measures [8, Thm 3.3.1, p. 108].

both tend to zero in probability, implying that $\Gamma_n(\delta)$ also tends to zero in probability:

$$|\Gamma_n(\delta)| \le \sup_{t\in[0,T\wedge\tau_n^r]} \sum_{i=1}^d \left| \tilde{A}_n^{ii}(t+\delta) - \tilde{A}_n^{ii}(t) - \int_t^{t+\delta} a_{ii}(\tilde{Z}_n^r(s))\mathrm{d}s \right| \xrightarrow{\mathrm{P}} 0. \tag{4.8}$$

Since the upper bound in (4.7) is non-decreasing in t, we get

$$|\Gamma_n(\delta)| \le \sum_{i=1}^d \left(3a_i(T+\delta) + 2 + 2 \sup_{s\in[0,(T+\delta)\wedge\tau_n^r]} \left(A_n^{ii}(s) - A_n^{ii}(s-) \right) \right).$$

By (4.2) the right-hand side of this inequality converges in L^1 as $n\to\infty$. Thus the sequence $(\Gamma_n(\delta))_{n\in\mathbb{N}}$ must be uniformly integrable and hence by (4.8) converges to zero in L^1. By adding and subtracting the relevant term we find

$$\limsup_{n\to\infty} \mathbb{E}\,\gamma_n(\delta) \le \limsup_{n\to\infty} \mathbb{E}\,|\Gamma_n(\delta)| + \limsup_{n\to\infty} \mathbb{E} \sup_{t\in[0,T\wedge\tau_n^r]} \sum_{i=1}^d \int_t^{t+\delta} a_{ii}(\tilde{Z}_n^r(s))\mathrm{d}s \le \delta \sum_{i=1}^d a_i.$$

Hence it clearly holds that $\lim_{\delta\to0}\limsup_{n\to\infty}\mathbb{E}\,\gamma_n(\delta)=0$ and the relative compactness of $\tilde{Z}_n^r$ now follows from [8, Thm 3.8.6, p. 137–138] (see also [8, Rem 8.7(b), p. 138]). $\square$

For any path $x\in\mathcal{D}_d$, we define the time $\tau^r(x)$ of its first contact with the complement of the open ball of radius r in $\mathbb{R}^d$ (centred at the origin) by

$$\tau^r(x) := \inf\{t\ge0 \,:\, \|x(t)\|\ge r \quad\text{or}\quad \|x(t-)\|\ge r\}, \tag{4.9}$$

where $\inf\emptyset=\infty$. If it is clear from the context which path x we are considering, to simplify the notation we sometimes write τ^r for $\tau^r(x)$. Note that if x is continuous, then $\tau^r(x)=\inf\{t\ge0 : \|x(t)\|\ge r\}$. The following lemma is important in the proof of Theorem 4.1.

Lemma 4.3. *Let $\mathbb{P}$ be a probability measure on $\mathcal{D}_d$. Then the complement in $\mathbb{R}_+$ of the set $\{r\in\mathbb{R}_+ : \mathbb{P}[\lim_{s\to r}\tau^s=\tau^r]=1\}$ is at most countable, with τ^r defined in* (4.9)*.*

To prove Lemma 4.3 we first need to establish properties of the function $r\mapsto\tau^r$.

Lemma 4.4. *Fix $x\in\mathcal{D}_d$. The function $r\mapsto\tau^r(x)$, mapping $\mathbb{R}_+$ into $[0,\infty]$, is non-decreasing, has right limits and is left continuous. Put differently, for any $r\in\mathbb{R}_+$ the limit $\lim_{s\downarrow r}\tau^s =: \tau^{r+}$ exists in $[0,\infty]$ and, for $r>0$, it holds that $\lim_{s\uparrow r}\tau^s=\tau^r$. Furthermore, for any $r\in\mathbb{R}_+$ the following hold:*

(i) if $\tau^r=\infty$ then $\lim_{s\to r}\tau^s=\tau^r$;

(ii) if $\tau^r<\infty$ then for any $\varepsilon>0$ there are at most finitely many $s\in[0,r]$ such that $\tau^{s+}>\tau^s+\varepsilon$.

Remark 4.5. The topology on $[0,\infty]$ is that of the one-point compactification of $\mathbb{R}_+$. If $\tau^r(x)=\infty$, then the function $s\mapsto\tau^s(x)$ defined on $[0,r]$ may have an infinite number of jumps greater than any given positive constant. If $\tau^r(x)<\infty$, then the inequality $\tau^{r+}(x)>\tau^r(x)$ may hold invalidating the limit in Lemma 4.4(i).

Proof of Lemma 4.4. Observe that $\tau^r(x)=\inf\{t\ge0 : \sup_{0\le s\le t}\|x(s)\|\ge r\}$ is the generalized inverse [7] of the non-decreasing right-continuous function $t\mapsto\sup_{0\le s\le t}\|x(s)\|$. Thus [7, Prop 2.3] $r\mapsto\tau^r(x)$ is non-decreasing, has right limits and is left-continuous.

It follows from the left continuity and monotonicity that $\tau^r=\infty$ implies the limit in (i). Assume $\tau^r<\infty$ and pick $\varepsilon>0$. The intervals in the family $\{[\tau^s,\tau^{s+}) : s\in[0,r]\}$ are disjoint and contained in the bounded interval $[0,\tau^r]$. Hence there can only be finitely many $s\in[0,r]$ satisfying the condition in (ii). $\square$

Proof of Lemma 4.3. Let $A^r_{\varepsilon,\delta} := \{s \in [0,r] : \mathbb{P}[\tau^{s+} > \tau^s + \varepsilon] \geq \delta\}$ for arbitrary $\varepsilon, \delta > 0$, $r \in \mathbb{R}_+$.

Claim. $A^r_{\varepsilon,\delta}$ is at most countable.

Note first that the claim implies the lemma. By Lemma 4.4, the following equivalence holds for any $r \in \mathbb{R}_+$: $\lim_{s\to r} \tau^s = \tau^r \iff \tau^{r+} = \tau^r$. Hence it suffices to show the set

$$\{r \in \mathbb{R}_+ : \mathbb{P}[\tau^{r+} > \tau^r] > 0\} = \cup_{n=1}^\infty \cup_{k=1}^\infty \cup_{i=1}^\infty A^{s_n}_{\varepsilon_k,\delta_i}$$

is at most countable, which clearly holds by the claim, where $(\varepsilon_k)_{k\in\mathbb{N}}$, $(\delta_i)_{i\in\mathbb{N}}$ and $(s_n)_{n\in\mathbb{N}}$ are monotone sequences satisfying $\varepsilon_k \downarrow 0$, $\delta_i \downarrow 0$ and $s_n \uparrow \infty$.

Proof of Claim. Assume that $A^r_{\varepsilon,\delta}$ is uncountable and let I be the set of its isolated points (i.e. $x \in I$ if and only if $x \in A^r_{\varepsilon,\delta}$ and there exists a neighbourhood U of x in $\mathbb{R}_+$ such that $\{x\} = U \cap A^r_{\varepsilon,\delta}$). Then I is at most countable. To see this, note that for each $x \in I$ there exists a rational number $q_x \leq x$, such that $[q_x, x) \cap A^r_{\varepsilon,\delta} = \emptyset$ (for $x \in I \cap \mathbb{Q}$ we may take $q_x := x$). For any distinct points $x, y \in I$, it clearly holds $q_x \neq q_y$. Hence the cardinality of I is at most that of $\mathbb{Q}$ and the uncountable set $A^r_{\varepsilon,\delta} \setminus I$ has no isolated points.

