Degenerate Stochastic Differential Equations and Super-Markov Chains

S.R. Athreya[1,2] , M.T. Barlow[1] , R.F. Bass[3], and E.A. Perkins[1]

Abstract

We consider diffusions corresponding to the generator

$$\mathcal{L}f(x) = \sum_{i=1}^{d} x_i \gamma_i(x) \frac{\partial^2}{\partial {x_i}^2} f(x) + b_i(x) \frac{\partial}{\partial x_i} f(x),$$

$x \in \mathbb{R}_+^d$, for continuous $\gamma_i, b_i : \mathbb{R}_+^d \to \mathbb{R}$ with γ_i nonnegative. We show uniqueness for the corresponding martingale problem under certain non-degeneracy conditions on b_i, γ_i and present a counter-example when these conditions are not satisfied. As a special case, we establish uniqueness in law for some classes of super-Markov chains with state dependent branching rates and spatial motions.

1. Introduction

Let

$$\gamma_i, b_i : \mathbb{R}_+^d \to \mathbb{R}, \text{ be continuous functions and each } \gamma_i \text{ be strictly positive.} \tag{1.1}$$

We consider the operator $\mathcal{L}$ on $C^2(\mathbb{R}_+^d)$ defined by

$$\mathcal{L}f(x) = \sum_{i=1}^{d} x_i \gamma_i(x) \frac{\partial^2 f}{\partial {x_i}^2}(x) + b_i(x) \frac{\partial f}{\partial x_i}(x), \qquad x \in \mathbb{R}_+^d. \tag{1.2}$$

We also consider the diffusion X_t associated to $\mathcal{L}$; this is the process on $\mathbb{R}_+^d$ that solves the stochastic differential equation

$$dX_t^i = \sqrt{2X_t^i \gamma_i(X_t)} dB_t^i + b_i(X_t)\, dt, \qquad X_t^i \geq 0, \quad i = 1, \ldots, d, \tag{1.3}$$

where B_t is a standard d-dimensional Brownian motion. The purpose of this paper is to prove uniqueness of the martingale problem for the operator $\mathcal{L}$. As is well known, this is equivalent to proving weak uniqueness (i.e., uniqueness in law) to the solution of (1.3).

---

1. Research supported in part by an NSERC research grant.

2. Research supported in part by The Pacific Institute for Mathematical Sciences.

3. Research supported in part by NSF grant DMS9988496

Keywords: Stochastic differential equations, martingale problem, elliptic operators, degenerate operators, diffusions, Bessel processes. Classification: Primary 60H10.

Running Head: Stochastic differential equations.

Let $\Omega = C(\mathbb{R}_+, \mathbb{R}_+^d)$, let $X_t(\omega) = \omega(t)$ be the usual coordinate functions, let $\mathcal{F}$ be the Borel σ-field on Ω, and let $\mathcal{F}_t$ be the canonical right-continuous filtration on $(\Omega, \mathcal{F})$. If ν is a probability on $\mathbb{R}_+^d$, we say $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ with initial law ν (or $MP(\nu, \mathcal{L})$) if

$$\begin{aligned} \mathbb{P}(X_0 \in \cdot) = \nu(\cdot), \quad & \text{and } N_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds \\ & \text{is an } \mathcal{F}_t\text{-martingale under } \mathbb{P} \text{ for each } f \in C_b^2(\mathbb{R}_+^d, \mathbb{R}). \end{aligned} \tag{1.4}$$

We say that an $\mathbb{R}_+^d$-valued process $(Y_t, t \geq 0)$ with a.s. continuous paths is a solution to the martingale problem for $\mathcal{L}$ if its probability law is a solution in the above sense. Y (or its law) is a strong Markov solution if, in addition, it is a strong Markov process with respect to $\mathcal{F}_t^Y = \cap_{s>t}\sigma(Y_r, r \leq s)$. Add ∂ to $\mathbb{R}^d$ as a discrete "cemetery" point.

Here is our main result. Let $\|x\| = \max_{i=1,\dots,d} |x_i|$.

Theorem 1.1. *Let $\mathcal{L}$ be as in (1.2) and suppose (1.1) holds. Assume that for all $i = 1, \dots, d$,*

$$b_i(x) > 0, \qquad x \in \partial\mathbb{R}_+^d, \tag{1.5}$$

$$|\, b_i(x) \,| \leq C(1 + \|x\|), \qquad x \in \mathbb{R}_+^d. \tag{1.6}$$

(a) For any initial law ν, there exists a unique solution to the martingale problem for $\mathcal{L}$.
(b) If $\mathbb{P}^x$ is the solution in (a) with initial law δ_x, then $(\mathbb{P}^x, X_t)$ is a strong Markov process, and for any bounded measurable function f on $\mathbb{R}_+^d$, its resolvent

$$S_\lambda f(x) = \mathbb{E}^x\Big(\int_0^\infty e^{-\lambda t} f(X_t)\, dt\Big),$$

is a continuous function of x.

The following corollary of Theorem 1.1 is relevant for applications to superprocesses. Let

$$T_0 = T_0(X) = \inf\{t \geq 0 : X_t = 0\}.$$

Corollary 1.2. *Assume (1.1) and (1.6), and let $\mathcal{L}$ be as in* (1.2).
(a) If for some $C \geq 0$,

$$b_i(x) > -Cx_i \text{ on } \partial\mathbb{R}_+^d - \{0\}, \tag{1.7}$$

then for every initial law there is a solution $\mathbb{P}$ to the martingale problem for $\mathcal{L}$ and $\mathbb{P}(X(\cdot \wedge T_0) \in \cdot)$ is unique.
(b) If, in addition to (1.6) and (1.7),

$$\sum_{i=1}^d b_i(x) = 0 \text{ on } \mathbb{R}_+^d, \tag{1.8}$$

then there is a unique solution to the martingale problem for $\mathcal{L}$.

Note that (1.7) implies that $b_i(x) > 0$ if $x_i = 0$ and $x \neq 0$.

The degeneracy of the diffusion coefficients on the boundary means that one cannot apply the results of [SV79] directly to establish uniqueness for the martingale problem (1.4). Uniqueness of the martingale problem is of course equivalent to uniqueness in law for solutions (in $\mathbb{R}_+^d$) to the stochastic differential equation (1.3), and would follow from pathwise uniqueness. However, the presence of the square root in (1.3) means that the coefficients are not Lipschitz, the standard condition for pathwise uniqueness. In the special case where $\gamma_i(x) = \gamma_i(x_i)$ depends only on x_i, each $\sqrt{x_i \gamma_i(x_i)}$ is Hölder continuous of order $\frac{1}{2}$, and each b_i is Lipschitz continuous, pathwise uniqueness can be proved by a well-known local time argument (see [YW71] or Sec. V.40 of [RW87]). However, this method fails in general, and even in the case when γ_i, b_i are smooth and bounded away from zero, pathwise uniqueness for (1.3) remains an open question. (See [S00] for a related but special case where pathwise uniqueness can be established.) Therefore we needed to develop new techniques to handle (1.4).

Our principal reason for studying this problem comes from the theory of superprocesses with state dependent interactions. A superprocess on a state space E is a diffusion taking values in the space $M_F(E)$ of finite measures on E. To describe it more precisely consider a conservative generator A of a Hunt process ξ on E, a bounded continuous drift function $g : E \mapsto \mathbb{R}$, and a bounded continuous branching rate $2\gamma : E \mapsto \mathbb{R}_+$. Write $\mu(\varphi)$ for $\int \varphi d\mu$, and let $D(A)$ denote the domain of A. The Dawson-Watanabe superprocess with drift g, branching rate 2γ, and spatial motion A is the $M_F(E)$-valued diffusion X whose law on $C(\mathbb{R}_+, M_F(E))$ is characterized by the law of X_0 and the following martingale problem: for each $\varphi \in D(A)$

$$X_t(\varphi) = X_0(\varphi) + \int_0^t X_s((A+g)\varphi)ds + M_t(\varphi), \qquad (MP)_{X_0}$$

where $M_t(\varphi)$ is a continuous martingale with square function $\langle M(\varphi)\rangle_t = \int_0^t X_s(2\gamma\varphi^2)ds$. See [D93] and [P01] for this and further background on superprocesses.

These processes arise as the large population (N), small mass ($1/N$) limit of a system of branching ξ-processes. At $y \in E$ each particle branches with rate $N\gamma(y)$ and produces a random number of offspring with mean $1+g(y)/N$ and variance approaching 1 as $N \to \infty$. The independent behaviour of the individual particles makes these models amenable to detailed mathematical study and is the key fact underlying the usual exponential duality proof of uniqueness in $(MP)_{X_0}$. From the perspective of potential biological applications it is clearly desirable to have the individuals in the population interact through the drift g, spatial motion A, or branching rate 2γ. One could allow these quantities to depend on the current state X_t and hence introduce $g : M_F(E) \times E \mapsto \mathbb{R}$, $\gamma : M_F(E) \times E \mapsto \mathbb{R}_+$ and state dependent generators $(A_\mu)_{\mu \in M_F(E)}$ defined on a common domain, D. It is not hard to see that, under appropriate continuity conditions, the interactive analogues of the above branching particle systems in which b, γ and A are replaced by their state dependent analogues produce a tight sequence of processes whose limit points will satisfy, for each $\varphi \in D$,

$$X_t(\varphi) = X_0(\varphi) + \int_0^t X_s((A_{X_s} + g(X_s))\varphi)ds + M_t(\varphi), \qquad (IMP)_{X_0}$$

where $M_t(\varphi)$ is a continuous martingale with $\langle M(\varphi)\rangle_t = \int_0^t X_s(2\gamma(X_s)\varphi^2)ds$. The question then is: Are solutions to $(IMP)_{X_0}$ unique in law?

In the case when γ is constant, uniqueness can be proved for a wide class of g and A – see [D78], [P92], [P95], [P01], [DK99] and [K98]. The change of measure technique in [D78] allows one to assume $g \equiv 0$ in quite general settings, and we will do so below. The case when γ depends on X is much harder, although weak uniqueness has been proved in some special cases by duality methods – see [M98] and [DEFMPX00].

In the Fleming-Viot setting [DM95] established uniqueness in the martingale problem for some state-dependent sampling (i.e. branching) rates which are very close to constant. They used the Stroock-Varadhan perturbation technique in an infinite dimensional setting (using completely different methods than those in this work). However, the strength of their norms meant that it was not possible to localize and so obtain a general uniqueness result. Athreya and Tribe [AT00] used a particle dual to calculate the moments for the solutions of a class of parabolic stochastic PDEs, some of which could be interpreted as examples of (IMP) with a purely local branching interaction, $E = \mathbb{R}$ and $Af = f''/2$. These duality arguments can be used to show uniqueness for certain degenerate stochastic differential equations as well, though under rather strict conditions on γ_i and b_i.

If the state space E is the finite set $\{1, \ldots, d\}$ then (IMP) reduces to the d-dimensional stochastic differential equation in the following example.

Example 1 (Super-Markov Chains). When $E = \{1, \ldots, d\}$ is finite uniqueness in law for a class of the interactive branching mechanisms described above follows from our main result. By [D78] we may assume that the drift $g \equiv 0$. In this setting $M_F(E) = \mathbb{R}^d_+$ and one easily sees that $(IMP)_{X_0}$ reduces to (1.4) where

$$b_i(x) = \sum_{j=1}^{d} x_j q_{ji}(x), \quad x = (x_1, \ldots, x_d) \in \mathbb{R}^d_+, \tag{1.9}$$

and so

$$\mathcal{L}f(x) = \sum_{i=1}^{d} \Big[x_i \gamma_i(x) \frac{\partial^2 f}{\partial {x_i}^2}(x) + \sum_{j=1}^{d} x_j q_{ji}(x) \frac{\partial f}{\partial x_i}(x) \Big]. \tag{1.10}$$

Here $q_{ji}(x)$ is the jump rate from site j to site i in when the population is x, and $\gamma_i(x)$ is the corresponding branching rate at site i.

Corollary 1.3. *Let $q_{ij} : \mathbb{R}^d_+ \to \mathbb{R}$, for $i, j = 1, \ldots, d$, be bounded, continuous and satisfy*

$$\sum_{j=1}^{d} q_{ij}(x) = 0, \quad i = 1, \ldots, d, \qquad x \in \mathbb{R}^d_+, \tag{1.11}$$

$$q_{ij}(x) \geq 0 \quad \text{for all } i \neq j, \qquad x \in \mathbb{R}^d_+, \tag{1.12}$$

$$q_{ij}(x) > 0 \text{ for all } i \neq j, \text{ if } x_j = 0 \text{ and } x \neq 0, \qquad j = 1, \ldots, d. \tag{1.13}$$

If $\mathcal{L}$ is given by (1.10), then there exists a unique solution for the martingale problem for $\mathcal{L}$.

Proof. We apply Corollary 1.2 with $b_i(x) = \sum_{j=1}^{d} x_j q_{ji}(x)$. Clearly (1.1) and (1.6) hold, and (1.8) is immediate from (1.11). To verify (1.7), note that if $C > \sup_i \|q_{ii}\|_\infty$ and $x_i > 0$, then

$$b_i(x) \geq x_i q_{ii}(x) > -C x_i.$$

On the other hand if $x_i = 0$, then $q_{ji}(x) > 0$ for all $j \neq i$ by (1.12) and so $b_i(x) > 0$ unless $x = 0$. This verifies (1.7) in either case. Corollary 1.2 (b) now gives the desired result. $\square$

Remark 1.4. By using a stopping argument as in the proof of Corollary 1.2(a), (see Section 7), one can weaken the non-degeneracy condition on γ_i in (1.1) to $\gamma_i(x) > 0$ for non-zero $x \in \mathbb{R}_+^d$.

Conditions (1.11) and (1.12) simply assert that $(q_{ij}(\cdot))$ is a state dependent generator of a chain. Unfortunately, however, (1.13) rules out such simple chains as nearest neighbour random walks on $\{1, \ldots, d\}$. One might hope that the condition (1.5) could be replaced by

$$b_i(x) \geq 0, \text{ for all } i, \quad x \in \partial \mathbb{R}_+^d, \tag{1.14}$$

but this is not possible in general – we give a one-dimensional counter-example in Section 8. See, however, [BP01], which considers the case $b_i(x) \geq 0$ on ∂R_+^d with γ_i and b_i Hölder continuous.

Example 2 (Generalized Mutually Catalytic Branching). Assume $q_{ij}^k : \mathbb{R}_+^d \to \mathbb{R}$ for $1 \leq i, j \leq d$ are bounded and continuous, and for each $k \leq K$, $(q_{ij}^k(x))$ is the generator of a Markov chain on $\{1, \ldots, d\}$ (i.e., (1.11) and (1.12) hold for each q^k) and (1.13) holds for each of these K generators. Let $\gamma_{k,i} : \mathbb{R}_+^K \to (0, \infty)$ be continuous for $i = 1, \ldots, d$ and $k = 1, \ldots, K$. Consider the system of stochastic differential equations in $\mathbb{R}_+^d$ for $1 \leq i \leq d, \ 1 \leq k \leq K$,

$$dX_t^{k,i} = \sum_{j=1}^d X_t^{k,j} q_{ji}^k(X_t^k)dt + \sqrt{2\gamma_{k,i}(X_t^{1,i}, \ldots, X_t^{K,i})X_t^{k,i}}dB_t^{k,i}. \tag{1.15}$$

Here $X_t^k = (X_t^{k,1}, \ldots, X_t^{k,d})$, and $B^{1,1}, \ldots, B^{K,d}$ are Kd independent one-dimensional Brownian motions. This represents K populations undergoing state dependent migration on d sites where the branching rate of the kth population at site i is a function $(\gamma_{k,i})$ of the mass of the K populations at the same site i. The intuition is that the presence of the different types at a site effects the branching of the other types at the site.

We claim that Corollary 1.2 (and its proof) gives uniqueness in law for the solutions of (1.15). By a result of Krylov it suffices to prove uniqueness of strong Markov solutions starting from an arbitrary constant initial condition (see Theorem 12.2.4 of [SV79] and its proof which applies equally well to diffusions in $\mathbb{R}_+^d$). Note first that for some $C > 0$,

$$b_{k,i}(x^1, \ldots, x^K) = \sum_{j=1}^d x^{k,j} q_{ji}^k(x^k) > -Cx^{k,i} \text{ if } x^k \neq 0.$$

This follows exactly as in Corollary 1.3. Now let $T^k = \inf\{t : X_t^k = 0\}$ and $T = \min_{k \leq K} T^k$. As in the proof of Corollary 1.2 $X(\cdot \wedge T)$ is unique in law. Since the total mass of each population is a non-negative local martingale it will stick at zero when it hits zero. Hence after time T one population is identically zero and the other $K - 1$ will satisfy a martingale problem of the same type. The obvious induction now gives uniqueness in law of $X(T + \cdot)$. Piecing the solution together we obtain uniqueness in law of X, as required.