Consider $r_1 := \sup\{y \in A^r_{\varepsilon,\delta} \setminus I\} \leq r$. There exists a strictly increasing sequence $(p^1_i)_{i\in\mathbb{N}}$ in $A^r_{\varepsilon,\delta} \setminus I$ with limit $p^1_i \uparrow r_1$. It is also clear that any $x \in \{\tau^{p^1_i+} > \tau^{p^1_i} + \varepsilon\} \subset \mathcal{D}_d$ satisfies $\tau^{p^1_i}(x) < \infty$. Hence the event $B^{r_1} := \{\tau^{p^1_i+} > \tau^{p^1_i} + \varepsilon\}$ i.o. satisfies: $\mathbb{P}[B^{r_1}] \geq \delta$ and, for each path $x \in B^{r_1}$, the function $s \mapsto \tau^s(x)$ has infinitely many jumps of size at least ε on the interval $[0, r_1]$. Furthermore, since these jumps occur along a subsequence of $(p^1_i)_{i\in\mathbb{N}}$, Lemma 4.4 implies for any $x \in B^{r_1}$ that $\tau^s(x) < \infty$ for all $s \in [0, r_1)$ and $\tau^{r_1}(x) = \infty$.

Since $(A^r_{\varepsilon,\delta} \setminus I) \subseteq [0, r_1]$, it holds that $(A^r_{\varepsilon,\delta} \setminus I) \subseteq A^{r_1}_{\varepsilon,\delta}$ making $A^{r_1}_{\varepsilon,\delta}$ uncountable. Furthermore, since $A^{r_1}_{\varepsilon,\delta} \setminus \{r_1\} = \cup_{s<r_1} A^s_{\varepsilon,\delta}$, there exists $r' < r_1$ such that $A^{r'}_{\varepsilon,\delta}$ is uncountable. We can now repeat the construction above, with $A^r_{\varepsilon,\delta}$ substituted by $A^{r'}_{\varepsilon,\delta}$, to define the event B^{r_2} (for some $r_2 \in (0, r']$) with properties analogous to those of B^{r_1}. In particular $\mathbb{P}[B^{r_2}] \geq \delta$ and, since each $x \in B^{r_2}$ satisfies $\tau^{r_2}(x) = \infty$, it must hold $B^{r_1} \cap B^{r_2} = \emptyset$. As before, there exists $r'' < r_2$ such that $A^{r''}_{\varepsilon,\delta}$ is uncountable. By the same construction there exists $r_3 \in (0, r'']$ and an event B^{r_3} satisfying $\mathbb{P}[B^{r_3}] \geq \delta$ and $B^{r_3} \cap (B^{r_1} \cup B^{r_2}) = \emptyset$, since $x \in B^{r_3}$ satisfies $\tau^{r_3}(x) = \infty$ while for any $x \in B^{r_1} \cup B^{r_2}$ we have $\tau^{r_3}(x) < \infty$. We can thus inductively construct a sequence of pairwise disjoint events $(B^{r_n})_{n\in\mathbb{N}}$ in $\mathcal{D}_d$ each of which has probability at least $\delta > 0$. This contradicts the fact that the total mass of $\mathbb{P}$ is equal to one. □

Remark 4.6. The proof of the claim, contained in the proof of Lemma 4.3, shows that $A^r_{\varepsilon,\delta}$ is in fact locally finite.

In order to apply Lemma 4.3 in the proof of Theorem 4.1, we need another fact about the metric space $(\mathcal{D}_d, d)$, where the metric $d : \mathcal{D}_d \times \mathcal{D}_d \to \mathbb{R}_+$ that induces the Skorokhod topology is defined in [8, Eq. (5.2), p. 117] (see also [8, §3.5]).

Lemma 4.7. *Pick $r > 0$. Assume that $x \in \mathcal{D}_d$ satisfies $\lim_{s\to r} \tau^s(x) = \tau^r(x)$ (see (4.9) for definition of $\tau^r(x)$). Then the function $\mathcal{D}_d \to [0,\infty]$, given by $y \mapsto \tau^r(y)$, is continuous at x. If in addition it holds that either $\|x(\tau^r(x)-)\| < r$ or $\|x(\tau^r(x))\| \leq r$, then the map $\mathcal{D}_d \to \mathcal{D}_d$, given by $y \mapsto y(\,\cdot \wedge \tau^r(y))$, is continuous at x.*

Remark 4.8. (a) The lemma implies that if $x \in \mathcal{C}_d$ satisfies $\lim_{s\to r} \tau^s(x) = \tau^r(x)$, the map $\mathcal{D}_d \to \mathcal{D}_d \times [0,\infty]$, given by $y \mapsto (y(\,\cdot \wedge \tau^r(y)), \tau^r(y))$, is continuous at x.

(b) It is easy to construct $x \in \mathcal{C}_d$, such that both $y \mapsto \tau^r(y)$ and $y \mapsto y(\,\cdot \wedge \tau^r(y))$ are discontinuous at x. The key feature of such a function x is that $\tau^{r+}(x) > \tau^r(x)$ (see Lemma 4.4 for the definition of $\tau^{r+}(x)$).

(c) If $x \in \mathcal{D}_d \setminus \mathcal{C}_d$, then the additional assumption in the lemma is necessary for the continuity of $y \mapsto y(\cdot \wedge \tau^r(y))$ to hold at x. To see this, for any $r > 0$ and $\varepsilon \in [0,1)$, consider $x_\varepsilon(t) := (t+\varepsilon)\mathbf{1}(0 \le t < r) + (r+1)\mathbf{1}(r \le t < \infty)$. Then x_0 clearly satisfies the first assumption in the lemma but not the second one. Note that for any $\varepsilon \in (0,1)$ we have $d(x_0, x_\varepsilon) \le \varepsilon$ and $|x_0(t \wedge \tau^r(x_0)) - x_\varepsilon(t \wedge \tau^r(x_\varepsilon))| \ge \mathbf{1}(r \le t < \infty)$.

Proof of Lemma 4.7. Let $x \in \mathcal{D}_d$ satisfy $\lim_{s \to r} \tau^s(x) = \tau^r(x)$. We first prove that for any sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathcal{D}_d$, such that $d(x_n, x) \to 0$, it holds that $\tau^r(x_n) \to \tau^r(x)$. Note that $d(x_n, x) \to 0$ and the definition of d in [8, Eq. (5.2), p. 117] imply that there exists a sequence $(\lambda_n)_{n \in \mathbb{N}}$ of strictly increasing, Lipschitz continuous, surjective functions $\lambda_n : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying

$$\sup\{\|x_n(\lambda_n(t)) - x(t)\|, |\lambda_n(t) - t| : t \in [0,T]\} \to 0 \qquad \text{for any } T > 0. \tag{4.10}$$

If $\tau^r(x) = \infty$, then for any $T > 0$ $\exists \delta > 0$ such that $\sup_{t \in [0,T]}\{\|x(t)\|, \|x(t-)\|\} < r - \delta$. By (4.10), for all sufficiently large $n \in \mathbb{N}$ we have $\sup_{s \in [0, \lambda_n(T)]}\{\|x_n(s)\|\} < r - \delta/2$, implying $\tau^r(x_n) \ge T - 1$. Since T was arbitrary, it holds that $\tau^r(x_n) \to \infty$.

Assume now that $\tau^r(x) < \infty$ and that $(\tau^r(x_n))_{n \in \mathbb{N}}$ does not converge to $\tau^r(x)$. By passing to a subsequence (again denoted by $(x_n)_{n \in \mathbb{N}}$), we may assume that $\exists \varepsilon > 0$ such that $|\tau^r(x_n) - \tau^r(x)| > \varepsilon$ for all $n \in \mathbb{N}$. Pick $T > \tau^r(x) + \varepsilon$ and note that without loss of generality we may assume (for all $n \in \mathbb{N}$) that either $\tau^r(x_n) > \tau^r(x) + \varepsilon$ or $\tau^r(x_n) < \tau^r(x) - \varepsilon$. Consider first the former case. By Lemma 4.4, our assumption is equivalent to $\tau^{r+}(x) = \tau^r(x)$. Hence $\exists \delta > 0$ and an interval $[t_0, s_0]$ contained in $(\tau^r(x), \tau^r(x) + \varepsilon)$, such that $\inf_{t \in [t_0, s_0]} \|x(t)\| > r + \delta$. As $[t_0, s_0] \subset [0, T]$, by (4.10) there exists $n \in \mathbb{N}$ and $t \in (t_0, s_0)$ such that $\lambda_n(t) < s_0$ and $\|x_n(\lambda_n(t))\| \ge \|x(t)\| - \|x(t) - x_n(\lambda_n(t))\| > r + \delta/2$, contradicting $\tau^r(x_n) > \tau^r(x) + \varepsilon > \lambda_n(t)$.

Consider now the case $\tau^r(x_n) < \tau^r(x) - \varepsilon$ for all $n \in \mathbb{N}$. Then for a sequence $\delta_n \downarrow 0$ we have $\sup_{s \in [0, \tau^r(x) - \varepsilon)} \|x_n(s)\| > r - \delta_n$. Hence there exists a sequence $(t_n)_{n \in \mathbb{N}}$ in $(0, \tau^r(x) - \varepsilon)$ such that $\|x_n(t_n)\| \to r$. By (4.10) it holds that $\lambda_n^{-1}(t_n) < \tau^r(x) - \varepsilon/2$ for all sufficiently large (and thus wlog all) $n \in \mathbb{N}$. Furthermore, the triangle inequality and (4.10) imply $|\|x(\lambda_n^{-1}(t_n))\| - r| \le \|x(\lambda_n^{-1}(t_n)) - x_n(t_n)\| + |\|x_n(t_n)\| - r| \to 0$, since $\lambda_n^{-1}(t_n), t_n \in [0,T]$ for all $n \in \mathbb{N}$. By passing to a convergent subsequence, there exists $\alpha \le \tau^r(x) - \varepsilon/2$ such that either $\lambda_n^{-1}(t_n) \uparrow \alpha$ or $\lambda_n^{-1}(t_n) \downarrow \alpha$. Hence we either get $\|x(\alpha-)\| = r$ or $\|x(\alpha)\| = r$, contradicting the fact that $\alpha < \tau^r(x)$. This implies the continuity of the map $y \mapsto \tau^r(y)$ at x.

Consider the map $y \mapsto y(\cdot \wedge \tau^r(y))$ in the case $\tau^r(x) = \infty$. Then $x(\cdot \wedge \tau^r(x)) = x$ and, as we have already established, $\tau^r(x_n) \to \infty$. By the definition of the metric d (see [8, Eq. (5.2), p. 117]), we have $d(x_n(\cdot \wedge \tau^r(x_n)), x(\cdot \wedge \tau^r(x))) \le d(x_n, x) + d(x_n, x_n(\cdot \wedge \tau^r(x_n))) \le d(x_n, x) + e^{-\tau^r(x_n)} \to 0$.

In the case $\tau^r(x) < \infty$, we have already seen that $\tau^r(x_n) \to \tau^r(x)$. By definition [8, Eq. (5.2), p. 117], for any $y \in \mathcal{D}_d$, $t \in \mathbb{R}_+$ and a sequence $(t_n)_{n \in \mathbb{N}}$ converging to t we have

$$d(y(\cdot \wedge t_n), y(\cdot \wedge t)) \le \|y(t) - y(t_n)\| + |t - t_n| \sup_{s \in [0, t+1]} \|y(s)\|$$

for all large $n \in \mathbb{N}$. Recall that y is bounded on compact intervals. Hence if either $t_n \downarrow t$ or $t_n \to t$ and y is continuous at t, then $d(y(\cdot \wedge t_n), y(\cdot \wedge t)) \to 0$.[4] Therefore the estimate

$$\begin{aligned} d(x_n(\cdot \wedge \tau^r(x_n)), x(\cdot \wedge \tau^r(x))) &\le d(x_n(\cdot \wedge \tau^r(x_n)), x(\cdot \wedge \tau^r(x_n))) \\ &\quad + d(x(\cdot \wedge \tau^r(x_n)), x(\cdot \wedge \tau^r(x))) \\ &\le d(x_n, x) + d(x(\cdot \wedge \tau^r(x_n)), x(\cdot \wedge \tau^r(x))) \end{aligned}$$

[4] Note that if $t_n \uparrow t$, $d(y(\cdot \wedge t_n), y(\cdot \wedge t))$ may be bounded from below by a positive constant $\forall n \in \mathbb{N}$.

implies the lemma, except when $\tau^r(x_n) \uparrow \tau^r(x)$ and $x(\tau^r(x)-) \neq x(\tau^r(x))$.

Assuming $\tau^r(x_n) \uparrow \tau^r(x) < \infty$ and $x(\tau^r(x)-) \neq x(\tau^r(x))$, by $\lim_{s\to r} \tau^s(x) = \tau^r(x)$ it holds that $x(\tau^r(x)-) < x(\tau^r(x))$. Furthermore, since by assumption it either holds that $x(\tau^r(x)-) < r$ or $x(\tau^r(x)) \leq r$, we must have $x(\tau^r(x)-) < r$. Hence there exists $\delta > 0$ such that $\sup_{t\in[0,\tau^r(x))} \|x(t)\| < r - \delta$. Therefore by (4.10) $\exists N \in \mathbb{N}$ such that for all $n \geq N$ and $t \in [0, \tau^r(x))$ we have $\|x_n(\lambda_n(t))\| \leq \|x(t)\| + \|x_n(\lambda_n(t)) - x(t)\| < r - \delta/2$. Thus we obtain $\lambda_n(\tau^r(x)) \leq \tau^r(x_n)$ for all $n \geq N$. As λ_n is increasing, for every $t \in [0, \tau^r(x)]$ it holds that $\|x_n(\lambda_n(t) \wedge \tau^r(x_n)) - x(t \wedge \tau^r(x))\| = \|x_n(\lambda_n(t)) - x(t)\|$. Furthermore, since $\tau^r(x_n) \in [\lambda_n(\tau^r(x)), \tau^r(x)]$, for all $t \in (\tau^r(x), \lambda_n^{-1}(\tau^r(x_n))]$ we have

$$\begin{aligned}\|x_n(\lambda_n(t) \wedge \tau^r(x_n)) - x(t \wedge \tau^r(x))\| &= \|x_n(\lambda_n(t)) - x(\tau^r(x))\| \\ &\leq \|x(t) - x(\tau^r(x))\| + \|x_n(\lambda_n(t)) - x(t)\|.\end{aligned}$$

Hence, for any $T > \tau^r(x)$, it holds that

$$\begin{aligned}&\sup_{t\in[0,T]} \|x_n(\lambda_n(t) \wedge \tau^r(x_n)) - x(t \wedge \tau^r(x))\| \\ &\quad= \sup_{t\in[0,\tau^r(x)]} \|x_n(\lambda_n(t)) - x(t)\| + \sup_{t\in(\tau^r(x), T\wedge\lambda_n^{-1}(\tau^r(x_n))]} \|x_n(\lambda_n(t)) - x(\tau^r(x))\| \\ &\qquad\qquad \leq \sup_{t\in[0,T]} \|x_n(\lambda_n(t)) - x(t)\| + \sup_{t\in(\tau^r(x),\lambda_n^{-1}(\tau^r(x))]} \|x(t) - x(\tau^r(x))\|,\end{aligned}$$

where the inequality uses the assumption $\tau^r(x_n) \leq \tau^r(x)$. The first summand in the bound tends to zero by (4.10) and the second by the right continuity of x and $\lambda_n^{-1}(\tau^r(x)) \to \tau^r(x)$. Hence $d(x_n(\,\cdot\, \wedge \tau^r(x_n)), x(\,\cdot\, \wedge \tau^r(x))) \to 0$ by [8, Prop. 3.5.3, p. 119] and the lemma follows. □

The next task in the proof of Theorem 4.1 is to construct a limiting process.