The standard mutually catalytic branching model (see [DP98]) has $K = 2$, the branching rate of each type is given by the amount of the other type at the site, and so

$$\gamma_{1,i}(x^{1,i}, x^{2,i}) = x^{2,i}, \quad \gamma_{2,i}(x^{1,i}, x^{2,i}) = x^{1,i}.$$

For this model and constant (q_{ij}) uniqueness in law can be proved by duality, but the argument does not extend to more general branching rates. The nondegeneracy condition we have imposed on the $\gamma_{k,i}$ unfortunately excludes this model from those covered by our result. However, for more than two types (see Fleischmann and Xiong [FX00]) the result above seems to be the first uniqueness result which allows branching rate of one type to depend on the other types at the site.

Example 3 (Stepping Stone Models). Assume $q_{ij} : [0,1]^d \to \mathbb{R}$ for $1 \le i, j \le d$ are bounded continuous and for each x, $(q_{ij}(x))$ is the generator of a Markov chain on $\{1, \dots, d\}$ such that

$$\sum_{i=1}^{d} q_{ij}(x) = 0 \text{ for all } x, \tag{1.16}$$

and

$$q_{ij}(x) > 0 \text{ for all } i \neq j \text{ whenever } x_j = 0 \text{ or } 1. \tag{1.17}$$

For $i = 1, \dots, d$, let γ_i be a strictly positive continuous function on $[0,1]^d$. Then Corollary 1.2 implies that for any fixed $X_0 \in [0,1]^d$, there is a solution $\{X_t, t \ge 0\} \in [0,1]^d$ of

$$dX_t^i = \sum_{j=1}^{d} X_t^j q_{ji}(X_t)dt + \sqrt{\gamma_i(X_t)X_t^i(1 - X_t^i)}dB_t^i \tag{1.18}$$

that is unique in law. Here again B^i, $i = 1, \dots, d$ are independent one-dimensional Brownian motions. X_t^i represents the proportion of the population with a given genotype at site i, $q_{ij}(\cdot)$ is the state-dependent migration rate from state i to state j and $\gamma_i(\cdot)$ is the state-dependent sampling rate at site i. Existence of solutions is standard.

Uniqueness is a local result in that it suffices to show each starting point has a neighbourhood on which the coefficients equal other coefficients for which uniqueness holds. This follows as in the well-known Stroock-Varadhan localization result on $\mathbb{R}^d$ (see Theorem 6.6.1 of [SV79] or Theorem VI.3.4 in [B97]). For starting points in the interior of $[0,1]^d$ we may change the diffusion coefficient outside a small open ball so that it is uniformly elliptic and then apply standard results from [SV79]. For initial points x in $\partial[0,1]^d$ satisfying $\max x_i < 1$, local uniqueness is clear from Corollary 1.3. If x is in the boundary with some coordinates equal to 1, we want to make the transformation taking X_t^i to $1 - X_t^i$ for those i where $x_i = 1$. We do this by setting $\psi_i(y) = 1 - y$ if $x_i = 1$ and $\psi_i(y) = y$ otherwise. We then perform the transformation $(y_1, \dots, y_d) \to (\psi_1(y_1), \dots, \psi_d(y_d))$. After this transformation we have reduced the problem to the situation where the starting point satisfies $\max x_i < 1$.

The interested reader may now combine the previous two examples to obtain uniqueness in law for a multi-type stepping stone model in which each type migrates according to its own state dependent Q-matrix and the sampling rate at each site may depend on the proportion of each of the types at the particular site. This last example was motivated by recent work of Greven, Klenke, and Wakolbinger [GKW99].

In Section 2 we give an overview of the proof of Theorem 1.1. Section 3 contains the necessary resolvent bounds, while Section 4 establishes key properties of the zeros of Bessel functions which are needed in Section 3. Section 5 and Section 6 deal with norm-finiteness and continuity of the resolvent respectively. Theorem 1.1 and Corollary 1.2 are proved in Section 7, and in Section 8 we give the one-dimensional counterexample which shows that we cannot weaken the condition $b_i > 0$ on $\partial\mathbb{R}^d_+$. Constants which appear in the statements of Lemmas (and propositions), say Lemma 5.2 are denoted by $c_{5.2}$. Elsewhere in the paper, c, c_i denote constants whose value may change from line to line.

Acknowledgment. We thank D. Dawson for a number of helpful conversations on the uniqueness problem for interactive branching. We also thank the referee for a very careful reading of the paper.

2. Overview of proof

In this section we give an outline of the proof of Theorem 1.1, and state the main results that we will need. Since existence of a solution to the martingale problem for $\mathcal{L}$ is relatively straightforward (see Section 7), we concentrate here on uniqueness.

If X is a process in $\mathbb{R}^d_+$ and $B \subset \mathbb{R}^d$, write

$$T_B = \inf\{t \geq 0 : X_t \in B\}, \quad \tau_B = \inf\{t \geq 0 : X_t \in B^c\},$$

for the first hitting times of B and B^c. We will sometimes use the notation $\tau_B(X)$, $T_B(X)$ when the process X is not clear from the context. Fix $M > 0$, and define the upper boundary of $[0, M]^d$ by

$$U = U_M = \{(x_1, \dots, x_d) \in [0, M]^d : x_1 \vee \cdots \vee x_d = M\}.$$

Let $\tau_M = \tau_{[0,M)^d}$. If ν is a probability on $[0, M)^d$, we say that a continuous $\mathbb{R}^d_+ \cup \{\partial\}$-valued process X is a solution to the stopped martingale problem for $(\mathcal{L}, [0, M)^d)$ with initial law ν, (or $SMP(\nu, \mathcal{L}, [0, M)^d)$) if X_0 has law ν, $X_t = \partial$ for $t \geq \tau_M$, and for each $f \in C^2_b([0, M]^d)$ the process $N^f_{t\wedge(\tau_M-)}$ is a continuous martingale. (Here N^f is as in (1.4) and $N^f_{t\wedge(\tau_M-)}$ equals $N^f_{\tau_M-}$ if $t \geq \tau_M$.) If

$$\Omega^\partial = \{\omega \in C(\mathbb{R}_+, \mathbb{R}^d \cup \{\partial\}) : \text{ whenever } 0 \leq s < t,\ \omega(s) = \partial \text{ implies } \omega(t) = \partial\}$$

and $\mathcal{F}^\partial$ is its Borel σ-field, then we also say that the law of X on $(\Omega^\partial, \mathcal{F}^\partial)$ is a solution of the stopped martingale problem for $(\mathcal{L}, [0, M]^d)$ with initial law ν.

A localisation argument, similar to that in [SV79] or [B97], (see Section 7) reduces the proof of Theorem 1.1 to the following case.

Proposition 2.1. *For any $\varepsilon > 0$ there is a $K = K(\varepsilon, d)$ so that if $b_i(\cdot)$, $\gamma_i(\cdot)$, $1 \leq i \leq d$ are as in (1.1), and there exist constants $b^0_i > 0$, $\gamma^0_i > 0$, $\varepsilon > 0$ such that*

$$\|b^0_i - b_i(\cdot)\|_\infty \leq (2K)^{-1}, \quad \|\gamma^0_i - \gamma_i(\cdot)\|_\infty \leq (2K)^{-1}, \quad i = 1, \dots, d, \tag{2.1}$$

$$\varepsilon \le b_i(x), \gamma_i(x), b_i^0, \gamma_i^0 \le \varepsilon^{-1}, \qquad x \in \mathbb{R}_+^d, \quad i = 1, \dots, d, \tag{2.2}$$

$$2\frac{b_i(x)}{\gamma_i(x)} \ge \frac{b_i(y)}{\gamma_i(y)} + \frac{\varepsilon^2}{2}, \qquad x, y \in \mathbb{R}_+^d, \quad i = 1, \dots, d, \tag{2.3}$$

then uniqueness of solutions holds for $MP(\mathcal{L}, \nu)$ *for any law* ν *on* $\mathbb{R}_+^d$.

Most of the remainder of this paper will be concerned with proving Proposition 2.1. Let

$$\mathcal{L}^0 f(x) = \sum_{i=1}^d \gamma_i^0 x_i \frac{\partial^2 f}{\partial x_i^2}(x) + \sum_{i=1}^d b_i^0 \frac{\partial f}{\partial x_i}(x). \tag{2.4}$$

Note that $\mathcal{L}^0$ is the generator of a process whose components are independent scaled copies of the square of a Bessel process of dimension $2b_i^0/\gamma_i^0$ (see Sec. V.48 of [RW87]). We write Y for this process killed (i.e. set equal to the cemetery state ∂) at time T_U. Analytically this means we impose zero boundary conditions on U. Set $b^0 = (b_1^0, \dots b_d^0)$ and $\gamma^0 = (\gamma_1^0, \dots \gamma_d^0)$. Let $R_\lambda = R_\lambda^{b^0,\gamma^0}$ denote the resolvent of this killed process Y. The measure on $[0, M]^d$ which makes $\mathcal{L}^0$ with these boundary conditions formally self-adjoint is $\mu(dx) = \prod_{i=1}^d x_i^{(b_i^0/\gamma_i^0)-1} dx_i$. We write L^2 for $L^2([0, M]^d, \mu)$ and $\|\cdot\|_2$ will denote the associated norm, hence suppressing dependence on (b^0, γ^0) in our notation.

Those familiar with the localization technique (and those not) may find (2.3) rather puzzling. It arises because, unlike the Brownian case, the natural reference measures μ depend on the constants (b^0, γ^0). It is used in the proof of Proposition 2.3 below and more specifically in the proof of Lemma 5.3.

We now give the key perturbation estimate needed to carry out the Stroock-Varadhan argument. This result introduces the constant $K(\varepsilon, d)$ needed in Proposition 2.1. Set

$$C_0^2 = C_0^2([0, M]^d) = \{f \in C^2([0, M]^d) : f|_U = 0\}.$$

Proposition 2.2. *There exists a dense subspace* $\mathcal{D}_0 \subset L^2([0, M]^d, \mu)$ *with*

$$R_\lambda(\mathcal{D}_0) \subset \mathcal{D}_0 \subset C_0^2 \tag{2.5}$$

satisfying the properties below. For each $\varepsilon > 0$ *there exists* $K = K(\varepsilon, d)$, *independent of* M, *such that if* $\varepsilon \le b_i^0, \gamma_i^0 \le \varepsilon^{-1}$, *then (recall* $R_\lambda = R_\lambda^{b^0,\gamma^0}$*),*

$$\sum_{i=1}^d \Big(\big\| x_i \frac{\partial^2}{\partial x_i^2} R_\lambda f \big\|_2 + \big\| \frac{\partial}{\partial x_i} R_\lambda f \big\|_2 \Big) \le K \|f\|_2 \text{ for all } \lambda > 0 \text{ and } f \in \mathcal{D}_0. \tag{2.6}$$

In particular the operators $x_i(\partial^2/\partial x_i^2)R_\lambda$ *and* $(\partial/\partial x_i)R_\lambda$ *extend uniquely to bounded operators on* L^2 *satisfying (2.6) for all* $f \in L^2$.

Using Theorem 12.2.4 of [SV79], we will see that uniqueness in general will follow if we can prove uniqueness for Borel strong Markov solutions of the stopped martingale problem for any M. So let $(X_t, \mathbb{P}_k^x)$, $k = 1, 2$, be two Borel strong Markov processes, such that for each x

the probability $\mathbb{P}^x_k$ is a solution to the stopped martingale problem for $(\mathcal{L},[0,M)^d)$ started at x. Let

$$S^k_\lambda f(x)=\mathbb{E}^x_k\int_0^\infty e^{-\lambda t}f(X_t)dt=\mathbb{E}^x_k\int_0^{T_U} e^{-\lambda t}f(X_t)dt,\qquad k=1,2,$$

where the process X_t is killed (set equal to ∂) upon exiting $[0,M)^d$ and $f(\partial)=0$. Some elementary stochastic calculus (see Section 7) shows that for f in $\mathcal{D}_0$,

$$S^k_\lambda f(x)=R_\lambda f(x)+S^k_\lambda(\mathcal{L}-\mathcal{L}^0)R_\lambda f(x),\qquad k=1,2. \tag{2.7}$$

We want to use a perturbation argument in L^2, but $\sup\{|S^k_\lambda f(x)|:\|f\|_2\le 1\}$ will not be finite in general; in fact $|R_\lambda f(x)|$ can be infinite even if $\|f\|_2<\infty$. So we integrate (2.7) with respect to the measure $\nu(dx)=\rho(x)\mu(dx)$ for $\rho\in L^2$, take the difference for $k=1,2$, and obtain

$$\int(S^1_\lambda-S^2_\lambda)f(x)\nu(dx)=\int(S^1_\lambda-S^2_\lambda)(\mathcal{L}-\mathcal{L}^0)R_\lambda f(x)\nu(dx).$$

Set $\theta=\sup\{|\int(S^1_\lambda-S^2_\lambda)f(x)\nu(dx)|:\|f\|_2\le 1\}$. Using Proposition 2.2 and (2.1), we obtain

$$\Big|\int(S^1_\lambda-S^2_\lambda)f(x)\nu(dx)\Big|\le\frac{\theta}{2}\|f\|_2.$$

Taking the supremum over $f\in C^2([0,M]^d)$ such that $\|f\|_2\le 1$, we obtain

$$\theta\le\frac{\theta}{2}. \tag{2.8}$$

To eliminate the possibility that $\theta=\infty$ we apply the following proposition.

Proposition 2.3. *Let X be a solution of $SMP(\nu,\mathcal{L},[0,M)^d)$, where $\nu(dx)=\rho(x)d\mu(x)$ for some $\rho\in L^2([0,M]^d,\mu)$. Set $S_\lambda f=\mathbb{E}\int_0^\infty e^{-\lambda t}f(X_t)\,dt$, where $f(\partial)=0$. If there are constants $\varepsilon>0$, b^0_i, γ^0_i satisfying (2.1), (2.2), and (2.3), then for all $\lambda>0$*

$$\sup\{|S_\lambda f|:\|f\|_2\le 1\}\le\frac{2\|\rho\|_2}{\lambda}<\infty.$$

This implies that $\theta<\infty$, and so we conclude from (2.8) that $\theta=0$. It follows that $S^1_\lambda f(x)=S^2_\lambda f(x)$ for almost every x. To extend this to equality everywhere, we prove that $S^i_\lambda f$ are continuous.

Proposition 2.4. *Assume γ_i and b_i are as in Theorem 1.1. Let $M\in(0,\infty]$ and assume $\{\mathbb{P}^x:x\in[0,M)^d\cup\{\partial\}\}$ is a collection of probabilities on $(\Omega^\partial,\mathcal{F}^\partial)$ such that:*
(i) For each $x\in[0,M)^d$, $\mathbb{P}^x$ is a solution of the stopped martingale problem for $(\mathcal{L},[0,M)^d)$ with initial law δ_x, and $\omega(\cdot)\equiv\partial$ $\mathbb{P}^\partial$-a.s.,
(ii) $(\mathbb{P}^x,X_t)$ is a Borel strong Markov process.
Then for any bounded measurable function f on $[0,M)^d$, and any $\lambda\ge 0$,

$$S_\lambda f(x)=\mathbb{E}^x\Big(\int_0^\infty e^{-\lambda t}f(X_t)dt\Big)$$

is a continuous function in $x \in [0, M)^d$.

Note that if $M = \infty$, solutions to the stopped martingale problem for $(\mathcal{L}, [0, M)^d)$ are just solutions to the martingale problem for $\mathcal{L}$. This Proposition allows us to conclude that $S^1_\lambda f(x) = S^2_\lambda f(x)$ for every x. It is then standard to deduce from this the uniqueness of the solution to the martingale problem.

We say a few words about the proofs of Propositions 2.2, 2.3, and 2.4. To get the estimates we need for Proposition 2.2, we first consider the case of one dimension in Section 3. We look at an eigenfunction decomposition of L^2, and an explicit calculation shows that if V_λ is the resolvent operator for a one-dimensional scaled squared Bessel process, then $d(V_\lambda)/dx$ is a bounded operator on L^2. This entails some detailed estimates of Bessel functions and their zeros, which is done in Section 4. To handle the d-dimensional estimates, we use the fact that the transition density for the process corresponding to $\mathcal{L}^0$ factors into a product of transition densities for one dimensional processes and some eigenvalue analysis. After we have a bound on the first derivatives, a bound on $x_i \partial^2(R_\lambda)/\partial x_i^2$ is easily achieved using some more eigenvalue calculations and the diagonal form of the diffusion matrix.