Lemma 4.9. *Fix $r_0 > 0$. There exists a process Z^{r_0} with paths a.s. in $\mathcal{C}_d$, such that for all but countably many $r \in (0, r_0)$ it holds that*

$$(Z_{n_k}(\,\cdot\, \wedge \tau^r_{n_k}), \tau^r_{n_k}) \Rightarrow (Z^{r_0}(\,\cdot\, \wedge \tau^r), \tau^r), \tag{4.11}$$

where $\tau^r_n = \tau^r(Z_n)$ is given in Theorem 4.1, $\tau^r = \tau^r(Z^{r_0})$ is defined in (4.9) and $\Rightarrow$ denotes the weak convergence of probability measures on $\mathcal{D}_d \times [0, \infty]$. Furthermore, the law of $\|Z^{r_0}(\,\cdot\, \wedge \tau^r)\|$ equals that of a Bessel process (of dimension greater than one) stopped at level r. In particular it holds that $(Z^{r_0}(\,\cdot\, \wedge \tau^r), \tau^r) \in \mathcal{D}_d \times \mathbb{R}_+$ a.s.

Proof. Lemma 4.2 implies the existence of a convergent subsequence $(\tilde{Z}^{r_0}_{n_k})_{k\in\mathbb{N}}$ of the sequence $(\tilde{Z}^{r_0}_n)_{n\in\mathbb{N}}$ defined in (4.5). Denote its limit by Z^{r_0}. By (4.4) and the definition of the metric $d : \mathcal{D}_d \times \mathcal{D}_d \to \mathbb{R}_+$ in [8, Eq. (5.2), p. 117], which induces the Skorokhod topology, it holds that

$$d(\tilde{Z}^{r_0}_{n_k}, Z_{n_k}(\,\cdot\, \wedge \tau^{r_0}_{n_k})) \leq e^{-\eta_{n_k}} \xrightarrow{\mathrm{P}} 0 \qquad \text{as } k \to \infty.$$

It hence follows that the sequence $(Z_{n_k}(\,\cdot\, \wedge \tau^{r_0}_{n_k}))_{k\in\mathbb{N}}$ also converges weakly to Z^{r_0}. Furthermore, by [8, Thm 3.10.2, p. 148] and assumption (4.1), the process Z^{r_0} is continuous, i.e. the support of its law is contained in $\mathcal{C}_d$.

Pick $r \in (0, r_0)$. It follows from Lemmas 4.3 and 4.7 and the mapping theorem (see [4, p. 20]) that the joint convergence in (4.11) holds for all but countably many $r < r_0$. Furthermore, from (4.11) we have that $\|Z_{n_k}(\,\cdot\, \wedge \tau^r_{n_k})\| \Rightarrow \|Z^{r_0}(\,\cdot\, \wedge \tau^r)\|$ for all but countably many $r < r_0$. By assumption in Theorem 4.1, the weak limit of $\|Z_{n_k}\|$ is a Bessel process. Hence, again by Lemmas 4.3 and 4.7, the fact that a Bessel process has continuous trajectories and the mapping theorem [4, p. 20], the law of $\|Z^{r_0}(\,\cdot \wedge \tau^r)\|$ equals that of a Bessel process stopped at level r for all but countably many $r < r_0$. The final statement in the lemma is equivalent to saying that a Bessel process of dimension greater than one reaches every positive level with probability one. This is immediate in the transient case. In the recurrent case it follows from the fact that the height of excursions away from zero is not bounded. □

Define the function $F_{i,j} : \mathcal{D}_d \times \mathbb{R}_+ \to \mathbb{R}$ by the formula $F_{i,j}(y,T) := \int_0^T a_{ij}(y(s))\mathrm{d}s$ for any $i,j \in \{1,\dots,d\}$, where a_{ij} is a coefficient in the generator G in Theorem 4.1.

Lemma 4.10. *Fix $r_0 > 0$. Then for all but countably many $r \in (0, r_0)$, the sequence of processes $F_{i,j}(Z_{n_k}, \cdot \wedge \tau^r_{n_k}) = (F_{i,j}(Z_{n_k}, t \wedge \tau^r_{n_k}); t \geq 0)$ converges weakly to the process $F_{i,j}(Z^{r_0}, \cdot \wedge \tau^r) = (F_{i,j}(Z^{r_0}, t \wedge \tau^r); t \geq 0)$ as $k \to \infty$ for any $i,j \in \{1,\dots,d\}$,*

Remark 4.11. In the proof of [8, Thm 7.4.1, p. 355], the statement of the lemma is used implicitly and follows directly from the continuity assumption on a_{ij} in [8, Thm 7.4.1, p. 355] (which implies that $F_{i,j}$ is itself continuous at any continuous path) and the analogue of the the weak limit in (4.11). In our case the coefficient a_{ij} is discontinuous at the origin and the process $\|Z^{r_0}\|$ may visit zero infinitely many times. Hence we must rely on the more detailed information about the limit law $\|Z^{r_0}(\,\cdot \wedge \tau^r)\|$. In particular, we use the fact that the Bessel process of dimension greater than one is a continuous semimartingale and apply the occupation times formula to quantify the amount of time it spends around zero.

Proof of Lemma 4.10. Let $\varepsilon > 0$ and take smooth functions $\phi_1^\varepsilon, \phi_2^\varepsilon : \mathbb{R}_+ \to [0,1]$ satisfying $\phi_1^\varepsilon(u) = 1$ for all $u \geq \varepsilon$, $\phi_1^\varepsilon(u) = 0$ for all $u \leq \varepsilon/2$ and $\phi_1^\varepsilon(u) + \phi_2^\varepsilon(u) = 1$ for all $u \in \mathbb{R}_+$. Let

$$F_{i,j}^{k,\varepsilon}(x,T) := \int_0^T a_{ij}(x(s))\phi_k^\varepsilon(\|x(s)\|)\mathrm{d}s, \qquad \text{where } k \in \{1,2\}.$$

Then since a_{ij} is continuous on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ and ϕ_1^ε is continuous and vanishes in a neighbourhood of 0, we have that $F_{i,j}^{1,\varepsilon} : \mathcal{D}_d \times \mathbb{R}_+ \to \mathbb{R}$ is continuous at any point $(x,T) \in \mathcal{C}_d \times \mathbb{R}_+$. Hence (4.11) in Lemma 4.9 implies the convergence $F_{i,j}^{1,\varepsilon}(Z_{n_k}, \cdot \wedge \tau^r_{n_k}) \Rightarrow F_{i,j}^{1,\varepsilon}(Z^{r_0}, \cdot \wedge \tau^r)$ for all but countably many $r < r_0$.

Consider now $F_{i,j}^{2,\varepsilon} : \mathcal{D}_d \times \mathbb{R}_+ \to \mathbb{R}$. Since a_{ij} is globally bounded, there exists a constant $C > 0$ such that

$$|F_{i,j}^{2,\varepsilon}(x,T)| \leq C \int_0^T \phi_2^\varepsilon(\|x(s)\|)\mathrm{d}s \qquad \forall (x,T) \in \mathcal{D}_d \times \mathbb{R}_+. \tag{4.12}$$

By Lemma 4.9, we may assume that $\|Z^{r_0}(\,\cdot \wedge \tau^r)\|$ is a Bessel process (of dimension greater than one) stopped at level r. The random field $(L_t(a))_{t,a\in\mathbb{R}_+}$ of Bessel local times exists by [24, Thm VI.1.7] since the process is a continuous semimartingale with the local martingale component equal to Brownian motion. Furthermore, it is well known that $(L_t(a))_{t,a\in\mathbb{R}_+}$ has a bi-continuous modification, i.e. the map $(t,a) \mapsto L_t(a)$ is a.s. continuous on $\mathbb{R}_+^2$. Then, by the occupation times formula [24, p. 224] and (4.12) we get

$$\sup_{t\in\mathbb{R}_+} |F_{i,j}^{2,\varepsilon}(Z^{r_0}, t \wedge \tau^r)| \leq C \int_0^{\tau^r} \phi_2^\varepsilon(\|Z^{r_0}(s)\|)\mathrm{d}s = C \int_0^\varepsilon \phi_2^\varepsilon(a) L_{\tau^r}(a)\mathrm{d}a, \tag{4.13}$$

since the quadratic variation of $\|Z^{r_0}(\,\cdot \wedge \tau^r)\|$ is dominated by that of the Brownian motion and the support of ϕ_2^ε is contained in $[0,\varepsilon]$. Since $(x,t) \mapsto \int_0^t \phi_2^\varepsilon(\|x(s)\|)\mathrm{d}s$ is continuous on $\mathcal{D}_d \times \mathbb{R}_+$, Lemma 4.9 and the mapping theorem [4, p. 20] imply

$$\sup_{t\in\mathbb{R}_+} |F_{i,j}^{2,\varepsilon}(Z_{n_k}, t \wedge \tau^r_{n_k})| \leq C \int_0^{\tau^r_{n_k}} \phi_2^\varepsilon(\|Z_{n_k}(s)\|)\mathrm{d}s \Rightarrow C \int_0^{\tau^r} \phi_2^\varepsilon(\|Z^{r_0}(s)\|)\mathrm{d}s. \tag{4.14}$$

If the convergence in the lemma fails, there exists a bounded uniformly continuous map $h:\mathcal{C}_1\to\mathbb{R}$ (with the uniform topology on $\mathcal{C}_1$) and ε_0 such that

$$|\mathbb{E}\, h\circ F_{i,j}(Z^{r_0},\,\cdot\wedge\tau^r)-\mathbb{E}\, h\circ F_{i,j}(Z_{n_k},\,\cdot\wedge\tau^r_{n_k})|>\varepsilon_0\qquad\forall k\in\mathbb{N},\tag{4.15}$$

where we have passed to a subsequence without changing the notation. Then there exists $\delta>0$ such that if $x,y\in\mathcal{C}_1$ satisfy $\sup_{t\in\mathbb{R}_+}|x(t)-y(t)|<\delta$, then $|h(x)-h(y)|<\varepsilon_0/6$. Fix a monotone sequence $\varepsilon_n\downarrow 0$ and note that we may assume that δ/C is not an atom of $\int_0^{\varepsilon_n}\phi_2^{\varepsilon_n}(a)L_{\tau^r}(a)\mathrm{d}a$ for any $n\in\mathbb{N}$, where C is the constant in (4.13) and (4.14). Note that by the inequality in (4.14) and the fact that $F_{i,j}=F^{1,\varepsilon}_{i,j}+F^{2,\varepsilon}_{i,j}$ we have

$$|\mathbb{E}\, h\circ F_{i,j}(Z_{n_k},\,\cdot\wedge\tau^r_{n_k})-\mathbb{E}\, h\circ F^{1,\varepsilon}_{i,j}(Z_{n_k},\,\cdot\wedge\tau^r_{n_k})|\le\varepsilon_0/6+C_h\mathbb{P}\left[\int_0^{\tau^r_{n_k}}\phi_2^{\varepsilon}(\|Z_{n_k}(s)\|)\mathrm{d}s>\delta/C\right]$$

any $\varepsilon>0$ and some constant $C_h>0$. By the dominated convergence theorem there exists ε_n such that

$$\mathbb{P}\left[\int_0^{\varepsilon_n}\phi_2^{\varepsilon_n}(a)L_{\tau^r}(a)\mathrm{d}a>\delta/C\right]<\frac{\varepsilon_0}{12C_h}.\tag{4.16}$$

By Lemma 4.9 and since δ/C is not an atom of $\int_0^{\varepsilon_n}\phi_2^{\varepsilon_n}(a)L_{\tau^r}(a)\mathrm{d}a$, there exists $k_0\in\mathbb{N}$ such that for all $k\ge k_0$ we have

$$\mathbb{P}\left[\int_0^{\tau^r_{n_k}}\phi_2^{\varepsilon_n}(\|Z_{n_k}(s)\|)\mathrm{d}s>\delta/C\right]<\mathbb{P}\left[\int_0^{\varepsilon_n}\phi_2^{\varepsilon_n}(a)L_{\tau^r}(a)\mathrm{d}a>\delta/C\right]+\frac{\varepsilon_0}{12C_h}<\frac{\varepsilon_0}{6C_h}.$$

Hence it holds that

$$|\mathbb{E}\, h\circ F_{i,j}(Z_{n_k},\,\cdot\wedge\tau^r_{n_k})-\mathbb{E}\, h\circ F^{1,\varepsilon_n}_{i,j}(Z_{n_k},\,\cdot\wedge\tau^r_{n_k})|<\varepsilon_0/3\qquad\forall k\ge k_0.\tag{4.17}$$

Since we already know $F^{1,\varepsilon}_{i,j}(Z_{n_k},\,\cdot\wedge\tau^r_{n_k})\Rightarrow F^{1,\varepsilon}_{i,j}(Z^{r_0},\,\cdot\wedge\tau^r)$, there exists $k_1\ge k_0$, such that

$$|\mathbb{E}\, h\circ F^{1,\varepsilon_n}_{i,j}(Z_{n_k},\,\cdot\wedge\tau^r_{n_k})-\mathbb{E}\, h\circ F^{1,\varepsilon_n}_{i,j}(Z^{r_0},\,\cdot\wedge\tau^r)|<\varepsilon_0/3\qquad\forall k\ge k_1.\tag{4.18}$$

Similarly, by (4.13) and (4.16), we get

$$\begin{aligned}|\mathbb{E}\, h\circ F_{i,j}(Z^{r_0},\,\cdot\wedge\tau^r)-\mathbb{E}\, h\circ F^{1,\varepsilon_n}_{i,j}(Z^{r_0},\,\cdot\wedge\tau^r)|&<\frac{\varepsilon_0}{6}+C_h\mathbb{P}\left[\int_0^{\varepsilon_n}\phi_2^{\varepsilon_n}(a)L_{\tau^r}(a)\mathrm{d}a>\delta/C\right]\\&<\frac{\varepsilon_0}{3}.\end{aligned}$$

This inequality, coupled with (4.17), (4.18) and the triangle inequality, contradicts the statement in (4.15), which proves the lemma. □

Lemma 4.10 is key in proving that the processes in (4.19) are true martingales, which will in turn imply that the limit Z^{r_0} is a solution of the stopped martingale problem. We establish the martingale property in the next lemma.