The proof of Proposition 2.3, given in Section 5, is similar to the proof in [SV79] of the analogous estimate. We "freeze" the coefficients of (1.3) at a finite number of fixed times, and prove finiteness of the corresponding resolvent. Combining this with a uniform estimate on the resolvent obtained from Proposition 2.2, and using an analogue of (2.7), we then obtain bounds independent of the approximation, which allows us to take a limit. Some care must be taken here because, unlike the uniformly elliptic setting, the natural reference measures depend on the "frozen" coefficients. This complication leads to the odd-looking condition (2.3).

Proposition 2.4 is proved in Section 6. Its analogue for uniformly elliptic diffusions is a well-known result of Krylov and Safonov (see e.g., Section V.7, p. 116 of [B97]). The proof uses the classical Girsanov theorem, scaling and the result of Krylov and Safonov to prove that X_t enters certain sets with positive probability.

In Section 7 we carry out the details of the argument described above.

Remark 2.5. One should note that the above approach also simplifies the analytic part of the classical results of Stroock and Varadhan on uniformly elliptic diffusions [SV79]. Instead of using L^p estimates in the analogue of Proposition 2.2, which require some difficult estimates for singular operators, one can get by with much simpler L^2 estimates which follow easily from Parseval's equality – see for example Appendix A.0 and A.1 in [SV79]. The price for this is that one must use Krylov selection to reduce uniqueness to the Markovian setting and the Krylov-Safonov results to obtain continuity of the resolvent operators. Both of these, however, have nice probabilistic proofs.

3. Resolvent Bounds

Fix $M > 0$, let $b, \gamma \in (0, \infty)$, and let

$$\mathcal{A}f(x) = \gamma x f''(x) + b f'(x), \qquad x \in [0, M], \quad f \in D(\mathcal{A}),$$

be the infinitesimal generator of a scaled squared Bessel diffusion killed when it hits M. In this section γ and b are constants and do not depend on x. Let

$$J_a(x) = \sum_{m=0}^{\infty} \frac{(-1)^m (x/2)^{a+2m}}{m!\Gamma(a+m+1)}, \quad x \geq 0, \tag{3.1}$$

be the Bessel function of the first kind with parameter $a > -1$, and let

$$w_k = w_k(b') \text{ be the } k\text{th positive zero of } J_{b'-1}(\cdot) \text{ for } b' > 0,\ k \in \mathbb{N}. \tag{3.2}$$

Proposition 3.1. *Let $b' = b/\gamma$, and set*

$$\varphi_k(x) = \frac{J_{b'-1}(w_k\sqrt{\frac{x}{M}})}{\sqrt{M}x^{(b'-1)/2} \mid J_{b'}(w_k) \mid}, \quad x \in [0, M]. \tag{3.3}$$

Then $\varphi_k(x)$ is in $C^2([0, M])$ with $\varphi_k(M) = 0$, φ_k satisfies

$$\mathcal{A}\varphi_k = -\lambda_k\varphi_k \text{ on } [0, M], \tag{3.4}$$

where

$$\lambda_k = \frac{\gamma w_k^2}{4M}, \tag{3.5}$$

and $\{\varphi_k : k \in \mathbb{N}\}$ is a complete orthonormal basis in $L^2([0, M], x^{b'-1}dx)$.

Proof. Using (3.1) one can see that $\varphi_k \in C^2([0, M])$. The definition of w_k guarantees that $\varphi_k(M) = 0$. A direct calculation shows that φ_k satisfies (3.4); perhaps the easiest way to see this is to write φ_k as a power series using (3.1) and perform the differentiations term by term. The fact that the φ_k are orthonormal follows from the fact that $\{\sqrt{2z}J_{b'-1}(w_k z)/|J_{b'}(w_k)| : k \in \mathbb{N}\}$ is a complete orthonormal system in $L^2([0, 1]), dz)$ ([H71], p. 264) and the change of variables $z = \sqrt{x/M}$. To check completeness, suppose $f \in L^2([0, M], x^{b'-1}dx)$ is orthogonal to all of the φ_k. By the change of variables $z = \sqrt{x/M}$ the function $F(z) = f(z^2M)z^{b'-1}$ can be seen to belong to $L^2([0, 1], z\,dz)$ and to be orthogonal to $J_{b'-1}(w_k z)$ in this space for all k. Since $\{J_{b'-1}(w_k z)\}$ is a complete basis in $L^2([0, 1], z\,dz)$, then $F(z) = 0$ a.e., which implies that $f(x) = 0$ a.e. □

We will need three technical lemmas on Bessel functions and their zeros. We defer the proofs of Lemmas 3.2, 3.3, and 3.4 to the next section.

Lemma 3.2. *For each $\varepsilon > 0$ there exists $c_{3.2}$ depending only on ε such that for any $b' \in [\varepsilon^2, \varepsilon^{-2}]$ and all $1 \le j \le k$,*

$$\Big|\int_0^1 J_{b'}(w_k z)J_{b'}(w_j z)z^{-1}\,dz\Big| \le c_{3.2}(w_j/w_k)^{b' \wedge (1/4)}.$$

Lemma 3.3. *For each $\varepsilon > 0$ there exists $c_{3.3} > 0$ depending only on ε such that for $b' \in [\varepsilon^2, \varepsilon^{-2}]$ and all $k \in \mathbb{N}$,*

$$w_n \ge c_{3.3}k.$$

Lemma 3.4. *For each $\varepsilon > 0$ there exists $c_{3.4} > 0$ depending only on ε such that for $b' \in [\varepsilon^2, \varepsilon^{-2}]$ and all $k \in \mathbb{N}$,*

$$|J_{b'}(w_k)| \ge c_{3.4}w_k^{-1/2}.$$

We will also need the following classical analysis result–see Theorem 318 in [HLP34]. As it is neat, short and fun, we give an alternate proof.

Proposition 3.5. *Suppose $\nu > 0$ and $K(j,k) = 1_{(j\le k)} j^{\nu-\frac{1}{2}} k^{-\frac{1}{2}-\nu}$. Then*

$$\sum_{1\le j\le k<\infty} |a_j|\,|a_k| K(j,k) \le (\nu\wedge 1/2)^{-1} \sum_{j=1}^{\infty} |a_j|^2.$$

Proof. As the left side is clearly decreasing in ν it suffices to consider $\nu \le 1/2$. Fix N for the moment and consider the bounded linear operator K_N on ℓ^2 defined by $K_N(j,k) = K(j,k)1_{(k\le N)}$ and $(K_N a)_j = \sum_{k=1}^{\infty} K_N(j,k)a_k$. Let $K_N^*(j,k) = K_N(k,j)$ and note

$$\begin{aligned}(K_N^* K_N)(j,k) &= \sum_{m=1}^{\infty} K_N^*(j,m)K_N(m,k) = \sum_{m=1}^{N} K_N(m,j)K_N(m,k)\\ &= \sum_{m=1}^{j\wedge k} m^{\nu-\frac{1}{2}} j^{-\frac{1}{2}-\nu} m^{\nu-\frac{1}{2}} k^{-\frac{1}{2}-\nu} 1_{(j\vee k\le N)}\\ &\le \frac{1}{2\nu}(j\wedge k)^{2\nu} j^{-\frac{1}{2}-\nu} k^{-\frac{1}{2}-\nu} 1_{(j\vee k\le N)}\\ &\le \frac{1}{2\nu}(K_N^*(j,k) + K_N(j,k)).\end{aligned}$$

In the next to last inequality we have used the fact that $\nu \le 1/2$. If $\mathbf{x} = (x_1, x_2, \ldots) \in \ell^2$, let $\mathbf{y} = (|x_1|, |x_2|, \ldots)$. We have

$$|((K_N^* K_N)\mathbf{x})_j| \le ((K_N^* K_N)\mathbf{y})_j \le \frac{1}{2\nu}((K_N^* + K_N)\mathbf{y})_j,$$

so

$$\|(K_N^* K_N)\mathbf{x}\|_{\ell^2} \le \frac{1}{2\nu}(\|K_N^*\| + \|K_N\|)\|\mathbf{y}\|_{\ell^2} = \frac{1}{\nu}\|K_N\|\,\|\mathbf{x}\|_{\ell^2}.$$

Hence

$$\|K_N\|^2 = \|K_N^* K_N\| \le \frac{1}{\nu}\|K_N\|,$$

which implies $\|K_N\| \le \frac{1}{\nu}$. Let $b_j = |a_j|$. By Cauchy-Schwarz,

$$\sum_{1\le j\le k\le N} |a_j|\,|a_k| K(j,k) = \sum_j b_j (K_N b)_j \le \|b\|_{\ell^2}\|K_N b\|_{\ell^2} \le \frac{1}{\nu}\|b\|_{\ell^2}^2 = \frac{1}{\nu}\|a\|_{\ell^2}^2.$$

Now let $N\to\infty$. □

Now let V_λ denote the resolvent associated with the generator $\mathcal{A}$. Set $b' = b/\gamma$. The $L^2([0,M], x^{b'-1}dx)$ norm will be denoted by $\|\cdot\|_2$. Let $\mathcal{D} = \mathcal{D}(b,\gamma)$ be the dense subspace in L^2 consisting of finite linear combinations of the eigenfunctions φ_k. Note that the constant $c_{3.6}$ in the next result does not depend on M.

Proposition 3.6. *For each $\varepsilon > 0$ there exists $c_{3.6} > 0$ depending only on ε such that if $b, \gamma \in [\varepsilon, \varepsilon^{-1}]$, then*

$$\sup_{\lambda>0} \| \frac{dV_\lambda f}{dx} \|_2 \le c_{3.6} \|f\|_2 \quad \textit{for all } f \in \mathcal{D}.$$

Proof. Let $b, \gamma \in [\varepsilon, \varepsilon^{-1}]$, so that $b' = b/\gamma \in [\varepsilon^2, \varepsilon^{-2}]$. From p. 45 of [W44] we have

$$\frac{d}{dz}\Big(z^{-(b'-1)} J_{b'-1}(z) \Big) = -z^{-(b'-1)} J_{b'}(z).$$

From this and (3.3) we have

$$\varphi_k'(x) = \frac{-w_k}{2M|J_{b'}(w_k)|} J_{b'}(w_k \sqrt{x/M}) x^{-b'/2}.$$

This implies that if $f = \sum_{k=1}^N a_k \varphi_k$ then, since $V_\lambda \varphi_k = (\lambda + \lambda_k)^{-1} \varphi_k$,

$$\frac{dV_\lambda f(x)}{dx} = \sum_{k=1}^N \frac{a_k}{\lambda+\lambda_k} \varphi_k'(x) = \sum_{k=1}^N \frac{-a_k}{\lambda+\lambda_k} \Big(\frac{w_k J_{b'}(w_k \sqrt{x} M^{-\frac12})}{2M \mid J_{b'}(w_k) \mid x^{\frac{b'}{2}}} \Big),$$

where the λ_k are as in (3.4). Hence

$$\begin{aligned} &\int_0^M \mid \frac{dV_\lambda f(x)}{dx} \mid^2 x^{b'-1} dx \\ &= \int_0^M \sum_{1\le j,k\le N} \frac{a_k a_j}{(\lambda+\lambda_j)(\lambda+\lambda_k)} \frac{w_k w_j J_{b'}(w_k \sqrt{x} M^{-\frac12}) J_{b'}(w_j \sqrt{x} M^{-\frac12})}{4M^2 \mid J_{b'}(w_k) J_{b'}(w_j) \mid} x^{-1} dx \\ &= \sum_{1\le j,k\le N} \frac{a_k a_j w_k w_j}{4M^2 (\lambda+\lambda_j)(\lambda+\lambda_k) \mid J_{b'}(w_k) J_{b'}(w_j) \mid} \int_0^1 J_{b'}(w_k z) J_{b'}(w_j z) \frac{2dz}{z}. \end{aligned}$$

In the last line we substituted $z = \sqrt{x/M}$.

Set $\nu = b' \wedge (1/4)$ and use Lemmas 3.2, 3.3 and 3.4 to conclude that for some constants which depend only on ε,

$$\begin{aligned} \| \frac{d}{dx} V_\lambda f \|_2^2 &\le c_1 \sum_{1\le j\le k\le N} \frac{|a_j||a_k| w_j^{\frac32} w_k^{\frac32}}{M^2(\lambda+\lambda_j)(\lambda+\lambda_k)} \Big(\frac{w_j}{w_k} \Big)^\nu \\ &= c_1 \sum_{1\le j\le k\le N} \frac{|a_j||a_k| w_j^{\frac32} w_k^{\frac32}}{(M\lambda + \frac{\gamma}{4} w_j^2)(M\lambda + \frac{\gamma}{4} w_k^2)} \Big(\frac{w_j}{w_k} \Big)^\nu \\ &\le c_2 \sum_{j\le k\le N} |a_j|\,|a_k| \Big(\frac{w_j^{\nu-\frac12}}{w_k^{\nu+\frac12}} \Big) \\ &\le c_3 \sum_{j\le k\le N} \frac{|a_j|\,|a_k|}{k^{\nu+\frac12} j^{\frac12-\nu}} \le \nu^{-1} c_3 \sum_{j\le N} |a_j|^2. \end{aligned} \tag{3.6}$$

Here we used Proposition 3.5 in the final line. Since $||f||_2^2 = \sum |a_j|^2$ this completes the proof. $\square$

We now show that the one-dimensional result, Proposition 3.6, is all we need to handle the higher-dimensional situation. Let $b_i^0, \gamma_i^0 \in (0,\infty)$ for $i = 1,\dots,d$, fix $M > 0$, let $\mu_i(dx_i) = x_i^{b_i'-1}dx_i$, where $b_i' = b_i^0/\gamma_i^0$. Define $\mu(dx) = \prod_{i=1}^d x_i^{b_i'-1}dx_i$. Let $\|\cdot\|_2$ denote the $L^2([0,M]^d,\mu)$ norm.

Set

$$\mathcal{A}_j f(x) = \gamma_j^0 x_j \frac{\partial^2 f}{\partial x_j^2}(x) + b_j^0 \frac{\partial f}{\partial x_j}(x), \qquad x \in [0,M]^d,\ 1 \le j \le d,$$

for $f \in C^2([0,M]^d)$ such that $f(x) = 0$ whenever $x \in U_M$. We will also need

$$\overline{\mathcal{A}}_j f(x) = \gamma_j^0 x f''(x) + b_j^0 f'(x), \qquad x \in [0,M]$$

for $f \in C^2([0,M])$ with $f(M) = 0$. Thus $\overline{\mathcal{A}}_j$ is the operator $\mathcal{A}_j$ considered as an operator in one dimension. Let V_λ^j be the resolvent for $\overline{\mathcal{A}}_j$. For each j let $\{\varphi_k^j : k \in \mathbb{N}\}$ be the complete orthonormal system of eigenfunctions for $\overline{\mathcal{A}}_j$ on $L^2([0,M],\mu_j(dx))$ and let λ_k^j be the corresponding eigenvalues. If $\mathbf{k} = (k_1,\dots,k_d)$, then $\varphi_{\mathbf{k}}(x_1,\dots,x_d) = \prod_{j=1}^d \varphi_{k_j}^j(x_j)$ defines a complete orthonormal system in $L^2([0,M]^d,\mu)$. Let $\lambda(\mathbf{k}) = \sum_{j=1}^d \lambda_{k_j}^j$. Recall that

$$\mathcal{L}^0 f(x) = \sum_{j=1}^d \mathcal{A}_j f(x) = \sum_{j=1}^d x_j \gamma_j^0 \frac{\partial^2 f}{\partial x_j^2}(x) + b_j^0 \frac{\partial f}{\partial x_j}(x), \tag{3.7}$$

and therefore

$$\mathcal{L}^0 \varphi_{\mathbf{k}} = -\lambda(\mathbf{k})\varphi_{\mathbf{k}}. \tag{3.8}$$

Proof of Proposition 2.2. Recall that R_λ is the resolvent of the operator $\mathcal{L}^0$ with zero boundary conditions on U_M. Set

$$\mathcal{D}_0 = \{\sum_{\mathbf{k}} a_{\mathbf{k}}\varphi_{\mathbf{k}} : a_{\mathbf{k}} \ne 0 \text{ for only finitely many } \mathbf{k}\}. \tag{3.9}$$

Since $R_\lambda \varphi_{\mathbf{k}} = (\lambda + \lambda(\mathbf{k}))^{-1}\varphi_{\mathbf{k}}$, we have $R_\lambda(\mathcal{D}_0) \subset \mathcal{D}_0 \subset C_0^2$.