Lemma 4.12. *Fix $r_0>0$ and pick $r\in(0,r_0)$. Then the components of the process $Z^{r_0}(\,\cdot\wedge\tau^r)$ are martingales. Moreover, for any $i,j\in\{1,\dots,d\}$, the following process is a martingale:*

$$Z^{r_0,i}(\,\cdot\wedge\tau^r)Z^{r_0,j}(\,\cdot\wedge\tau^r)-\int_0^{\cdot\wedge\tau^r}a_{ij}(Z^{r_0}(s))\mathrm{d}s\tag{4.19}$$

Proof. Recall that the sequence $(\tilde{Z}_n^{r_0})_{n\in\mathbb{N}}$, defined in (4.5), is relatively compact by Lemma 4.2. Furthermore, the process Z^{r_0} was defined as a weak limit of a convergent subsequence $(\tilde{Z}_{n_k}^{r_0})_{k\in\mathbb{N}}$. For any $i,j\in\{1,\dots,d\}$ the processes $\tilde{Z}_{n_k}^{r_0,i}$ and $\tilde{A}_{n_k}^{ij}$ (see (4.5) for definition) give rise to martingales $\tilde{Z}_{n_k}^{r_0,i}\tilde{Z}_{n_k}^{r_0,j}-\tilde{A}_{n_k}^{ij}$ (see the argument following the display in (4.6)). Hence, for any index $i\in\{1,\dots,d\}$ and $k\in\mathbb{N}$, we have that

$$\mathbb{E}[(\tilde{Z}_{n_k}^{r_0,i}(t))^2]=\mathbb{E}[Z_{n_k}^i(0)^2]+\mathbb{E}\left[\tilde{A}_{n_k}^{ii}(t)\right]\qquad\text{for all } t\geq 0.$$

Thus by (4.2), (4.7) and the assumption on the square integrability of $Z_{n_k}(0)$ in Theorem 4.1, we have that $\sup_{k\in\mathbb{N}}\mathbb{E}[\|\tilde{Z}_{n_k}^{r_0}(t)\|^2]<\infty$ and hence the family $(\|\tilde{Z}_{n_k}^{r_0}(t)\|)_{k\in\mathbb{N}}$ is uniformly integrable for every $t\geq 0$.

To prove that the components of Z^{r_0} are martingales with respect to the natural filtration $(\sigma(Z_u^{r_0}:u\in[0,s]),s\in\mathbb{R}_+)$, note first that each σ-algebra $\sigma(Z_u^{r_0}:u\in[0,s])$ is generated by the π-system of events of the form $\{Z^{r_0}(s_1)\in A_1,\dots,Z^{r_0}(s_p)\in A_p\}$ for any $p\in\mathbb{N}$ and $s_1,\dots,s_p\in[0,s]$, where $A_1,\dots,A_p$ are rectangular boxes in $\mathbb{R}^d$. Hence it is sufficient to show that for any $0\leq s_1<\dots s_p\leq s<t$ and a non-negative, bounded, continuous $f:\mathbb{R}^d\otimes\mathbb{R}^p\to\mathbb{R}$ it holds that

$$\mathbb{E}[\left(Z^{r_0,i}(t)-Z^{r_0,i}(s)\right)f(Z^{r_0}(s_1),\dots,Z^{r_0}(s_p))]=0. \tag{4.20}$$

By the Skorokhod representation theorem [8, Thm 3.1.8, p. 102] we may assume that the zero mean random variables $\left(\tilde{Z}_{n_k}^{r_0,i}(t)-\tilde{Z}_{n_k}^{r_0,i}(s)\right)f(\tilde{Z}_{n_k}^{r_0}(s_1),\dots,\tilde{Z}_{n_k}^{r_0}(s_p))$ converge almost surely as $k\to\infty$ to the random variable in (4.20). Furthermore, since f is bounded, this sequence is uniformly integrable by the argument in the first paragraph of this proof. This implies the convergence in L^1 and hence the identity in (4.20). Since Z^{r_0} is a martingale, so is $Z^{r_0}(\,\cdot\wedge\tau^r)$ for any $r\in(0,r_0)$.

Consider now the process in (4.19). We start by establishing the following fact.

Claim. For any $i,j\in\{1,\dots,d\}$ and all but countably many $r\in(0,r_0)$ it holds that

$$\tilde{Z}_{n_k}^{r_0,i}(\cdot\wedge\tau_{n_k}^r)\tilde{Z}_{n_k}^{r_0,j}(\cdot\wedge\tau_{n_k}^r)-\tilde{A}_{n_k}^{ij}(\cdot\wedge\tau_{n_k}^r)\Rightarrow Z^{r_0,i}(\cdot\wedge\tau^r)Z^{r_0,j}(\cdot\wedge\tau^r)-\int_0^{\cdot\wedge\tau^r}a_{ij}(Z^{r_0}(s))\mathrm{d}s,$$

where the stopping times $\tau_n^r=\tau^r(Z_n)$ and $\tau^r=\tau^r(Z^{r_0})$ are as in Lemma 4.9.

Proof of Claim. By definition it holds that $\tilde{Z}_{n_k}^{r_0}\Rightarrow Z^{r_0}$. Hence, as in the proof of Lemma 4.9, since Z^{r_0} has continuous trajectories it follows from Lemmas 4.3 and 4.7 and the mapping theorem [4, p. 20] that $\tilde{Z}_{n_k}^{r_0}(\,\cdot\wedge\tau_{n_k}^r)\Rightarrow Z^{r_0}(\,\cdot\wedge\tau^r)$. Thus it holds that $\tilde{Z}_{n_k}^{r_0,i}(\,\cdot\wedge\tau_{n_k}^r)\tilde{Z}_{n_k}^{r_0,j}(\,\cdot\wedge\tau_{n_k}^r)\Rightarrow Z^{r_0,i}(\,\cdot\wedge\tau^r)Z^{r_0,j}(\,\cdot\wedge\tau^r)$.

To prove the claim it therefore suffices to show that $\tilde{A}_{n_k}^{ij}(\cdot\wedge\tau_{n_k}^r)\Rightarrow\int_0^{\cdot\wedge\tau^r}a_{ij}(Z^{r_0}(s))\mathrm{d}s$. With this in mind, we note that

$$\tilde{A}_{n_k}^{ij}(\,\cdot\wedge\tau_{n_k}^r)=U_k+V_k+F_{i,j}(Z_{n_k},\,\cdot\wedge\tau_{n_k}^r), \tag{4.21}$$

where $U_k:=\tilde{A}_{n_k}^{ij}(\,\cdot\wedge\tau_{n_k}^r)-A_{n_k}^{ij}(\,\cdot\wedge\tau_{n_k}^r)\xrightarrow{\mathrm{P}}0$ by (4.4)–(4.5) and $V_k:=A_{n_k}^{ij}(\,\cdot\wedge\tau_{n_k}^r)-F_{i,j}(Z_{n_k},\,\cdot\wedge\tau_{n_k}^r)\xrightarrow{\mathrm{P}}0$ by the assumption in (4.3). The representation of $\tilde{A}_{n_k}^{ij}(\,\cdot\wedge\tau_{n_k}^r)$ in (4.21), [8, Cor. 3.3.3, p. 110] and Lemma 4.10 imply

$$\tilde{A}_{n_k}^{ij}(\,\cdot\wedge\tau_{n_k}^r)\Rightarrow\int_0^{\cdot\wedge\tau^r}a_{ij}(Z^{r_0}(s))\mathrm{d}s, \tag{4.22}$$

and the claim follows.