We begin by proving that

$$\Big\|\frac{\partial R_\lambda f}{\partial x_j}\Big\|_2 \le c_{3.6}\|f\|_2 \quad \text{for all } f \in \mathcal{D}_0. \tag{3.10}$$

We will do the case $j = 1$; the proof for other j is exactly the same. Suppose

$$f = \sum_{k_1,\dots,k_d=1}^N a_{\mathbf{k}}\varphi_{\mathbf{k}}.$$

Set

$$g(x_1;k_2,\dots,k_d)=\sum_{k_1=1}^N a_{\mathbf{k}}\varphi^1_{k_1}(x_1).$$

Set $\sigma(\mathbf{k})=\lambda^2_{k_2}+\cdots+\lambda^d_{k_d}$. We have

$$V^1_{\lambda+\sigma(\mathbf{k})}g(x_1;k_2,\dots,k_d)=\sum_{k_1=1}^N a_{\mathbf{k}}\frac{1}{\lambda+\lambda^2_{k_2}+\cdots+\lambda^d_{k_d}+\lambda^1_{k_1}}\varphi^1_{k_1}(x_1)=\sum_{k_1=1}^N\frac{a_{\mathbf{k}}}{\lambda+\lambda(\mathbf{k})}\varphi^1_{k_1}(x_1).$$

It follows that

$$\begin{aligned}R_\lambda f(x)&=\sum_{k_1,\dots,k_d=1}^N\frac{a_{\mathbf{k}}}{\lambda+\lambda(\mathbf{k})}\varphi_{\mathbf{k}}(x)\\&=\sum_{k_2,\dots,k_d=1}^N\varphi^2_{k_2}(x_2)\cdots\varphi^d_{k_d}(x_d)(V^1_{\lambda+\sigma(\mathbf{k})}(g(\cdot;k_2,\dots,k_d)))(x_1),\end{aligned}$$

and hence that

$$\frac{\partial R_\lambda f}{\partial x_1}(x)=\sum_{k_2,\dots,k_d=1}^N\varphi^2_{k_2}(x_2)\cdots\varphi^d_{k_d}(x_d)\frac{d}{dx_1}(V^1_{\lambda+\sigma(\mathbf{k})}(g(\cdot;k_2,\dots,k_d)))(x_1).$$

If $\mathbf{m}=(m_1,\dots,m_d)$,

$$\begin{aligned}\|\frac{\partial}{\partial x_1}R_\lambda f\|^2_{L^2(\mu)}=\int\cdots\int&\sum_{k_2,\dots,k_d=1}^N\ \sum_{m_2,\dots,m_d=1}^N\varphi^2_{k_2}(x_2)\varphi^2_{m_2}(x_2)\cdots\varphi^d_{k_d}(x_d)\varphi^d_{m_d}(x_d)\\&\times\frac{d}{dx_1}(V^1_{\lambda+\sigma(\mathbf{k})}(g(\cdot;k_2,\dots,k_d)))(x_1)\frac{d}{dx_1}(V^1_{\lambda+\sigma(\mathbf{m})}(g(\cdot;m_2,\dots,m_d)))(x_1)\\&\times\mu_2(dx_2)\cdots\mu_d(dx_d)\mu_1(dx_1).\end{aligned}$$

Since $\int\varphi^i_{k_i}(x_i)\varphi^i_{m_i}(x_i)\mu_i(dx_i)=1$ if $k_i=m_i$ and 0 otherwise,

$$\begin{aligned}\|\frac{\partial}{\partial x_1}R_\lambda f\|^2_2&=\sum_{k_2,\dots,k_d=1}^N\int|\frac{d}{dx_1}(V^1_{\lambda+\sigma(\mathbf{k})}(g(\cdot;k_2,\dots,k_d)))(x_1)|^2\mu_1(dx_1)\\&=\sum_{k_2,\dots,k_d=1}^N\|\frac{d}{dx_1}(V^1_{\lambda+\sigma(\mathbf{k})}(g(\cdot,;k_2,\dots,k_d)))\|^2_{L^2(\mu_1)}\\&\le c_{3.6}\sum_{k_2,\dots,k_d=1}^N\|g(\cdot;k_2,\dots,k_d)\|^2_{L^2(\mu_1)},\end{aligned}$$

using Proposition 3.6. But

$$\|g(\cdot;k_2,\dots,k_d)\|^2_{L^2(\mu_1)} = \|\sum_{k_1=1}^N a_{\mathbf{k}}\varphi^1_{k_1}\|^2_{L^2(\mu_1)} = \sum_{k_1=1}^N |a_{\mathbf{k}}|^2.$$

Therefore

$$\|\frac{\partial}{\partial x_1}R_\lambda f\|_2^2 \le c_{3.6}\sum_{k_2,\dots,k_d=1}^N\sum_{k_1=1}^N |a_{\mathbf{k}}|^2 = c_{3.6}\|f\|_2^2,$$

and so (3.10) is proved.

If $f=\sum_{\mathbf{k}} a_{\mathbf{k}}\varphi_{\mathbf{k}}\in\mathcal{D}_0$, then

$$\mathcal{A}_j R_\lambda f = \sum_{\mathbf{k}} a_{\mathbf{k}}\frac{-\lambda^j_{k_j}}{\lambda+\lambda(\mathbf{k})}\varphi_{\mathbf{k}},$$

and so

$$\|\mathcal{A}_j R_\lambda f\|_2^2 = \sum_{\mathbf{k}} a^2_{\mathbf{k}}\Big(\frac{\lambda^j_{k_j}}{\lambda+\lambda(\mathbf{k})}\Big)^2 \le \sum_{\mathbf{k}} a^2_{\mathbf{k}} = \|f\|_2^2. \tag{3.11}$$

Finally, note that for $f\in\mathcal{D}_0$,

$$x_j\frac{\partial^2}{\partial x_j^2}R_\lambda f(x) = \frac{1}{\gamma^0_j}\mathcal{A}_j R_\lambda f(x) - \frac{b^0_j}{\gamma^0_j}\frac{\partial}{\partial x_j}R_\lambda f(x);$$

the proposition therefore follows by the bounds (3.10) and (3.11). □

4. Bessel functions and their zeros

In this section we prove Lemmas 3.2–3.4. Each is standard for a fixed b', but we need estimates that are uniform over $b'\in[\varepsilon,\varepsilon^{-1}]$. We first prove

Lemma 4.1. *Let $J_{b'}$ denote the Bessel function of the first kind with parameter $b'>-1$.*
(a) $J_{b'}(x)\le \frac{(x/2)^{b'}}{\Gamma(b'+1)}\exp\Big(\frac{x^2}{2(b'+1)}\Big)$ for all $x>0$.
(b) For any $\varepsilon>0$ there is a $c_{4.1}(\varepsilon)$ such that for all $-1<b'\le\varepsilon^{-2}$

$$J_{b'}(x) = \sqrt{\frac{2}{\pi x}}\cos(x-b'\pi/2-\pi/4)+E_{b'}(x) \text{ where } |E_{b'}(x)|\le c_{4.1}x^{-3/2} \text{ for all } x\ge 1.$$

Proof. (a) follows from the series expansion of $J_{b'}$ (see (1) on p. 44 of [W44]).
(b) This is a very simple case of the asymptotic expansions on p. 206 of [W44]. We let $(x)_n = x(x+1)\dots(x+n-1)$ and $\{x\}$ be the least integer $k\ge x$. Define

$$a_m(b') = (-1)^{\{m/2\}}\frac{(1/2-b')_m(1/2+b')_m}{m!2^m}.$$

Choose the smallest positive integer p so that $2p > \varepsilon^{-2} - 1/2$. Then (1) on p. 206 of [W44] gives (for $-1 < b' \le \varepsilon^{-2}$)

$$J_{b'}(x) = \sqrt{\frac{2}{\pi x}} \left(\cos(x - \frac{b'\pi}{2} - \frac{\pi}{4})P(x,b') - \sin(x - \frac{b'\pi}{2} - \frac{\pi}{4})Q(x,b') \right),$$

where

$$P(x,b') = \sum_{m=0}^{p-1} a_{2m}(b')x^{-2m} + R_{2p}(b',x),$$

$$Q(x,b') = \sum_{m=0}^{p-1} a_{2m+1}(b')x^{-(2m+1)} + R_{2p+1}(b',x),$$

and

$$|R_q(b',x)| \le |a_q(b')|x^{-q} \text{ for } q = 2p \text{ or } 2p+1.$$

This shows that (b) holds with

$$E_{b'}(x) = \sqrt{\frac{2}{\pi x}} \Big(\cos(x - \frac{b'\pi}{2} - \frac{\pi}{4})(P(x,b') - 1) - \sin(x - \frac{b'\pi}{2} - \frac{\pi}{4})Q(x,b') \Big).$$

The above bounds now give the required bound on $E_{b'}(x)$. □

Proof of Lemma 3.2. Let $\nu = b' \wedge \frac{1}{4}$ and combine the bounds in Lemma 4.1 to see that for $\varepsilon^2 \le b' \le \varepsilon^{-2}$,

$$J_{b'}(x) \le c_0 x^{\nu} 1_{(x \le 1)} + c_1 x^{-1/2} 1_{(x>1)},$$

where c_0 and c_1 are constants depending only on ε. This implies that for $j \le k$,

$$\begin{aligned}
&| \int_0^1 J_{b'}(w_k z) J_{b'}(w_j z) \frac{dz}{z} \, | \\
&\quad \le \int_0^{\frac{1}{w_k}} c_0^2 (w_k w_j)^{\nu} z^{2\nu - 1} dz + \int_{\frac{1}{w_k}}^{\frac{1}{w_j}} c_0 c_1 (w_k z)^{-\frac{1}{2}} (w_j z)^{\nu} \frac{dz}{z} \\
&\qquad + c_1^2 \int_{\frac{1}{w_j}}^{\infty} (w_j w_k)^{-1/2} z^{-2} dz \\
&\quad \le c_0^2 (2\nu)^{-1} (w_j/w_k)^{\nu} + c_0 c_1 (1/2 - \nu)^{-1} (w_j/w_k)^{\nu} + c_1^2 (w_j/w_k)^{1/2} \\
&\quad \le \Big[c_0^2 (2\varepsilon^2 \wedge 1/2)^{-1} + 4 c_0 c_1 + c_1^2 \Big] (w_j/w_k)^{\nu}.
\end{aligned}$$

□

Proof of Lemma 3.3. First consider $k = 1$. Note that

$$f(b', w) \equiv J_{b'-1}(w)/w^{b'-1} = \sum_{m=0}^{\infty} (-1)^m \frac{(w/2)^{2m}}{m!\Gamma(b'+m)} 2^{1-b'}$$

is a jointly continuous function of $(b', w) \in (0, \infty) \times [0, \infty)$. Since

$$\inf_{\varepsilon^2 \le b' \le \varepsilon^{-2}} f(b', 0) = \inf_{\varepsilon^2 \le b' \le \varepsilon^{-2}} 2^{1-b'} \Gamma(b')^{-1} = \eta > 0, \tag{4.1}$$

by uniform continuity there exists $\delta > 0$ depending only on ε such that

$$\inf_{(b',w) \in [\varepsilon^2, \varepsilon^{-2}] \times [0,\delta]} f(b', w) \ge \eta/2.$$

This implies

$$\inf_{\varepsilon^2 \le b' \le \varepsilon^{-2}} w_1(b') > \delta. \tag{4.2}$$

Let $N(R, b')$ be the number of zeros of $J_{b'-1}$ in $[0, R]$. We claim that

$$\sup_{\varepsilon^2 \le b' \le \varepsilon^{-2}} N(R, b') < \infty \text{ for all } R > 0. \tag{4.3}$$

If this is false there is a sequence $\{b'_n\}$ in $[\varepsilon^2, \varepsilon^{-2}]$ which converges to b_0 such that the number of zeros of $f(b'_n, \cdot)$ in $[0, R]$ approaches ∞. Now the series defining $f(b', \cdot)$ shows that we may extend it to an analytic function in the complex plane for each $b' > 0$. Moreover as $n \to \infty$, $f(b'_n, \cdot)$ converges to $f(b_0, \cdot)$ uniformly on compact subsets in the complex plane. Rouché's Theorem now shows that the number of zeros of $f(b'_n, \cdot)$ inside a smooth simple closed curve on which $f(b_0, \cdot)$ does not vanish approaches the number of zeros of $f(b_0, \cdot)$ inside the same curve. This implies that $f(b_0, \cdot)$ has infinitely many zeros in $[0, R]$ which is impossible for a non-constant analytic function.

We claim that there exist $R, c_0 > 0$ depending only on ε such that if $w'_n(b')$ is the nth largest zero of $J_{b'-1}$ in $(R, \infty]$, then

$$w'_n(b') \ge c_0 n \quad \text{for all } b' \in [\varepsilon^2, \varepsilon^{-2}] \text{ and all } n \in \mathbb{N}. \tag{4.4}$$

Let $I_n = [(b'-1)\pi/2 + \pi/2 + n\pi, (b'-1)\pi/2 + \pi + n\pi]$. It follows from Lemma 4.1(b) that there is an $R > 0$ depending only on ε so that all the zeros of $J_{b'-1}$ are included in $\cup_{n=1}^{\infty} I_n \cup [0, R]$. Suppose $J_{b'-1}$ has two zeros in $I_n \cap (R, \infty)$. Since the zeros of $J_{b'-1}$ and $J_{b'}$ are interleaved (see p. 479 of [W44]), this means $J_{b'}$ must have a zero in I_n. Now on I_n,

$$\begin{aligned}
|J_{b'}(x)| &= |\sqrt{\frac{2}{\pi x}} \cos(x - b'\pi/2 - \pi/4) + E_{b'}(x)| \\
&= |\sqrt{\frac{2}{\pi x}} \sin(x - (b'-1)\pi/2 - \pi/4) + E_{b'}(x)| \\
&\ge (\pi x)^{-1/2} - c_{4.1} x^{-3/2} > (\pi x)^{-1/2}/2 > 0,
\end{aligned}$$

where the last line holds if we take R sufficiently large (depending only on ε). This contradiction proves that $J_{b'-1}$ has at most one zero in $I_n \cap (R, \infty)$. (It is in fact easy to use Lemma 4.1(b) to see it has exactly one zero in this interval but we will not need this.) This implies that

$$w'_n(b') \ge (b'-1)\pi/2 + \pi/2 + n\pi,$$

and so (4.4) is proved.

Now use (4.3) to obtain an integer N depending on ε which bounds the number of zeros of J_{b-1} in $[0,R]$ for all $b' \in [\varepsilon^2, \varepsilon^{-2}]$. Therefore (4.4) implies that

$$w_{n+N}(b') \geq w'_n(b') \geq c_0 n \geq \frac{c_0}{N+1}(n+N) \text{ for all } n \in \mathbb{N}. \tag{4.5}$$

For $n \leq N$ note that (4.2) implies that for all $b' \in [\varepsilon^2, \varepsilon^{-2}]$,

$$w_n(b') \geq \delta \geq \frac{\delta}{N} n.$$

This and (4.5) complete the proof. □

Proof of Lemma 3.4. Let $b' \in [\varepsilon^2, \varepsilon^{-2}]$ and assume $w_n(b') \geq R \geq 1$, where R depending only on ε will be chosen below. By Lemma 4.1(b) we have

$$0 = J_{b'-1}(w_n) = \sqrt{\frac{2}{\pi w_n}} \cos(w_n - \frac{(b'-1)\pi}{2} - \frac{\pi}{4}) + E_{b'-1}(w_n)$$

and so, as $w_n \geq 1$, Lemma 4.1(b) implies

$$|\cos(w_n - \frac{(b'-1)\pi}{2} - \frac{\pi}{4})| \leq \sqrt{\frac{\pi}{2}} \frac{c_{4.1}}{w_n} = c_0 w_n^{-1}. \tag{4.6}$$

Use the equality $\cos(x - \pi/2) = \sin x$ to see that for $w_n(b') \geq R \geq 1$ and $b' \in [\varepsilon^2, \varepsilon^{-2}]$,

$$\begin{aligned} |J_{b'}(w_n)| &= \left| \sqrt{\frac{2}{\pi w_n}} \sin\Big(w_n - \frac{(b'-1)\pi}{2} - \frac{\pi}{4}\Big) + E_{b'}(w_n) \right| \\ &\geq \sqrt{\frac{2}{\pi w_n}} \sqrt{1 - c_0^2 w_n^{-2}} - c_{4.1} w_n^{-3/2} \geq \frac{1}{2\sqrt{w_n}}. \end{aligned} \tag{4.7}$$

We have chosen R large enough so that the last inequality holds, and in the next to last inequality we have used (4.6) and Lemma 4.1(b).

Assume next that b' is as above and $w_n(b') < R$. Implicit in the normalization in (3.3) (or see p. 258 of [H]) is the fact that

$$\tfrac{1}{2} J_{b'}(w_n)^2 = \int_0^1 z J_{b'-1}^2(w_n z)\, dz = w_n^{-2} \int_0^{w_n} u J_{b'-1}(u)^2\, du.$$

Therefore

$$\begin{aligned} \tfrac{1}{2} J_{b'}(w_n)^2 &\geq R^{-2} \int_0^{w_1(b')} u J_{b'-1}(u)^2\, du \\ &\geq R^{-2} \int_0^{\delta} u^{2b'-1} \Big(\frac{J_{b'-1}(u)}{u^{b'-1}}\Big)^2\, du. \end{aligned} \tag{4.8}$$

In the last line $\delta = \delta(\varepsilon)$ is as in (4.2). If η is as in (4.1) then (4.1) shows that $(J_{b'-1}(u)/u^{b'-1})^2 \geq \eta^2/4$ for $(b', u) \in [\varepsilon^2, \varepsilon^{-2}] \times [0, \delta]$. So (4.8) now implies

$$|J_{b'}(w_n)| \geq R^{-1}\frac{\eta\delta^{b'}}{2b'} \geq c(\varepsilon) \geq c(\varepsilon)\sqrt{\delta}w_n^{-1/2}$$

(the last by (4.2)). This together with (4.7) gives the required lower bound. □

5. Finiteness of resolvents

In this Section we prove Proposition 2.3. Our strategy is to approximate X_t by processes X_t^n of the form (5.1) below, obtain the analogue of Proposition 2.3 for these processes, and then pass to the limit; cf. [SV79], Theorem 7.1.4. One complication that arises here and that is not present in the work of Stroock and Varadhan is the following. In order to ensure that $|S_\lambda f| < \infty$ for $f \in L^2$ we must start the process with an initial distribution that is absolutely continuous with respect to the measure μ. This requires us to show that the distribution of X_t^n at each of the times $[s]_n$ is also absolutely continuous with respect to μ; this is used in Lemma 5.3.

Let $\mathcal{L}$ be given by (1.2), $M > 0$, and assume X is a solution of the stopped martingale problem for $(\mathcal{L}, [0, M)^d)$ with initial distribution ν on $[0, M)^d$. Recall that we write $\tau = \tau_M$ for the first time $\|X_t\| = M$ and if $f : [0, M]^d \to \mathbb{R}$ we set $f(\partial) = 0$ and let

$$S_\lambda f = E\Big(\int_0^\infty e^{-\lambda t} f(X_t)dt\Big)$$

be the resolvent operator associated with X. Let

$$\|S_\lambda\|_\nu = \sup\{|S_\lambda f| : \|f\|_2 \leq 1\}.$$

We may assume there is a d-dimensional $(\mathcal{F}_t)$-Brownian motion $B_t = (B_t^1, \dots B_t^d)$ such that

$$X_t^i = X_0^i + \int_0^t \sqrt{2\gamma_i(X_s)X_s^i}\, dB_s^i + \int_0^t b_i(X_s)ds \quad \text{for } t < \tau \text{ and } i = 1, \dots, d.$$

Let $\gamma_i(\partial) = \gamma_i^0$, $b_i(\partial) = b_i^0$, and define $X_s = \partial$ for $s < 0$. Set $[s]_n = ([ns] - 1)/n$, and approximate X by the unique solution of

$$X_t^{n,i} = X_0^i + \int_0^t \sqrt{2\gamma_i(X_{[s]_n})X_s^{n,i}}\, dB_s^i + \int_0^t b_i(X_{[s]_n})ds, \ t \geq 0 \text{ and } i = 1, \dots, d. \tag{5.1}$$

Note that for $j \geq 0$ on $[\frac{j}{n}, \frac{j+1}{n}]$ and conditional on $\mathcal{F}_{j/n}$, $X^n = (X^{n,1}, \dots, X^{n,d})$ has generator of the form (3.7) but with γ_i^0 and b_i^0 replaced by the random coefficients $\gamma_i(X_{(j-1)/n})$ and $b_i(X_{(j-1)/n})$ (which will equal these constants for $j = 0$ or j large enough), respectively. With this in mind, one sees that pathwise uniqueness in the above equation for X^n is immediate from the classical result in [YW71]. Note also that, unlike X, $X_t^n \neq \partial$ for all $t \geq 0$.

Lemma 5.1. *For any $T > 0$, $\sup_{t<T\wedge\tau} \|X_t^n - X_t\| \xrightarrow{P} 0$ as $n \to \infty$.*

Proof. Since the b_i and γ_i are bounded on $[0, M]^d$, it is routine to show that the sequence $\{(X^n(\cdot \wedge (\tau-)), X(\cdot \wedge (\tau-)), B) : n \in \mathbb{N}\}$ is tight in $C(\mathbb{R}_+, \mathbb{R}_+^d)^2 \times C(\mathbb{R}_+, \mathbb{R}^d)$. Let (X^∞, X, B) denote any weak limit point. Then standard arguments (e.g. see [Ke84]) show that on an appropriate filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, B is an $\mathcal{F}_t$-Brownian motion and (X^∞, X, B) satisfies for $i = 1, \ldots, d$

$$X^{\infty,i}(t \wedge (\tau-)) = X_0^i + \int_0^{t\wedge(\tau-)} \sqrt{2\gamma_i(X_s)X_s^{\infty,i}}\, dB_s^i + \int_0^{t\wedge(\tau-)} b_i(X_s)ds$$
$$X^i(t \wedge (\tau-)) = X_0^i + \int_0^{t\wedge(\tau-)} \sqrt{2\gamma_i(X_s)X_s^i}\, dB_s^i + \int_0^{t\wedge(\tau-)} b_i(X_s)ds.$$

Using the inequality $\sqrt{x} - \sqrt{y} \le \sqrt{x-y}$ for $0 \le y \le x$, we can bound the local time at zero of $X^{\infty,i} - X^i$ (see [RY91]) by

$$\begin{aligned} L^0_{t\wedge(\tau-)}(X^{\infty,i} - X^i) &= \lim_{\varepsilon\downarrow 0} \varepsilon^{-1} \int_0^{t\wedge\tau} 2\gamma_i(X_s)(\sqrt{X_s^{\infty,i}} - \sqrt{X_s^i})^2 1_{(X_s^{\infty,i} - X_s^i \in (0,\varepsilon))}\, ds \\ &\le 2\|\gamma_i\|_\infty \lim_{\varepsilon\downarrow 0} \varepsilon^{-1} \int_0^{t\wedge\tau} (X_s^{\infty,i} - X_s^i) 1_{(X_s^{\infty,i} - X_s^i \in (0,\varepsilon))}\, ds \\ &\le 2\|\gamma_i\|_\infty \lim_{\varepsilon\downarrow 0} \int_0^{t\wedge\tau} 1_{(X_s^{\infty,i} - X_s^i \in (0,\varepsilon))}\, ds = 0. \end{aligned}$$

By Tanaka's formula this means that $\mathbb{E}(|X^{\infty,i}_{t\wedge(\tau-)} - X^i_{t\wedge(\tau-)}|) = 0$ and so $X^\infty = X$ a.s. It follows that $\sup_{s<T\wedge\tau} \|X_s^n - X_s\| \xrightarrow{w} 0$, and hence also converges to 0 in probability. □

Now let $T_n = \inf\{t : \|X_t^n\| = M\}$, set

$$Y_t^n = \begin{cases} X_t^n & \text{if } t < T_n \\ \partial & \text{if } t \ge T_n, \end{cases}$$

and let $S_\lambda^n f = \mathbb{E}\Big(\int_0^\infty f(Y_t^n)e^{-\lambda t}\, dt\Big)$ be the resolvent associated with Y^n (with the fixed initial law ν), where as usual $f(\partial) = 0$. We will prove Proposition 2.3 by obtaining an upper bound on $\|S_\lambda^n\|_\nu = \sup\{|S_\lambda^n f| : \|f\|_2 \le 1\}$ which is uniform in n.

Let $\mathcal{L}^0$ be as in (3.7) for $b_i^0, \gamma_i^0 > 0$ as in the statement of Proposition 2.3, let X^{γ^0,b^0} be a solution to the stopped martingale problem for $(\mathcal{L}^0, [0, M)^d)$, denote the corresponding resolvent operators by $R_\lambda^{\gamma^0,b^0}$, and recall $\mu(dx) = \prod_{i=1}^d x_i^{b_i^0/\gamma_i^0 - 1} dx_i$. The notation $\|R_\lambda^{\gamma^0,b^0}\|_2$ will refer to the norm as an operator on $L^2([0, M]^d, \mu)$. A trivial eigenfunction expansion shows that

$$\|R_\lambda^{\gamma^0,b^0}\|_2 \le \lambda^{-1} \text{ for all } \gamma^0, b^0 \in (0,\infty)^d \text{ and } \lambda > 0. \tag{5.2}$$

Lemma 5.2. *For any $\delta > 0$ there exists $c_{5.2} = c_{5.2}(\delta)$ such that if $q_t^{\gamma^0,b^0}(x, y)$ is the transition density of X^{γ^0,b^0} with respect to μ and $t \ge \delta$, $\gamma_i^0 \ge \delta$, and $\delta^{-1} \ge b_i^0/\gamma_i^0 \ge \delta$, then*

$$\|q_t^{\gamma^0,b^0}\|_\infty \le c_{5.2}.$$

Proof. It suffices to consider $d = 1$ since $q_t^{\gamma^0,b^0}$ factors into the product of its one-dimensional marginals. It also suffices to show the required bound holds for the corresponding transition density, $p_t^{\gamma^0,b^0}(x,y)$, of the process which is not killed upon exiting $[0,M)$. Dropping the superscript zeros, and using the above notation for these unkilled processes, we note first that $\frac{2}{\gamma}X_t^{\gamma,b}$ is equal in law to $X^{2,2b/\gamma}$ and so is the square of a Bessel process of dimension $2b/\gamma$. If $\alpha = b/\gamma - 1$, and $h_\alpha(z) = \sum_{m=0}^\infty \frac{z^m}{m!\Gamma(m+\alpha+1)}$, then (see p. 411 of [RY91])

$$p_t^{\gamma^0,b^0}(x,y) = (\gamma t)^{-b/\gamma}\exp(-(x+y)/\gamma t)h_\alpha(xy/\gamma^2t^2).$$

The fact that $\sup_{\alpha\geq -1+\delta, z\in[0,K]} h_\alpha(z) < \infty$ for any $\delta, K > 0$ now implies the result. □

Lemma 5.3. *Assume $\nu(dx) = \rho(x)\mu(dx)$ for some $\rho \in L^2(\mu)$. Suppose (2.2) and (2.3) hold. Then $\|S_\lambda^n\|_\nu < \infty$ for each $\lambda > 0$, $n \geq 1$.*

Proof. Set $\gamma^k = \gamma(X_{(k-1)/n})$, $b^k = b(X_{(k-1)/n})$, and $\mu_k(dy) = \prod_{i=1}^d x_i^{(b_i^k/\gamma_i^k)-1}dx_i$ for $k \geq 0$. Let f be a non-negative function in $L^2(\mu)$ and let $\mathbb{E}^x(f(X^{\gamma^k,b^k}))$ denote the (random) expectation with respect to the solution of the stopped martingale problem for $(\mathcal{L}^0, [0,M)^d)$, with initial law δ_x, where $(\gamma^0, b^0) = (\gamma^k, b^k)$ are now random. By the definition of S_λ^n we have,

$$\begin{aligned}
S_\lambda^n f &= \sum_{k\geq 0} e^{-\lambda k/n}\mathbb{E}\Big(\mathbb{E}_{X_{k/n}^n}\Big(\int_0^{1/n} e^{-\lambda t} f(X_t^{\gamma^k,b^k})dt\Big)\Big)\\
&\leq \int R_\lambda^{\gamma^0,b^0} f(x)\rho(x)d\mu(x) + \sum_{k=1}^\infty e^{-\lambda k/n}\mathbb{E}\big(\mathbb{E}\big(R_\lambda^{\gamma^k,b^k} f(X_{k/n}^n)|\mathcal{F}_{(k-1)/n}\big)\big)\\
&\leq \lambda^{-1}\|f\|_2\|\rho\|_2\\
&\qquad + \sum_{k=1}^\infty e^{-\lambda k/n}\mathbb{E}\Big(\int R_\lambda^{\gamma^k,b^k} f(x) q_{1/n}^{\gamma^{k-1},b^{k-1}}(X_{(k-1)/n}^n, x)\mu_{k-1}(dx)\Big),
\end{aligned} \tag{5.3}$$

where (5.2) is used in the last line and $q_{1/n}^{\gamma^{k-1},b^{k-1}}(\partial,\cdot) \equiv 0$ by definition. Now use Lemma 5.2 and (2.2) to bound the (random) transition density in the above summation and conclude that the expectation in the above summation is at most

$$\begin{aligned}
&c_n\mathbb{E}\Big(\int R_\lambda^{\gamma^k,b^k} f(x)\mu_{k-1}(dx)\Big)\\
&\leq c_n\mathbb{E}\Big(\Big(\int R_\lambda^{\gamma^k,b^k} f(x)^2\mu_k(dx)\Big)^{1/2}\Big(\int_{[0,M]^d}\prod_{i=1}^d x_i^{1-(b_i^k/\gamma_i^k)}\prod_{i=1}^d x_i^{2(b_i^{k-1}/\gamma_i^{k-1}-1)}dx\Big)^{1/2}\Big)\\
&\leq c_n\lambda^{-1}\|f\|_2\mathbb{E}\Big(\Big(\int_{[0,M]^d}\prod_{i=1}^d x_i^{2(b_i^{k-1}/\gamma_i^{k-1})-(b_i^k/\gamma_i^k)-1}dx\Big)^{1/2}\Big).
\end{aligned} \tag{5.4}$$

In the first inequality we have used Hölder's inequality and in the second we have used (5.2). Now use (2.2) and (2.3) to see that

$$2\varepsilon^{-2} \geq \frac{2b_i^{k-1}}{\gamma_i^{k-1}} \geq \frac{b_i^k}{\gamma_i^k} + \frac{\varepsilon^2}{2},$$

and hence (5.4) is at most

$$c_n\lambda^{-1}\|f\|_2\Big(\int_{[0,M]^d}(\prod_{i=1}^d x_i^{(\varepsilon^2/2)-1}+\prod_{i=1}^d x_i^{2\varepsilon^{-2}})dx\Big)^{1/2}.$$

We use this bound in (5.3) to conclude that

$$S_\lambda^n f\le\Big[\lambda^{-1}\|\rho\|_2+\lambda^{-1}c(n,\varepsilon,M)\Big]\|f\|_2.$$

The required result follows by applying the above to the positive and negative parts of an arbitrary function in $L^2(\mu)$. ☐

Proof of Proposition 2.3. Using the notation $f_{ii}=\partial^2 f/\partial x_i^2$, define

$$\mathcal{L}^{(y)}f(x)=\sum_{i=1}^d\gamma_i(y)x_if_{ii}(x)+b_i(y)f_i(x),\ \ f\in C^2(\mathbb{R}_+^d),$$

so that in particular $\mathcal{L}^{(\partial)}f=\mathcal{L}^0 f$, and let $\mathcal{B}^{(y)}f(x)=(\mathcal{L}^{(y)}-\mathcal{L}^0)f(x)$. If $f\in C^2([0,M]^d)$, $f|_U=0$, and $f(\partial)=0$, then by Itô's lemma,

$$f(Y_t^n)=f(X_0)+N_{t\wedge\tau}^f+\int_0^t\mathcal{L}^{(X([s]_n))}f(Y_s^n)ds,$$

where N_t^f is a martingale. The above is clear for $t<\tau$ but then follows for all $t\ge0$ because our assumption on f implies that $f(Y_{T_n-}^n)=0$ and so both sides vanish for $t\ge\tau$. Therefore if $\lambda>0$,

$$\begin{aligned}S_\lambda^n f&=\mathbb{E}\Big(\int_0^\infty e^{-\lambda t}\Big[f(X_0)+\int_0^t\mathcal{L}^{(X([s]_n))}f(Y_s^n)ds\Big]dt\Big)\\&=\lambda^{-1}\mathbb{E}(f(X_0))+\lambda^{-1}\mathbb{E}\Big(\int_0^\infty e^{-\lambda s}\mathcal{L}^{(X([s]_n))}f(Y_s^n)ds\Big).\end{aligned}\tag{5.5}$$