Since $\tilde{Z}_{n_k}^{r_0,i}\tilde{Z}_{n_k}^{r_0,j}-\tilde{A}_{n_k}^{ij}$ is a martingale by the argument following (4.6), the stopped process $M_k:=\tilde{Z}_{n_k}^{r_0,i}(\,\cdot\wedge\tau_{n_k}^r)\tilde{Z}_{n_k}^{r_0,j}(\,\cdot\wedge\tau_{n_k}^r)-\tilde{A}_{n_k}^{ij}(\,\cdot\wedge\tau_{n_k}^r)$ is also a martingale for every

$k \in \mathbb{N}$. Hence the process in (4.19) will be a martingale by the analogous argument to the one that established the martingale property of $Z^{r_0,i}$ above, if we prove that for any $t \geq 0$ the family of random variables $\{M_k(t) : k \in \mathbb{N}\}$ is uniformly integrable. With this in mind, note that $2|\tilde{A}^{ij}_{n_k}| \leq \tilde{A}^{ii}_{n_k} + \tilde{A}^{jj}_{n_k}$ since the matrix $\tilde{A}_{n_k}$ is non-negative definite. The elementary inequality $2|\tilde{Z}^{r_0,i}_{n_k}\tilde{Z}^{r_0,j}_{n_k}| \leq (\tilde{Z}^{r_0,i}_{n_k})^2 + (\tilde{Z}^{r_0,j}_{n_k})^2$ implies

$$|M_k(t)| \leq \tilde{Z}^{r_0,i}_{n_k}(t \wedge \tau^r_{n_k})^2 + \tilde{Z}^{r_0,j}_{n_k}(t \wedge \tau^r_{n_k})^2 + \tilde{A}^{ii}_{n_k}(t \wedge \tau^r_{n_k}) + \tilde{A}^{jj}_{n_k}(t \wedge \tau^r_{n_k}).$$

Since the sequence $(\tilde{A}^{ii}_{n_k}(t \wedge \tau^r_{n_k}) + \tilde{A}^{jj}_{n_k}(t \wedge \tau^r_{n_k}))_{k\in\mathbb{N}}$ is bounded in L^1 by (4.2) and (4.7), $\{M_k(t) : k \in \mathbb{N}\}$ will be uniformly integrable if $\{\tilde{Z}^{r_0,i}_{n_k}(t \wedge \tau^r_{n_k})^2 : k \in \mathbb{N}\}$ is uniformly integrable for all $i \in \{1, \dots, d\}$. Note that by (4.6), for any $r \in (0, r_0)$, we have that

$$\tilde{Z}^{r_0,i}_{n_k}(t \wedge \tau^r_{n_k})^2 \leq 3\left(\sup_{n\in\mathbb{N}} \|Z_n(0)\|^2 + 4r_0^2 + \sup_{0\leq s\leq t\wedge\tau^r_{n_k}} \|Z_{n_k}(s) - Z_{n_k}(s-)\|^2\right).$$

The right-hand side converges in L^1 by (4.1). Hence $\{\tilde{Z}^{r_0,i}_{n_k}(t \wedge \tau^r_{n_k})^2 : k \in \mathbb{N}\}$ is uniformly integrable and the lemma follows for all but countably many $r \in (0, r_0)$. Note however that there exist $r_n \uparrow r_0$ such that the martingale properties in the lemma hold for all r_n. Since a stopped martingale is a martingale, the lemma follows for all $r \in (0, r_0)$. □

Proof of Theorem 4.1. By Lemma 4.12 and Itô's formula for continuous semimartingales, the process Z^{r_0} constructed in the proof of Lemma 4.9 solves the stopped martingale problem (see [8, p. 216] for the precise definition) $(G, v, \{x \in \mathbb{R}^d : \|x\| < r\})$ for any $r \in (0, r_0)$. Since the martingale problem (G, v) is well-posed, by [8, Thm 4.6.1, p. 216] there exists a unique solution to the stopped martingale problem. Furthermore, if Z is a solution of the martingale problem (G, v) on $\mathcal{D}_d$, then $Z(\,\cdot \wedge \tau^r(Z))$ must be a solution to the stopped martingale problem by the optional sampling theorem (cf. [8, pp. 216–217]), where $\tau^r(Z)$ is defined in (4.9). In particular (since $r_0 > 0$ is arbitrary) for all but countably many $r > 0$, any subsequence of $Z_n(\,\cdot \wedge \tau^r_n)$, where τ^r_n is defined in Lemma 4.9, has by Lemma 4.9 a further subsequence that converges weakly to the law of the process $Z(\,\cdot \wedge \tau^r(Z))$. It hence follows that the entire sequence must be convergent, $Z_n(\,\cdot \wedge \tau^r_n) \Rightarrow Z(\,\cdot \wedge \tau^r(Z))$, for all but at most countably many $r > 0$.

In order to prove that this implies $Z_n \Rightarrow Z$, note that $\tau^r(Z) \to \infty$ a.s. as $r \to \infty$, since the paths of Z are in $\mathcal{D}_d$ (in fact in $\mathcal{C}_d$), and it holds that

$$d(Z, , Z(\,\cdot \wedge \tau^r(Z)) \leq e^{-\tau^r(Z)} \to 0 \qquad \text{a.s. as } r \to \infty,$$

where $d : \mathcal{D}_d \times \mathcal{D}_d \to \mathbb{R}_+$, defined in [8, Eq. (5.2), p. 117], is the Skorokhod metric. Pick any uniformly continuous and bounded map $h : \mathcal{D}_d \to \mathbb{R}$. This class of maps is convergence determining [8, Prop. 3.4.4, p. 112]. Pick $\varepsilon > 0$ and let $\delta \in (0, 1)$ satisfy: if $d(x, y) < \delta$ then $|h(x) - h(y)| < \varepsilon/6$. Let $C_h > 0$ satisfy $\sup_{x\in\mathcal{D}_d} |h(x)| < C_h$. By Lemmas 4.3 and 4.7 and the mapping theorem (see [4, p. 20]), there exists $r > 0$ such that $\tau^r_n \Rightarrow \tau^r(Z)$ and $\mathbb{P}[\tau^r(Z) \leq \log(1/\delta)] < \varepsilon/(12C_h)$. Without loss of generality we may assume that $\log(1/\delta)$ is not an atom of $\tau^r(Z)$. Hence we may choose $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$ we have $\mathbb{P}[\tau^r_n \leq \log(1/\delta)] < \varepsilon/(6C_h)$ and $|\,\mathbb{E}\, h(Z_n(\,\cdot \wedge \tau^r_n)) - \mathbb{E}\, h(Z(\,\cdot \wedge \tau^r(Z)))| < \varepsilon/6$. This implies the inequalities

$$\begin{aligned}
|\,\mathbb{E}\, h(Z_n) - \mathbb{E}\, h(Z)| &\leq |\,\mathbb{E}\, h(Z_n) - \mathbb{E}\, h(Z_n(\cdot \wedge \tau^r_n))| + |\,\mathbb{E}\, h(Z_n(\,\cdot \wedge \tau^r_n)) - \mathbb{E}\, h(Z(\,\cdot \wedge \tau^r(Z)))| \\
&\quad + |\,\mathbb{E}\, h(Z(\,\cdot \wedge \tau^r(Z))) - \mathbb{E}\, h(Z)| \\
&\leq \mathbb{P}[\tau^r_n > \log(1/\delta)]\frac{\varepsilon}{6} + \frac{\varepsilon}{3} + \mathbb{P}[\tau^r(Z) > \log(1/\delta)]\frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} \leq \varepsilon,
\end{aligned}$$

which completes the proof. □

4.2 Proof of Theorem 1.2

Recall the definition of the scaled process $\widetilde{X}_n = (\widetilde{X}_n(t), t \geq 0)$ in (1.2) in terms of the chain $X = (X_m, m \in \mathbb{Z}_+)$, $\widetilde{X}_n(t) = n^{-1/2} X_{\lfloor nt \rfloor}$ for $t \in \mathbb{R}_+$. Theorem 1.2 will follow from Theorem 4.1 and the main result of [10]:

Lemma 4.13. *Suppose that* (A0)–(A4) *hold. Without loss of generality assume that* $U = 1$. *Then* $\|\widetilde{X}_n\|$ *converges weakly to the* V*-dimensional Bessel process started at* 0.

Proof of Theorem 1.2. Define $A_n(t) = \frac{1}{n} \sum_{m=0}^{\lfloor nt \rfloor - 1} M(X_m)$, where $M(\mathbf{x})$ is the covariance matrix of the increment at $\mathbf{x} \in \mathbb{X}$ and, as usual, an empty sum is 0. Define $Z_n := \widetilde{X}_n$ and note that $Z_n^i Z_n^j - A_n^{ij}$ is a local martingale for all $i, j \in \{1, \ldots, d\}$. By Lemma 4.13 we have $\|Z_n\| \Rightarrow \mathrm{BES}^V(0)$ as $n \to \infty$. Let $a(\mathbf{x}) := \sigma^2(\hat{\mathbf{x}})$ be a non-negative definite matrix-valued function on $\mathbb{R}^d$, where σ^2 satisfies (A3)–(A6). Let the generator G be defined as in Theorem 4.1 for this coefficient a. Then the $\mathcal{C}_d$ martingale problem for $(G, \delta_{\mathbf{0}})$ is well-posed by Theorem 1.1, where $\delta_{\mathbf{0}}$ denotes the delta measure on $\mathbb{R}^d$ concentrated at the origin. In order to apply Theorem 4.1, it remains to establish the assumptions (4.1), (4.2) and (4.3) for Z_n and A_n. Condition (4.1) follows from [10, Lem 2]. Since by assumption $|M_{ij}(\mathbf{y})| \leq \sup_{\mathbf{x} \in \mathbb{X}: \|x\| \geq r} \|M(\mathbf{x})\| < \infty$ for a sufficiently large $r > 0$ and any $\mathbf{y} \in \mathbb{X}$ with $\|\mathbf{y}\| \geq r$, condition (4.2) follows from $\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \max_{0 \leq m \leq \lfloor nT \rfloor} |M_{ij}(X_m)| = 0$. Finally (4.3) is verified by [10, Lem 5] for the coordinate functional $\phi : \mathbb{R}^d \otimes \mathbb{R}^d \to \mathbb{R}$, $\phi(B) = B_{ij}$. Thus Theorem 4.1 yields Theorem 1.2. □

References

[1] Alili, L., Chaumont, L., Graczyk, P. and Żak, T.: Inversion, duality and Doob h-transforms for self-similar Markov processes. *Electron. J. Probab.* **22** (2017), paper no. 20, 18 pp. MR-3622890

[2] Barlow, M., Pitman, J. and Yor, M.: On Walsh's Brownian motions. Séminaire de Probabilités, XXIII, pp. 275–293, Lecture Notes in Math. **1372**, *Springer-Verlag*, Berlin, 1989. MR-1022917

[3] Bertoin, J.: Lévy Processes. *Cambridge University Press*, Cambridge, 1996. MR-1406564

[4] Billingsley, P.: Convergence of Probability Measures. 2nd ed., *John Wiley & Sons, Inc.*, New York, 1999. MR-1700749

[5] Cherny, A.S.: Convergence of some integrals associated with Bessel processes. *Theory Probab. Appl.* **45** (2001) 195–209. Translated from *Teor. Veroyatnost. i Primenen* **45** (2000) 251–267 (Russian). MR-1967756

[6] Dugundji, J.: Topology. *Allyn and Becon, Inc.*, Boston, 1966. MR-0193606

[7] Embrechts, P. and Hofert, M.: A note on generalized inverses. *Math. Meth. Oper. Res.* **77** (2013) 423–432. MR-3072795

[8] Ethier, S.N. and Kurtz, T.G.: Markov Processes. Characterization and Convergence. *John Wiley & Sons, Inc.*, New York, 1986. MR-0838085

[9] Georgiou, N., Menshikov, M.V., Mijatović, A. and Wade, A.R.: Anomalous recurrence properties of many-dimensional zero-drift random walks. *Adv. in Appl. Probab.* **48A** (2016) 99–118. MR-3539299

[10] Georgiou, N., Mijatović, A. and Wade, A.R.: A radial invariance principle for non-homogeneous random walks. *Electron. Commun. Probab.* **23** (2018) paper no. 56, 11 pp. MR-3863912

[11] Hsu, E.P.: Stochastic Analysis on Manifolds. *American Mathematical Society*, Providence, 2002. MR-1882015

[12] Ikeda, N. and Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. 2nd ed., *North-Holland Publishing Company*, Amsterdam, 1989. MR-0637061

[13] Itô, K. and McKean Jr., H.P.: Diffusion Processes and Their Sample Paths. 2nd corr. printing, *Springer-Verlag*, Berlin, 1974. MR-0345224

[14] Jost, J.: Riemannian Geometry and Geometric Analysis. 4th ed., *Springer-Verlag*, Berlin, 2005. MR-2165400

[15] Kent, J.: Time-reversible diffusions. *Adv. in Appl. Probab.* **10** (1978) 819–835. MR-0509218

[16] Kent, J.: Eigenvalue expansion for diffusion hitting times. *Z. Wahr. ver. Gebiete* **52** (1980) 309–319. MR-0576891

[17] Kingman, J.F.C.: Poisson Processes. *Oxford University Press*, Oxford, 1993. MR-1207584

[18] Krylov, N.V.: Controlled Diffusion Processes. Reprint of the 1980 edition, *Springer-Verlag*, Berlin, 2009. MR-2723141

[19] Lamperti, J.: A new class of probability limit theorems. *J. Math. Mech.* **11** (1962) 749–772. MR-0148120

[20] Menshikov, M., Popov, S. and Wade, A.: Non-homogeneous Random Walks. *Cambridge University Press*, Cambridge, 2016. MR-3587911

[21] Mijatović, A. and Urusov, M.: Convergence of integral functionals of one-dimensional diffusions. *Electron. Commun. Probab.* **17** (2012) paper no. 61, 13 pp. MR-3005734

[22] Pinsky, R.G.: Positive Harmonic Functions and Diffusion. *Cambridge University Press*, Cambridge, 1995. MR-1326606

[23] Pitman, J. and Yor, J.: A decomposition of Bessel bridges. *Z. Wahr. ver. Gebiete* **59** (1982) 425–457. MR-0656509

[24] Revuz, D. and Yor, M.: Continuous Martingales and Brownian Motion. 3rd ed., *Springer-Verlag*, Berlin, 1999. MR-1725357

[25] Rogers, L.C.G. and Williams, D.: Diffusions, Markov Processes, and Martingales. Vol. 2. Itô Calculus. Reprint of the second (1994) edition. *Cambridge University Press*, Cambridge, 2000. MR-1780932

[26] Stroock, D.W. and Varadhan, S.R.S.: Multidimensional Diffusion Processes. *Springer-Verlag*, Berlin-New York, 1979. MR-0532498

[27] Stroock, D.W. and Yor, M.: Some remarkable martingales. *Séminaire de Probabilités XV 1979/80*. Lecture Notes in Mathematics **850** (1981) pp. 590–603. MR-0622590

[28] Vuolle-Apiala, J.: Excursion theory for rotation invariant Markov processes. *Probab. Theory Relat. Fields* **93** (1992) 153–158. MR-1176722

Acknowledgments. NG was supported by the Heilbronn Institute for Mathematical Research. NG and AW were supported in part by the EPSRC grant EP/J021784/1. AM is supported by the EPSRC grant EP/P003818/1 and a Fellowship at The Alan Turing Institute, sponsored by the Programme on Data-Centric Engineering funded by Lloyd's Register Foundation. AM is also supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1. The authors are grateful for the comments and suggestions of an anonymous referee, which have led to several improvements in the paper.