Recall the definition of $\mathcal{D}_0$ from (3.9), and write $R_\lambda g$ for $R_\lambda^{\gamma^0,b^0}g$. Let $g\in\mathcal{D}_0$. Since $(\lambda-\mathcal{L}^0)R_\lambda g=g$, we have

$$\mathcal{L}^{(y)}R_\lambda g=\mathcal{B}^{(y)}R_\lambda g+\mathcal{L}^0R_\lambda g=\mathcal{B}^{(y)}R_\lambda g+\lambda R_\lambda g-g.\tag{5.6}$$

Now set $f=R_\lambda g$ in (5.5): using (5.6) we have

$$\begin{aligned}\lambda S_\lambda^n(R_\lambda g)&=\mathbb{E}(R_\lambda g(X_0))+\mathbb{E}\Big(\int_0^\infty e^{-\lambda s}\mathcal{L}^{(X([s]_n))}R_\lambda g(Y_s^n)ds\Big)\\&=\mathbb{E}(R_\lambda g(X_0))+\mathbb{E}\Big(\int_0^\infty\mathcal{B}^{(X([s]_n)}R_\lambda g(Y_s^n)e^{-\lambda s}ds\Big)+\lambda S_\lambda^n(R_\lambda g)-S_\lambda^n g.\end{aligned}$$

Therefore by (2.1) we have

$$\begin{aligned}
|S^n_\lambda g| &= \Big|\mathbb{E}(R_\lambda g(X_0)) + \mathbb{E}\Big(\int_0^\infty \mathcal{B}^{(X([s]_n)} R_\lambda g(Y^n_s) e^{-\lambda s} ds\Big)\Big| \\
&\le |\int R_\lambda g(x)\rho(x)\mu(dx)| + \frac{1}{2K}\mathbb{E}\Big(\int_0^\infty \Big[\sum_{i=1}^d Y^{n,i}_s |(R_\lambda g)_{ii}(Y^n_s)| + |(R_\lambda g)_i(Y^n_s)|\Big] e^{-\lambda s} ds\Big) \\
&\le \|\rho\|_2 \lambda^{-1}\|g\|_2 + (2K)^{-1}\|S^n_\lambda\|_\nu \Big(\sum_1^d \|x_i (R_\lambda g)_{ii}\|_2 + \|(R_\lambda g)_i\|_2\Big),
\end{aligned}$$

where (5.2) is used in the last line. By (2.2) we may apply Proposition 2.2 and conclude that for all $g \in \mathcal{D}_0$,

$$\begin{aligned}
|S^n_\lambda g| &\le \lambda^{-1}\|\rho\|_2\|g\|_2 + (2K)^{-1}\|S^n_\lambda\|_\nu K\|g\|_2 \\
&= \lambda^{-1}\|\rho\|_2\|g\|_2 + \tfrac{1}{2}\|S^n_\lambda\|_\nu \|g\|_2.
\end{aligned}$$

Since S^n_λ is a bounded linear functional on L^2 by Lemma 5.3 and $\mathcal{D}_0$ is dense in L^2, we see that $\|S^n_\lambda\|_\nu \le \frac{1}{2}\|S^n_\lambda\|_\nu + \|\rho\|_2\lambda^{-1}$. Lemma 5.3 implies that $\|S^n_\lambda\|_\nu < \infty$ and so the above implies

$$\|S^n_\lambda\|_\nu \le \frac{2\|\rho\|_2}{\lambda}. \tag{5.7}$$

By Lemma 5.1 and by taking a subsequence, if necessary, we may assume

$$\sup_{t<\tau\wedge n} \|X^n_t - X_t\| \overset{a.s.}{\to} 0.$$

This implies $\liminf T_n \ge \tau$ a.s. and therefore $\lim_{n\to\infty} Y^n_s = X_s$ for all $0 \le s < \tau$. Therefore, Fatou's Lemma and (5.7) show that for $f \in C([0,M]^d)$,

$$S_\lambda(|f|) \le \liminf_{n\to\infty} S^n_\lambda(|f|) \le \frac{2\|\rho\|_2}{\lambda}\|f\|_2.$$

The above inequality now follows easily for all $f \in L^2$. □

6. Continuity of resolvents

In this section we prove Proposition 2.4. The main step in the argument (in Theorem 6.4) is the prove the continuity of harmonic functions for the process X. This is done by adapting an argument of Krylov and Safonov. All the difficulty is at the boundary; if $X_0 \in \partial\mathbb{R}^d_+$ then we have to control the behaviour of X as it leaves the boundary. Using a comparison with Bessel processes (Lemma 6.2) we show X leaves the boundary sufficiently rapidly so that the other components of X do not change much. Once X has left $\partial\mathbb{R}^d_+$ we can use the estimates of Krylov and Safonov (extended in Proposition 6.1 to diffusions with bounded drift) to deduce an oscillation bound, which is then used to imply the continuity of harmonic functions.

The Lebesgue measure of a Borel set G will be denoted by $|G|$. Recall that T_G and τ_G are the first hitting times of G and G^c, respectively. A closed box is a set of the form $[a_1, b_1] \times \cdots \times [a_d, b_d]$. We use $Int(Q)$ to denote the interior of Q.

Proposition 6.1. *Let Q_0 and Q_1 be closed boxes in $\mathbb{R}^d$ with $Q_1 \subset Int(Q_0)$. Let Y satisfy the stochastic differential equation*

$$dY_t^i = \alpha_i(Y_t)dB_t^i + \beta_i(Y_t)dt, \quad i = 1, \dots, d, \tag{6.1}$$

where B is a d-dimensional Brownian motion, α_i and β_i are functions on $\mathbb{R}^d$, the α_i are continuous and the β_i are Borel. Let $\{\mathbb{P}^y, y \in \mathbb{R}^d\}$ be the laws of Y with $Y_0 = y$. Suppose there exists $\lambda > 0$ such that

$$\lambda^{-1} \le \alpha_i(y) \le \lambda, \quad |\beta_i(y)| \le \lambda, \quad y \in \mathbb{R}^d. \tag{6.2}$$

Let $G \subset Q_0$ with $|G| \ge \frac{1}{2}|Q_0|$. Then there exists a constant $p_1 > 0$, depending only on d, Q_0, Q_1 and λ such that

$$\mathbb{P}^y(T_G < \tau_{Q_0}) \ge p_1 \quad \textit{for all } y \in Q_1.$$

Proof. Let $\mathbb{Q}^y$ denote the unique solution to the martingale problem corresponding to the stochastic differential equation

$$dY_t^i = \alpha_i(Y_t)dB_t^i, \quad i = 1, \dots, d.$$

Since the α_i are continuous and uniformly elliptic, both $\mathbb{P}^y$ and $\mathbb{Q}^y$ are uniquely defined by [SV79]. The Girsanov theorem tells us that the Radon-Nikodym density of $\mathbb{Q}^y$ with respect to $\mathbb{P}^y$ on $\mathcal{F}_{t \wedge \tau_{Q_0}}$ equals

$$M_t = \exp\Big(\sum_i (\int_0^t H_s^i dB_s^i - \frac{1}{2}\int_0^t |H_s^i|^2 ds)\Big),$$

where the H^i are adapted processes with $|H_s^i| \le \lambda^2$.

By the theorem of Krylov and Safonov (see, e.g., [B97], Theorem V.7.4) there exists $p_2 > 0$ (depending only on d, λ, Q_i) such that

$$\mathbb{Q}^y(T_G < \tau_{Q_0}) \ge p_2 \quad \text{for all } y \in Q_1.$$

Also, by the Dubins-Schwarz theorem ([B95], Theorem I.5.11) there exists $t > 0$ sufficiently large so that

$$\mathbb{Q}^y(\tau_{Q_0} \ge t) \le \tfrac{1}{2}p_2 \quad \text{for all } y \in Q_1.$$

Therefore if $F = \{T_G < \tau_{Q_0} < t\}$, we have $\mathbb{Q}^y(F) \ge \frac{1}{2}p_2$ for all $y \in Q_1$. So, writing $T = t \wedge \tau_{Q_0}$,

$$\tfrac{1}{2}p_2 \le \mathbb{Q}^y(F) = \mathbb{E}_{\mathbb{P}}^y 1_F M_T \le \mathbb{P}^y(F)^{1/2}(\mathbb{E}_{\mathbb{P}}^y M_T^2)^{1/2} \quad \text{for all } y \in Q_1.$$

Since $\mathbb{E}_{\mathbb{P}}^y M_T^2 \le \exp(2d\lambda^4 t)$, the result now follows. □

Remarks. 1. Since this result only concerns the behaviour of the process Y up to its first exit from Q_0, we only need assume that Y satisfies (6.1) for $0 \le t \le \tau_{Q_0}(Y)$.
2. As the proof of Proposition 6.1 is invariant under translations in $\mathbb{R}^d$, the constant $p_1 = p_1(d, \lambda, Q_0, Q_1)$ can be chosen so that it is not affected by a translation of the boxes Q_i.

Lemma 6.2. *Let H and J be predictable processes satisfying, for some constant $1 \le \kappa < \infty$*

$$\kappa^{-1} \le H_s \le \kappa, \quad \kappa^{-1} \le J_s \le \kappa, \quad s \ge 0.$$

Let B be a Brownian motion and let V satisfy, for some $v_0 \ge 0$,

$$V_t = v_0 + \int_0^t H_s \sqrt{2V_s} dB_s + \int_0^t J_s ds.$$

Let $t_0 > 0$, and S be a random variable uniformly distributed on $[\frac{1}{2}t_0, t_0]$, independent of B, H and J. Then for each $\varepsilon > 0$ there exists a constant $\delta > 0$, depending only on t_0, κ and ε, such that

$$\mathbb{P}(V_S \le \delta) \le \varepsilon. \tag{6.3}$$

Proof. Let $A_t = \int_0^t H_s^2 ds$ and σ_t be the inverse of A. For any process X we write $\widetilde{X}_t = X_{\sigma_t}$. Note that we have the bounds

$$\kappa^{-2} t \le A_t \le \kappa^2 t, \quad \kappa^{-2} t \le \sigma_t \le \kappa^2 t.$$

The process $\widetilde{V}$ satisfies, for a Brownian motion B',

$$\widetilde{V}_t = v_0 + \int_0^t \sqrt{2\widetilde{V}_s} dB'_s + \int_0^t \widetilde{J}_s \widetilde{H}_s^{-2} ds.$$

Let $\widetilde{U}$ be defined by

$$\widetilde{U}_t = \int_0^t \sqrt{2\widetilde{U}_s} dB'_s + \int_0^t \kappa^{-3} ds;$$

then as $\kappa^{-3} \le \widetilde{J}_s \widetilde{H}_s^{-2}$ by a comparison theorem (see Theorem V.43.1 in [RW87]) we have $\widetilde{U}_t \le \widetilde{V}_t$ for all $t \ge 0$. Set $U_t = \widetilde{U}_{A_t}$. Write $p(s) = (2/t_0) 1_{(t_0/2, t_0)}(s)$ for the density of S. Since S is independent of $\widetilde{U}$ and A we have, for any $\lambda > 0$

$$\begin{aligned}
\mathbb{P}(V_S \le \lambda) \le \mathbb{P}(U_S \le \lambda) = \mathbb{P}(\widetilde{U}_{A_S} \le \lambda) \\
&= \mathbb{E} \int_0^\infty 1_{[0,\lambda]}(\widetilde{U}_{A_s}) p(s) ds \\
&= \mathbb{E} \int_0^\infty 1_{[0,\lambda]}(\widetilde{U}_t) p(\sigma_t) \widetilde{H}_t^{-2} dt \\
&\le \kappa^2 \mathbb{E} \int_{t_0/2\kappa^2}^{t_0 \kappa^2} 1_{[0,\lambda]}(\widetilde{U}_t)(2/t_0) dt.
\end{aligned}$$

The process $\widetilde{U}$ starts at 0 and has transition density $q_t^{\gamma,b}$ as in Lemma 5.2 with respect to μ, with $d = 1$, $\gamma = 1$ and $b = \kappa^{-3}$. Therefore

$$\mathbb{P}(V_S \le \lambda) \le (2\kappa^2/t_0) \int_0^\lambda dx \int_{t_0/2\kappa^2}^{t_0\kappa^2} q_s^{1,\kappa^{-3}}(0,x) x^{\kappa^{-3}-1} ds. \tag{6.4}$$

The integral (6.4) converges to 0 as $\lambda \to 0$ by Lemma 5.2, and so, taking δ small enough we have proved (6.3). $\square$

Remark 6.3. Write $T_M = \inf\{t : V_t > M\}$, and $T'_M = \inf\{t : \widetilde{U}_t > M\}$. Then we have, using the comparison between U and V above, that $T_M \le h^2 T'_M$. Since a squared Bessel process with a positive parameter has integrable hitting times, we obtain

$$\mathbb{E}T_M \le h^2 \mathbb{E}T'_M \le h^2 c(h^{-3}, M) < \infty.$$

Definition. Assume D_0 is a domain in $\mathbb{R}^d_+$ and $((\mathbb{P}^x)_{x \in D_0 \cup \{\partial\}}, X_t)$ is a Borel strong Markov process taking values in $D_0 \cup \{\partial\}$. If D is open in D_0, a Borel function $h : \overline{D} \to \mathbb{R}$ is X-harmonic in D if $h(X(t \wedge \tau_D))$ is a $\mathbb{P}^x$-martingale for every $x \in D_0$.

Theorem 6.4. *Suppose that $M \in (0, \infty]$, and $((\mathbb{P}^x, x \in [0, M)^d), X_t)$ satisfies the hypotheses of Proposition 2.4. Let D be open in $[0, M)^d$, and h be a bounded X-harmonic function in D. Then h is continuous on D.*

Proof. It is enough to prove that h is continuous at each $z \in D$. If $z \in (0, M)^d$, then by changing the diffusion coefficients outside of a small ball B, centered at z, we may assume that h is X'-harmonic on B where X' is a diffusion with bounded, continuous, and uniformly elliptic diffusion coefficients. It is then well-known that h is continuous on B – see [B97], Theorem V.7.5.

Now let $z \in \partial\mathbb{R}^d_+ \cap D$; by permuting the axes if necessary we can assume that $z = (z_1, \dots, z_k, 0, \dots, 0)$ where $0 \le k < d$ and $z_i > 0$ for $i \le k$. (If $z = 0$ then $k = 0$: this is covered in the calculations below, but some of the estimates required for the general mixed case are not needed.) Choose $q \in (0, 1)$, and for $\eta \ge 0$, set

$$R_n(z, \eta) = \prod_{i=1}^{k} [z_i - q^n, z_i + q^n] \times [\eta, \eta + q^{2n}]^{d-k}.$$

Choose $N = N(z)$ so that $R_N(z, 0) \subset D$ and $2q^N \le \min(z_1, \dots z_k)$. The hypotheses on b_i and γ_i imply that there exists a constant $\varepsilon_1 = \varepsilon(z) > 0$, (depending on z) such that

$$\varepsilon_1 \le b_i(x) \le \varepsilon_1^{-1}, \quad \varepsilon_1 \le \gamma_i(x) \le \varepsilon_1^{-1}, \quad x \in R_N(z, 0), \quad 1 \le i \le d.$$

If $D' \subset D$ let $\mathrm{Osc}\,(D', h) = \sup\{|h(x) - h(y)| : x, y \in D'\}$. To prove that h is continuous at z it is enough to prove that there exists $\rho = \rho(z) < 1$ such that

$$\mathrm{Osc}\,(R_{n+2}(z, 0), h) \le \rho\, \mathrm{Osc}\,(R_n(z, 0), h), \quad n \ge N(z). \tag{6.5}$$

Fix $n \ge N(z)$. By looking at $c_1 h + c_2$ for suitable c_1 and c_2, we may assume $\sup_{R_n} h = 1$ and $\inf_{R_n} h = 0$. Define $\psi : \mathbb{R}^d_+ \to \mathbb{R}^d_+$ by $\psi(x) = (q^{-n} x_1, \dots, q^{-n} x_k, q^{-2n} x_{k+1}, \dots, q^{-2n} x_d)$. Note that if $w_n = \psi(z)$ then $\psi(R_{n+m}(z, 0)) = R_m(w_n, 0)$. For any function $g : \mathbb{R}^d_+ \to \mathbb{R}$ write $\widehat{g}(y) = g(\psi^{-1}(y))$. Let $Y_t = \psi(X_{q^{2n} t})$, $\widehat{\mathbb{P}}^y(\cdot) = \mathbb{P}^{\psi^{-1}(y)}(Y \in \cdot)$, and $\widehat{\mathbb{P}}^\partial = \mathbb{P}^\partial$; then $((\widehat{\mathbb{P}}^y, y \in \psi([0, M)^d) \cup \{\partial\}), Y)$ is a Borel strong Markov process, and $\widehat{h}$ is Y-harmonic in $R_0(w_n, 0)$.

We use $\tau_\cdot$, $T_\cdot$ for the exit and hitting times of the process Y. To prove (6.5) we only need consider Y up to the time $\sigma = \tau_{R_0(w_n,0)}$. The process Y satisfies

$$dY_t^i = q^{n/2}\sqrt{2Y_t^i\widehat{\gamma}_i(Y_t)}d\widehat{B}_t^i + q^n\widehat{b}_i(Y_t)dt, \quad i \le k, \quad 0 \le t < \sigma \tag{6.6}$$

$$dY_t^i = \sqrt{2Y_t^i\widehat{\gamma}_i(Y_t)}d\widehat{B}_t^i + \widehat{b}_i(Y_t)dt, \quad k+1 \le i \le d, \quad 0 \le t < \sigma \tag{6.7}$$

Here $\widehat{B}$ is a d-dimensional Brownian motion. For $i \le k$, set $\alpha_i(y) = q^{n/2}(2y_i\widehat{\gamma}_i(y))^{1/2}$. Then since for $x \in R_n(z,0)$,

$$\tfrac{1}{2}z_i \le z_i - q^n \le x_i \le z_i + q^n \le 2z_i,$$

we have $\frac{1}{2}q^{-n}z_i \le y_i \le 2q^{-n}z_i$ for $y \in R_0(w_n,0)$. Thus

$$(2\varepsilon_1 z_i)^{1/2} \le \alpha_i(y) \le (2\varepsilon_1^{-1}z_i)^{1/2}, \quad y \in R_0(w_n,0), \quad i \le k. \tag{6.8}$$

The k-dimensional process on $R_0(w_n,0)$ defined by (6.6) therefore has uniformly elliptic diffusion coefficients, but with the ellipticity bounds depending on z. To prove (6.5) it is sufficient to prove

$$\operatorname{Osc}(R_2(w_n,0),\widehat{h}) \le \rho \operatorname{Osc}(R_0(w_n,0),\widehat{h}), \tag{6.9}$$

for a constant $\rho = \rho(z) < 1$. We do this by, first, finding $\eta > 0$ such that Y hits $R_1(w_n,2\eta)$ with high probability, and then using Proposition 6.1 to handle the behaviour of Y in the cubes $R_1(w_n,2\eta) \subset R_0(w_n,\eta)$.

Standard estimates for semimartingales imply that there exists $t_0 > 0$, depending only on z and $\varepsilon_1(z)$, such that

$$\widehat{\mathbb{P}}^y\big(\max_{1\le i\le k}\sup_{0\le s\le t_0}|Y_{s\wedge\sigma}^i - Y_0^i| > \tfrac{1}{2}(q-q^2)\big) \le \tfrac{1}{4}, \quad y \in R_2(w_n,0). \tag{6.10}$$

This controls the oscillation of the process $(Y^1,\dots,Y^k)$. We now look at $(Y^{k+1},\dots,Y^d)$. Define processes $\overline{Y}_t^i$ for $k+1 \le i \le d$ by setting $\overline{Y}_t^i = Y_t^i$ for $t \le \sigma$, and

$$d\overline{Y}_t^i = \sqrt{2\overline{Y}_t^i\widehat{\gamma}_i(Y_\sigma)}d\widehat{B}_t^i + \widehat{b}_i(Y_\sigma)dt, \quad t \ge \sigma.$$

So $\overline{Y}^i$ satisfies the hypotheses of Lemma 6.2 with $\kappa = \varepsilon_1(z)^{-1}$, and therefore there exists $1/2 > \eta = \eta(z) > 0$ small enough so that if S is independent of $\overline{Y}$ and is uniformly distributed on $[t_0/2, t_0]$ then

$$\widehat{\mathbb{P}}^y(\min_{k<i\le d}\overline{Y}_S^i > 2\eta) \ge \tfrac{3}{4}, \quad y \in R_2(w_n,0). \tag{6.11}$$

Write $\tau_j = \tau_{R_j(w_n,0)}(Y)$, $j = 0,1$, and $T_1 = T_{R_1(w_n,2\eta)}(Y)$. From (6.10) and (6.11) we obtain

$$\widehat{\mathbb{P}}^y(T_1 < \tau_1) \ge \tfrac{1}{2}, \quad y \in R_2(w_n,0). \tag{6.12}$$

Now let $Q_0^n = R_0(w_n,\eta)$, $Q_1^n = R_1(w_n,2\eta)$, and $G = \{y \in Q_0^n : \widehat{h}(y) \ge \frac{1}{2}\}$; replacing h by $1-h$ if necessary we can assume $|G| \ge \frac{1}{2}|Q_0^n|$. Use (6.8) for $i \le k$, and for $i > k$ note that on Q_0^n,

$$\sqrt{y_i\widehat{\gamma}_i(y)} \ge \sqrt{2\eta\widehat{\gamma}_i(y)} \ge c\sqrt{\eta}$$

to see that the coefficients of (6.6), (6.7), restricted to Q_0^n, satisfy the hypotheses (6.2) of Proposition 6.1 with a λ which may depend on z but is independent of n.

Using Proposition 6.1 and the remarks following we deduce that there exists a constant $p_3 > 0$, depending on z but not n, such that

$$\widehat{\mathbb{P}}^y(T_G < \tau_{R_0(w_n,\eta)}) > p_3, \quad y \in R_1(w_n, 2\eta). \tag{6.13}$$

So for $y \in R_2(w_n, 0)$, (6.12), (6.13) and the strong Markov property imply

$$\widehat{\mathbb{P}}^y(T_G < \tau_0) \geq \widehat{\mathbb{E}}^y 1_{(T_1<\tau_1)} \widehat{\mathbb{P}}^{Y_{T_1}}(T_G < \tau_0) \geq \tfrac{1}{2} p_3.$$

Thus if $y \in R_2(w_n, 0)$, the Y-harmonicity of $\widehat{h}$ on $R_0(w_n, 0)$ gives

$$\widehat{h}(y) = \widehat{\mathbb{E}}^y \widehat{h}(Y_{T_G \wedge \tau_0}) \geq \widehat{\mathbb{E}}^y 1_{(T_G<\tau_0)} \widehat{h}(Y_{T_G}) \geq \tfrac{1}{4} p_3.$$

So, taking $\rho = 1 - \frac{1}{4}p_3$ and recalling that $\operatorname{Osc}(R_0(w_n, 0), \widehat{h}) = 1$, we have proved (6.9). □

Proof of Proposition 2.4. Fix $\lambda > 0$, and a bounded Borel measurable function f: we can assume $||f||_\infty = 1$. Let $B(x, \delta)$ denote the set of points in $\mathbb{R}_+^d$ within a distance δ of $x \in \mathbb{R}_+^d$. Fix $x \in [0, M)^d$ and $\varepsilon > 0$, and choose $\delta > 0$ so that $B = B(x, \delta) \subset [0, M)^d$. We claim that we can choose $\delta > 0$ sufficiently small so that

$$\sup_{y \in B} \mathbb{E}^y \tau_B \leq \varepsilon. \tag{6.14}$$

If $x \in (0, M)^d$, then we can take δ so that $B(x, \delta)$ is bounded away from $\partial\mathbb{R}_+^d$ and so the diffusion coefficients are uniformly elliptic on $B(x, \delta)$. A simple application of the Dubins-Schwarz theorem (see Theorem I.5.11 of [B95]) now gives the required δ. If $x \in \partial\mathbb{R}_+^d$, then we can argue as in Remark 6.3 to bound the left-hand side of (6.14) by the mean hitting time of δ by a squared Bessel process with some positive parameter starting at 0. This can be made arbitrarily small by making δ small by dominated convergence and so (6.14) is proved in either case.

Let $D = B(x, \delta)$. For $x \in D$ set

$$h_D(x) = \mathbb{E}^x S_\lambda f(X_{\tau_D}).$$

Note that as $S_\lambda f(x) \leq \lambda^{-1}$, h_D is bounded, and since it is X-harmonic in D, h_D is continuous in D by Theorem 6.4. If $y \in D$ we have

$$S_\lambda f(y) = \mathbb{E}^y \int_0^{\tau_D} e^{-\lambda s} f(X_s) ds + \mathbb{E}^y e^{-\lambda \tau_D} S_\lambda f(X_{\tau_D}).$$

Therefore

$$|S_\lambda f(y) - h_D(y)| \leq \mathbb{E}^y \tau_D + \lambda^{-1} \mathbb{E}^y (1 - e^{-\lambda \tau_D}) \leq 2\mathbb{E}^y \tau_D \leq 2\varepsilon, \tag{6.15}$$

where (6.14) is used in the last line. By the continuity of h_D we can choose $0 < \delta' < \delta$ so that $y \in B(x, \delta')$ implies that $|h_D(y) - h_D(x)| < \varepsilon$. This together with (6.15) shows that $|S_\lambda f(x) - S_\lambda f(y)| < 5\varepsilon$ if $y \in B(x, \delta')$, and hence $S_\lambda f$ is continuous at x. □

7. Proof of main theorem

Proof of Proposition 2.1. Let $\mathbb{P}_i$, $i = 1,2$ be distinct solutions to $MP(\nu, \mathcal{L})$. A standard argument (see p. 136 of [B97]) shows that for ν-a.a. x, the regular conditional probability $\mathbb{P}_i(\ \cdot\ |X_0 = x)$ solves $MP(\delta_x, \mathcal{L})$ and so it is enough to consider $\nu = \delta_x$. The construction in Theorem 12.2.4 of [SV79] (Krylov's Markov selection theorem) now gives a pair of Borel strong Markov processes $(\mathbb{Q}_i^x, X_t)$, $i = 1, 2$, so that $\mathbb{Q}_i^x$ solves $MP(\delta_x, \mathcal{L})$ for each $x \in \mathbb{R}_+^d$, and $\mathbb{Q}_1^{x_0} \neq \mathbb{Q}_2^{x_0}$ for some x_0. Here we are applying this result in the positive orthant rather than $\mathbb{R}^d$ but one can extend the coefficients to all of $\mathbb{R}^d$ by replacing x_i with x_i^+ and note that solutions starting in the orthant must remain there by a comparison argument as in the proof of Lemma 6.2. Recall that τ_M is the exit time from $[0, M)^d$ and let

$$Y_t = \begin{cases} X_t, & \text{if } t < \tau_M \\ \partial & \text{if } t \geq \tau_M. \end{cases} \tag{7.1}$$

Then Y is a Borel measurable function of X because τ_M is the hitting time of a closed set by a continuous path X (it is the increasing limit of a sequence of hitting times of open sets). If $\mathbb{P}_i^x(\cdot) = \mathbb{Q}_i^x(Y \in \cdot)$, and $\mathbb{P}_i^\partial$ is point mass at the trivial path, then $(\mathbb{P}_i^x, X_t)$ is a Borel strong Markov process for $i = 1, 2$. This follows as in Section III.3 of [BG68], but universal completions can be avoided by the Borel measurability noted above. It is also clear that $\mathbb{P}_i^x$ solves $SMP(\delta_x, \mathcal{L}, [0, M)^d)$ for each $x \in [0, M)^d$ and that we may take M sufficiently large to ensure $\mathbb{P}_1^{x_0} \neq \mathbb{P}_2^{x_0}$. To obtain a contradiction and hence complete the proof we now show that these Borel strong Markov solutions to the stopped martingale problem must coincide.

First consider an initial law $\nu(dx) = \rho(x)d\mu(x)$ for some $\rho \in L^2(\mu)$. Extend any function f on $[0, M]^d$ to $[0, M]^d \cup \partial$ by setting $f(\partial) = 0$. Recall that C_0^2 is the set of functions in $C^2([0, M]^d)$ such that $f(x) = 0$ if $x \in U_M$. Let $\mathcal{D}_0$ be defined as in (3.9). For $f \in C_0^2$ we have by Itô's formula

$$f(X_t) = f(X_0) + \int_0^t \mathcal{L}f(X_s)ds + N^f_{t\wedge(\tau_M-)}, \tag{7.2}$$

where the last term is an $\mathcal{F}_t^X$-martingale under each $\mathbb{P}_i^x$. Note that as $f(X_{\tau_M-}) = 0$ both sides of (7.2) are zero for $t \geq \tau_M$. Let

$$\overline{S}_\lambda^k f = \int S_\lambda^k f(x)\,\nu(dx) = \int \mathbb{E}_k^x\Big(\int_0^\infty e^{-\lambda t} f(X_t)dt\Big)\nu(dx).$$

Taking expectations and integrating (7.2) we have,

$$\lambda \overline{S}_\lambda^k f = \int f d\nu + \overline{S}_\lambda^k \mathcal{L} f = \int f d\nu + \overline{S}_\lambda^k (\mathcal{L} - \mathcal{L}^0) f + \overline{S}_\lambda^k \mathcal{L}^0 f. \tag{7.3}$$

Let $g \in \mathcal{D}_0$, and set $f = R_\lambda g$. Then $f \in \mathcal{D}_0$ by Proposition 3.1, and rearranging (7.3) by using the fact that $(\lambda - \mathcal{L}^0)R_\lambda g = g$, we get

$$\overline{S}_\lambda^k g = \int R_\lambda g(x) d\nu(x) + \overline{S}_\lambda^k (\mathcal{L} - \mathcal{L}^0) R_\lambda g.$$

Hence, writing $Wg(x) = (\mathcal{L} - \mathcal{L}^0)R_\lambda g(x)$, we have

$$(\overline{S}_\lambda^1 - \overline{S}_\lambda^2)g = (\overline{S}_\lambda^1 - \overline{S}_\lambda^2)Wg \text{ for all } g \in \mathcal{D}_0. \tag{7.4}$$

Apply Proposition 2.2, and (2.1) to see that for $g \in \mathcal{D}_0$,

$$|(\overline{S}_\lambda^1 - \overline{S}_\lambda^2)g| \le ||\overline{S}_\lambda^1 - \overline{S}_\lambda^2||_\nu \|Wg\|_2 \le \tfrac{1}{2}\|\overline{S}_\lambda^1 - \overline{S}_\lambda^2\|_\nu \|g\|_2.$$

Since the $\overline{S}_\lambda^k$ are bounded linear functionals on L^2 by Proposition 2.3 and $\mathcal{D}_0$ is dense in L^2 by (2.5), this implies $\|\overline{S}_\lambda^1 - \overline{S}_\lambda^2\|_\nu \le \frac{1}{2}\|\overline{S}_\lambda^1 - \overline{S}_\lambda^2\|_\nu$. By Proposition 2.3 again we conclude that

$$\int S_\lambda^1 f(x)\rho(x)d\mu(x) = \int S_\lambda^2 f(x)\rho(x)d\mu(x) \text{ for all } f, \rho \in L^2. \tag{7.5}$$

We therefore conclude that $S_\lambda^1 f(x) = S_\lambda^2 f(x)$ for almost every x (with respect to μ). But by Proposition 2.4 $S_\lambda^k f$ is continuous, $k = 1, 2$, hence we have equality for all $x \in [0, M)^d$. With this fact, we now appeal to Corollary 6.2.4 and Lemma 6.5.1 of [SV79] or Theorem VI.3.2 of [B97] to conclude that $\mathbb{P}_1^x = \mathbb{P}_2^x$ for all x, and we are done. □

Proof of Theorem 1.1. We first prove existence in the case when γ_i, $i = 1, \ldots, d$, are bounded. This is a well-known result of Skorokhod (see Theorem 6.1.6 of [SV79]) when the state space is all of $\mathbb{R}^d$ (rather than the non-negative orthant) and b_i is bounded. The latter condition is easily weakened to our linear growth condition (1.6). To apply the above result to the orthant, extend $x_i\gamma_i(x)$ and $b_i(x)$ to all of $\mathbb{R}^d$ by replacing x_j by x_j^+ for $j \le d$. In particular, $x_i\gamma_i(x) = 0$ if $x_i < 0$ and $b_i(x) > 0$ for x outside the positive orthant, the latter by (1.5). Since these extended functions are continuous, we have a solution $\mathbb{P}$ to the martingale problem for $\mathcal{L}$ on $\mathbb{R}^d$ by the above. As in the proof of Lemma 6.2 we can apply a comparison argument to see that each coordinate of X is locally bounded below by the square of a Bessel process. This implies that if X starts in $\mathbb{R}_+^d$, it will remain there and so $\mathbb{P}$ is a solution to the martingale problem for $\mathcal{L}$ on $\mathbb{R}_+^d$.

To remove the boundedness condition on the γ_i, let $X^M = (X^{M,i})$ be a solution of the above martingale problem with γ_i^M in place of γ_i, where these functions agree on $[0, M]^d$ and γ_i^M is bounded. We will assume X^M satisfies the associated stochastic differential equation driven by Brownian motions B^i. Set

$$Y^{M,i}(t) = X_0^i + \int_0^t |b_i(X_s^M)|ds + \int_0^t \sqrt{2X_s^{M,i}\gamma_i^M(X_s^M)}dB^i(s).$$

Then $Y^{M,i}$ is a non-negative submartingale dominating $X^{M,i}$ and a standard argument using the linear growth of b_i shows that first $\mathbb{E}(\sum_i |X_t^{M,i}|)$ and then $\mathbb{E}(\sum_i Y_T^{M,i})$ is bounded above uniformly in M. An application of the weak maximal inequality to $\sum_i Y^{M,i}$ now shows that $\sup_{t\le T}\sum_i X_t^{M,i}$ is bounded in probability uniformly in M for each $T > 0$. It is now standard to establish tightness of $\{X^M, M \in \mathbb{N}\}$ and show that any weak limit point satisfies the martingale problem for $\mathcal{L}$ with initial law ν.

We now turn to uniqueness. As in the proof of Proposition 2.1 we may assume $\nu = \delta_{x_0}$. Since the γ_i and b_i are continuous, the γ_i are strictly positive, and the b_i are strictly positive on $\partial\mathbb{R}^d_+$, for some $\varepsilon > 0$, every point $y \in [0, M]^d$ has a neighborhood $V_y = B(y, \eta(y)) \cap \mathbb{R}^d_+$ in $\mathbb{R}^d_+$ such that either:

(a) $\bar{V}_y \cap \partial\mathbb{R}^d_+ = \emptyset$,

or

(b) (i) $\varepsilon \le b_i(x), \gamma_i(x) \le \varepsilon^{-1}$ for $x \in V_y$, $i = 1, \ldots, d$;

(ii) $|b_i(x) - b_i(y)|, |\gamma_i(x) - \gamma_i(y)| \le (2K(\varepsilon, d))^{-1}$ for $x \in V_y$, $i = 1, \ldots, d$, where $K(\varepsilon, d)$ is given by Proposition 2.1;

(iii) $2b_i(x)/\gamma_i(x) \ge b_i(z)/\gamma_i(z) + (\varepsilon^2/2)$ for $x, z \in V_y$, $i = 1, \ldots, d$.

In case (a) an appropriate truncation will allow us to define bounded continuous coefficients $a_{ii}(x)$ and $\tilde{b}_i(x)$ on all of $\mathbb{R}^d$ which agree with $x_i\gamma_i(x)$ and $b_i(x)$, respectively, on V_y, and for which the matrix a is uniformly continuous and positive definite. Existence and uniqueness in law of solutions to the martingale problem for $\tilde{\mathcal{L}}f = \sum_i a_{ii}f_{ii} + \tilde{b}_i f_i$ follows from the classical theorem of Stroock and Varadhan [SV79].

Assume now that case (b) holds. It is easy to use the values of these coefficients on ∂V_y to define coefficients $\tilde{\gamma}_i$, $\tilde{b}_i$ on all of $\mathbb{R}^d_+$ which agree with γ_i and b_i, respectively, on V_y, and which satisfy the hypotheses of Theorem 1.1 and (b) on all of the non-negative orthant, not just V_y. Now apply Proposition 2.1 with $b^0_i = b_i(y)$ and $\gamma^0_i = \gamma_i(y)$ to see that there is uniqueness in law for solutions to the martingale problem for $\tilde{\mathcal{L}}f = \sum_i x_i\tilde{\gamma}_i(x)f_{ii}(x) + \tilde{b}_i(x)f_i(x)$. Existence of solutions was already established above.

We therefore have shown that in a neighborhood of each point we can find diffusion coefficients which agree with our given coefficients and for which the martingale problem is well-posed. We now apply Stroock-Varadhan's localization argument (Theorem 6.6.1 of [SV79]–see also Theorem VI.3.4 in [B97]), trivially modified to our positive orthant setting, to see that solutions to the martingale problem for $\mathcal{L}$ are unique. (Note the measurability required in Theorem 6.6.1 follows from the uniqueness of the martingale problem as in Ex. 6.7.4 of [SV79]).

The Borel and strong Markov properties now follow from the uniqueness and existence established above by well-known arguments (see Theorem 6.2.2 of [SV79] and the ensuing comments). The claimed continuity of the resolvent operators associated with this Markov process follows from Proposition 2.4 with $M = \infty$. $\square$

Proof of Corollary 1.2. The existence of a solution follows by a minor modification of the proof of Theorem 1.1 (existence only requires (1.1), (1.6) and (1.7)). To prove the uniqueness assertion in (a), first assume (1.7) with $C = 0$. For $M > 0$ one can suitably change γ_i, b_i outside the set $\{x \in \mathbb{R}^d_+ : \frac{1}{M} \le \|x\|\}$ so that (1.5) and (1.6) hold. If

$$T(M) = \inf\{t \ge 0 : \|X_t\| \le 1/M\},$$

then $T(M) \uparrow T_0 \le \infty$ a.s. Apply Theorem 1.1 to the martingale problem for these modified coefficients to see that $\mathbb{P}(X(\cdot \wedge T(M)) \in \cdot)$ is uniquely determined, and hence so is $\mathbb{P}(X(\cdot \wedge T_0) \in \cdot)$. Turning now to the general case under (1.7), note first that by considering the solutions up to the first time they exit from $[0, M]^d$ we can assume without loss of generality that b_i and γ_i are bounded and γ_i is bounded away from zero for all i. With Girsanov's theorem in mind, set $\hat{\mathcal{L}} = \mathcal{L} + \sum_i Cx_i \frac{\partial}{\partial x_i}$, where C is as in (1.7). Then $\hat{b}_i(x) = b_i(x) + Cx_i > 0$ for all $x \in \partial\mathbb{R}^d_+/\{0\}$. If $\hat{\mathbb{P}}$ is any solution of the martingale problem for $\hat{\mathcal{L}}$, then the $C = 0$ case proved above shows uniqueness in law of $\hat{\mathbb{P}}(X(\cdot \wedge T_0) \in \cdot)$ and so the same is true for any solution $\mathbb{P}$ of the martingale problem for $\mathcal{L}$ by Girsanov's theorem. Here we have used the fact that γ_i is bounded away from zero and b_i and γ_i are all bounded. This establishes (a).

For (b), note that $\sum_i X^i_t$ is a non-negative supermartingale by (1.8) and so must be identically 0 after T_0. Therefore X is a.s. equal to a fixed Borel function of $X(\cdot \wedge T_0)$ and so uniqueness of the solution of the martingale problem for $\mathcal{L}$ follows from (a). ☐

8. A Counter-example

The following example shows that, even if $d = 1$, we do not have uniqueness in Theorem 1.1 if we weaken (1.5) and only assume $b \geq 0$.

Proposition 8.1. *Let $b(x) = \left(c/\log^+ \frac{1}{x}\right) \wedge 1$ for $x > 0$, and let $b(0) = b(0+) = 0$. If $c > 1$ then the stochastic differential equation*

$$dX_t = (2X_t)^{1/2} dB_t + b(X_t)dt, \quad X_0 = 0, \tag{8.1}$$

has a solution $X \geq 0$ which is not identically 0. Since 0 is also a solution, uniqueness in law fails for solutions of (8.1).

Remark 8.2. X will solve (8.1) if and only if X is a solution of the martingale problem for $\mathcal{L} = x\frac{d^2}{dx^2} + b(x)\frac{d}{dx}$. This Proposition is an exercise in the classification of boundary points for one-dimensional diffusions (see, e.g., Section 16.7 of [B68]). 0 is a regular boundary point for the generator $\mathcal{L}$ on $(0, \infty)$ if and only if $c > 1$. It follows that there are non-trivial solutions to (8.1) if and only if $c > 1$. We give a direct construction of a non-trivial solution for the sake of completeness.

Proof. Let $x_0 = e^{-c}$ and $u(x) = \int_0^x \exp\Big\{-\int_{x_0}^y (b(z)/z)dz\Big\}dy$. Then $u : [0, \infty) \to [0, \infty)$ is a strictly increasing continuous function whose range is all of $[0, \infty)$. Let $s(y)$ denote the inverse function to u, and set

$$\sigma^2(y) = 2s(y) \exp\Big\{-\int_{x_0}^{s(y)} \frac{2b(z)}{z} dz\Big\}, \; y > 0.$$

Clearly σ^2 is a strictly positive continuous function on $(0, \infty)$. It is easy to see that $\sigma^2(0) \equiv \sigma^2(0+) = 0$. Note that $u'(x)$ is strictly positive and continuous on $(0, \infty)$ and has right-hand limit ∞ at $x = 0$. It follows easily that s' is continuous on $[0, \infty)$ and

$$s'(y) = u'(s(y))^{-1} = \exp\Big\{\int_{x_0}^{s(y)} \frac{b(z)}{z} dz\Big\} \quad \text{for all } y > 0, \quad s'(0) = 0. \tag{8.2}$$

It follows from the above that

$$s'(y)\sigma(y) = \sqrt{2s(y)} \quad \text{for all } y \geq 0. \tag{8.3}$$

It is clear from (8.2) that s' is differentiable on $(0,\infty)$ and if we differentiate both sides of (8.2) we easily derive

$$s''(y)\sigma^2(y)/2 = b(s(y)) \quad \text{for all } y > 0. \tag{8.4}$$

Now let $Z_t = W_t + L_t^0(Z)$ be a reflecting Brownian motion in $\mathbb{R}_+$ starting at 0, where W_t is a Brownian motion starting at 0 and $L_t^a(Z)$ is the local time at a of Z_t. Set $A_t = \int_0^t \sigma^{-2}(Z_s)ds = \int_0^\infty \sigma(a)^{-2}L_t^a(Z)da$. An easy calculation shows that

$$\int_0^{x_0} \sigma^{-2}(a)da = \int_0^{x_0} \frac{1}{2x}(\log(1/x)/c)^{-c}dx < \infty$$

because $c > 1$. Therefore A is finite. Clearly A is strictly increasing and as Z is recurrent, $\lim_{t\to\infty} A_t = \infty$ a.s. Therefore A has a continuous inverse τ_t, $t \geq 0$. Let $Y_t = Z_{\tau_t}$, so that Y satisfies

$$Y_t = M_t + L_t^0(Y).$$

Here $L_t^0(Y)$ is the local time of Y at 0, and M is a martingale with $\langle M\rangle_t = \int_0^t \sigma^2(Y_s)ds$. Hence we can write $dM_t = \sigma(Y_t)dB_t$, where B is a Brownian motion.

We now define $X_t = s(Y_t)$. Since s' is increasing (by (8.2)), we may apply Tanaka's formula to see that

$$\begin{aligned}
X_t &= \int_0^t s'(Y_r)dY_r + \frac{1}{2}\int_0^\infty L_t^a ds'(a) \\
&= \int_0^t s'(Y_r)\sigma(Y_r)dB_r + \int_0^t s'(Y_r)dL_r^0(Y) + \frac{1}{2}\lim_{\delta\downarrow 0}\int_\delta^\infty L_t^a s''(a)da \\
&= \int_0^t s'(Y_r)\sigma(Y_r)dB_r + \frac{1}{2}\lim_{\delta\downarrow 0}\int_0^t s''(Y_r)\sigma^2(Y_r)1_{(Y_r\geq\delta)}dr.
\end{aligned}$$

In the last line we have used the fact that $s'(0) = 0$ (see (8.2)). Now use (8.3) and (8.4) to see that

$$X_t = \int_0^t \sqrt{2X_s}dB_s + \lim_{\delta\downarrow 0}\int_0^t b(X_s)1_{(Y_s\geq\delta)}ds.$$

Since $b(0) = b(0+) = 0$, it is clear that (8.1) follows from the above. As it is clear that X is not identically 0 from its definition, we are done. □

References

[AT00] S. Athreya and R. Tribe, Uniqueness for a class of one-dimensional stochastic PDEs using moment duality. *Ann. Probab.* **28**, 1711–1734, 2000.

[B95] R. F. Bass, *Probabilistic Techniques in Analysis*, Springer, New York, 1995.

[B97] R. F. Bass, *Diffusions and Elliptic Operators*, Springer, New York, 1997.

[BP01] R.F. Bass, E.A. Perkins, Degenerate stochastic differential equations with Hölder continuous coefficients and super-Markov chains. Preprint.

[BG68] R. M. Blumenthal and R.K. Getoor, *Markov Processes and Potential Theory*, Academic Press, New York, 1968.

[B68] L. Breiman, *Probability*, Addison-Wesley, Reading, Mass, 1968.

[D78] D. Dawson, Geostochastic calculus, *Can. J. Statistics* **6**, 143–168, 1978.

[D93] D. Dawson, Measure-valued Markov Processes, Ecole d'Eté de Probabilités de Saint Flour 1991, Lect. Notes in Math. 1541, Springer, Berlin, 1993.

[DEFMPX00] D. Dawson, A. Etheridge, K. Fleischmann, L. Mytnik, E. Perkins, and J. Xiong, Mutually catalytic branching in the plane: finite measure states, preprint.

[DK99] P. Donnelly and T. Kurtz, Particle representations for measure-valued population models. *Ann. Probab.* **27** 166–205, 1999.

[DM95] D. Dawson and P. March, Resolvent estimates for Fleming-Viot operators and uniqueness of solutions to related martingale problems, *J. Funct. Anal.* **132**, 417–472, 1995.

[DP98] D. Dawson and E. Perkins, Longtime behaviour and coexistence in a mutually catalytic branching model, *Ann. Probab.* **26**, 1088–1138, 1998.

[FX00] K. Fleischmann and J. Xiong, A cyclically catalytic super-Brownian motion. *Ann. Probab.* **29**, 820–861, 2001.

[GKW99] A.Greven, A.Klenke, and A. Wakolbinger, Interacting Fisher-Wright diffusions in a catalytic medium. *Probab. Th. Rel. Fields* **120**, 85-117, 2001.

[HLP34] G.H. Hardy, J.E. Littlewood and R.E.A.C. Paley, *Inequalities*, Cambridge Univ. Press, Cambridge, 1934.

[H71] H. Hochstadt, *Functions of Mathematical Physics*, Wiley, New York, 1971.

[K98] T. Kurtz, Martingale problems for conditional distributions of Markov processes, *Elect. J. Prob.* **3**, paper 9, 1998.

[Ke84] H. J. Keisler, An infinitesimal approach to stochastic analysis, *Mem. Amer. Math. Soc.* no. 297, 1984.

[M98] L. Mytnik, Weak uniqueness for the heat equation with noise, *Ann. Probab.* **26**, 968–984, 1998.

[P92] E. Perkins, Measure-valued branching diffusions with spatial interactions, *Probab. Th. Rel. Fields* **94**, 189-245, 1992.

[P95] E. Perkins, On the martingale problem for interactive measure-valued diffusions, *Mem. Amer. Math. Soc.* **115** no. 549, 1995.

[P01] E. Perkins, Dawson-Watanabe Superprocesses and Measure-Valued Diffusions, to appear, Ecole d'Eté de Probabilités de Saint Flour 1999, Lect. Notes in Math. Springer, Berlin, 2001.

[RW87] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales, Vol. 2*, Wiley, Chichester, 1987.

[RY91] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer, Berlin, 1991.

[S00] J. Swart, On pathwise uniqueness. *To appear in Stoch. Proc. App.*

[SV79] D.W. Stroock, S.R.S. Varadhan, *Multidimensional diffusion processes*, Springer, Berlin, 1979.

[W44] G.N. Watson, *A Treatise on the Theory of Bessel Functions, Second Ed.*, University Press, Cambridge, 1944.

[YW71] T. Yamada and S. Watanabe, On the uniqueness of solutions of stochastic differential equations, *J. Math. Kyoto U.* **11**, 155–167, 1971.

Addresses:
S.R. Athreya: Stat-Math unit, Indian Statistical Institute, New Delhi, India - 110016.

M.T. Barlow, E.A. Perkins: Department of Mathematics, University of British Columbia, Vancouver, B.C., Canada V6T 1Z2.

R.F. Bass: Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA.