Constantinos Kardaras and Scott Robertson
Continuous-time perpetuities and time reversal of diffusions

Article (Accepted version)
(Refereed)

http://eprints.lse.ac.uk

CrossMark

Continuous-time perpetuities and time reversal of diffusions

Constantinos Kardaras[1] · Scott Robertson[2]

Abstract We consider the problem of estimating the joint distribution of a continuous-time perpetuity and the underlying factors which govern the cash flow rate, in an ergodic Markovian model. Two approaches are used to obtain the distribution. The first identifies a partial differential equation for the conditional cumulative distribution function of the perpetuity given the initial factor value, which under certain conditions ensures the existence of a density for the perpetuity. The second (and more general) approach, using techniques of time reversal, identifies the joint law as the stationary distribution of an ergodic multidimensional diffusion. This latter approach allows efficient use of Monte Carlo simulation, as the distribution is obtained by sampling a single path of the reversed process.

✉ C. Kardaras
k.kardaras@lse.ac.uk

S. Robertson
scottrob@andrew.cmu.edu

[1] Department of Statistics, London School of Economics, 10 Houghton Street, London,
WC2A 2AE, England, UK

[2] Department of Mathematical Sciences, Carnegie Mellon University, Wean Hall 6113,
Pittsburgh, PA 15213, USA

⚛ Springer

1 Introduction

1.1 Discussion

In this article, we consider a *continuous-time perpetuity* given by the random variable

$$X_0 := \int_0^\infty D_t f(Z_t) \, dt. \tag{1.1}$$

Above, $Z = (Z_t)_{t \in \mathbb{R}_+}$ represents the value of an economic factor that determines a cash flow rate $(f(Z_t))_{t \in \mathbb{R}_+}$. Cash flows are discounted according to $D = (D_t)_{t \in \mathbb{R}_+}$; therefore, X_0 represents the whole payment in units of account at time zero. Our main concern is the identification of an efficient way to obtain the joint distribution of (Z_0, X_0), as naive estimation of the distribution by simulating sample paths of Z and approximating X_0 through numerical integration may be prohibitively slow. As Z_0 is typically observable, the joint distribution of (Z_0, X_0) also allows us to obtain the conditional distribution of X_0 given Z_0.

In order to make the problem tractable, we work in a diffusive, Markovian environment where Z and D are solutions to the respective stochastic differential equations (written in integrated form)[1]

$$Z = Z_0 + \int_0^{\cdot} m(Z_t) \, dt + \int_0^{\cdot} \sigma(Z_t) \, dW_t, \tag{1.2}$$

$$D = 1 - \int_0^{\cdot} D_t \big(a(Z_t) \, dt + \theta(Z_t)' \sigma(Z_t) \, dW_t + \eta(Z_t)' \, dB_t \big). \tag{1.3}$$

In the above equations, W and B are independent Brownian motions of dimension d and k, respectively, while m, σ, a, θ and η are given functions. (Precise assumptions on all the model coefficients are given in Sect. 2.) We assume Z is stationary and ergodic with invariant density p. Equation (1.3) includes in particular the case when D is smooth, in other words, $D = \exp(- \int_0^{\cdot} a(Z_t) \, dt)$, where a represents a short-rate function. However, the more general form of (1.3) is considered to accommodate a broader range of situations; for example,

– when payment streams are denominated in different units of account (for example, another currency, or financial assets), in which case discounting has to take into account the "exchange rate";
– when for pricing purposes, the payment stream, though denominated in domestic currency, must incorporate both traditional discounting and the density of the pricing kernel.

The two main results of the paper—Theorems 3.1 and 3.4—identify the distribution of (Z_0, X_0) in different ways. First, in the case where η in (1.3) is non-degenerate and f in (1.1) is sufficiently regular, the conditional cumulative distribution function of X_0 given Z_0 is shown to coincide with the explosion probability of an associated locally elliptic diffusion and hence, through the Feynman–Kac formula, satisfies a

---

[1]Throughout the text, the prime symbol $'$ denotes transposition.

partial differential equation (PDE); see Theorem 3.1. Second, for general η and f, using methods of diffusion time reversal, we identify an "ergodic" process (ζ, χ) whose invariant distribution coincides with the joint distribution of (Z_0, X_0). In particular, for any fixed starting point $x > 0$ of χ, the (random) empirical time-average law of (ζ, χ) on $[0, T]$ almost surely converges to the joint distribution of (Z_0, X_0) in the weak topology; see Theorem 3.4. The time-reversal result has the advantage of leading to an efficient method for obtaining the distribution via simulation, as the ergodic theorem enables estimation of the entire distribution based upon a single realization of (ζ, χ); a numerical example in Sect. 4 dramatically reinforces this point. However, it must be noted that the invariant distribution p for Z appears in the reversed dynamics, and hence must be known to perform simulation. When Z is one-dimensional, or more generally reversing, in the sense that the second order linear differential operator associated to its generator is symmetric on a certain Hilbert space (see [27, Sect. 4.10]), p is given in explicit form with respect to the model parameters. In the general multidimensional setup, lack of knowledge of p could pose an issue; however, we provide a potential way to amend the situation in the discussion after Theorem 3.4. Note also that in the PDE result in Theorem 3.1, explicit knowledge of p is not necessary.

1.2 Existing literature and connections

Obtaining the distribution of the perpetuity X_0 is of great importance in the areas of finance and actuarial science; for this reason, perpetuities with a form similar to X_0 have been extensively studied. For example, [11] deals with the case where

$$X_0 = \int_0^\infty e^{-\sigma B_t - \nu t} \, \mathrm{d}t,$$

establishing that X_0 has an inverse gamma distribution. This fits into the setup of (1.2), (1.3) by taking $a = \nu - \sigma^2/2$, $f = 1$, $\theta = 0$ and $\eta = \sigma$. Note that here Z plays no role. In a similar manner, [31, Chap. 5] and [9, 10] consider the case

$$X_0 = \int_0^\infty e^{-\int_0^t Z_u \, \mathrm{d}u} \, \mathrm{d}t, \qquad \mathrm{d}Z_t = \kappa(\theta - Z_t)\mathrm{d}t + \xi\sqrt{Z_t}\mathrm{d}W_t,$$

and obtain the first moment, along with bounds for other moments, of X_0. In [16], the perpetuity takes the form

$$X_0 = \int_0^\infty e^{-Q_t} \, \mathrm{d}P_t, \quad \text{with } P \text{ and } Q \text{ being independent Lévy processes.} \quad (1.4)$$

Under certain conditions on P and Q, the distribution of X_0 is implicitly calculated by identifying the characteristic function and/or Laplace transform for X_0. In fact, the results of [16] are pre-dated (for highly particular P and Q) in [24, 21]. The Laplace transform method is also used in [26, 25] to treat (1.4) when $P_t = t$ and Q is a diffusion. In addition to identifying a degenerate elliptic partial differential equation for the Laplace transform, they propose a candidate recurrent Markov chain whose invariant distribution has the law of X_0. Lastly, the setup of [16] is significantly

extended in [6] where under minimal assumptions on P and Q, the distribution of X_0 is shown to coincide with the unique invariant measure for a certain generalized Ornstein–Uhlenbeck process, a relationship that is confirmed in our current setting in Proposition 9.2.

The use of time reversal to identify the distribution of a discrete-time perpetuity is well known, dating at least back to [12], where X_0 takes the form

$$X_0 = \sum_{n=1}^{\infty} \left(\prod_{i=1}^{n} D_i \right) f_n,$$

where the discount factors $(D_n)_{n \in \mathbb{N}}$ and cash flows $(f_n)_{n \in \mathbb{N}}$ are two independent sequences of independent, identically distributed (iid) random variables. To provide insight, the time-reversal argument in [12] is briefly presented here. With

$$X_0^{(N)} := \sum_{n=1}^{N} \left(\prod_{i=1}^{n} D_i \right) f_n,$$

it is clear by the iid property that $X_0^{(N)}$ has the same distribution as

$$\widetilde{X}_N := D_N f_N + D_N D_{N-1} f_{N-1} + \cdots + \left(\prod_{j=1}^{N} D_j \right) f_1.$$

Straightforward calculations show that the reversed process $(\widetilde{X}_n)_{n \in \mathbb{N}}$ satisfies the recursive equation $\widetilde{X}_n = D_n(\widetilde{X}_{n-1} + f_n)$. Thus, assuming that $(\widetilde{X}_n)_{n \in \mathbb{N}}$ converges to a random variable $\widetilde{X}$ in distribution, $\widetilde{X}$ must solve the distributional equation $\widetilde{X} = D(\widetilde{X} + f)$, where D, f and $\widetilde{X}$ are independent, D has the same law as D_1 and f has the same law as f_1. In [30], solutions to that distributional equation are obtained based upon the expectation of $\log |D|$ and $\log^+ |Df|$. The tails of $\widetilde{X}$, as well as convergence of iterative schemes, are studied in [14]; furthermore, [17] gives "almost" if and only if conditions for the convergence of iterative schemes.

In a continuous-time setting, we employ an argument similar in spirit, but rather different in execution, to [12]. Specifically, we extend X_0 to a whole "forward" process $X := (1/D) \int_{\cdot}^{\infty} D_t f(Z_t) dt$, and then for each $T > 0$ define the reversed process (ζ^T, χ^T) on $[0, T]$ by $\zeta_t^T := Z_{T-t}$, $\chi_t^T := X_{T-t}$; see (3.6), (3.7). Using results on time reversal of diffusions from [19] (alternatively, see [23, 3, 7, 13]) as well as additional elementary calculations, we obtain the dynamics for (ζ^T, χ^T). In fact, Proposition 8.5 shows that the generator of (ζ^T, χ^T) does not depend upon T and ergodicity can be studied for the process (ζ, χ) with the given generator. When $|\eta| > 0$ and f is sufficiently regular, this generator is locally elliptic and the associated process (ζ, χ) is ergodic with invariant distribution equal to that of (Z_0, X_0); see Proposition 9.2. In the general case, a slightly weaker (but still sufficient) form of ergodicity still holds: starting ζ from its invariant distribution p and χ from any starting point $x > 0$, the (random) empirical time-average laws of (ζ, χ) converge almost surely in the weak topology to the distribution of (Z_0, X_0).

1.3 Structure

This paper is organized as follows. In Sect. 2, we precisely state the assumptions on the processes Z and D, as well as the function f, paying particular attention to deriving sharp conditions under which X_0 is almost surely finite or infinite. The main results are then presented in Sect. 3. First, when $|\eta| > 0$ and f is sufficiently regular, the conditional cumulative distribution function of X given $Z_0 = z$ is shown to satisfy a certain partial differential equation. Then, using the method of time reversal, we construct a probability space and diffusion (ζ, χ) such that with probability one, its empirical time-average laws weakly converge to the joint distribution of (Z_0, X_0) for all starting points of χ. Section 3 concludes with a brief discussion how the distribution may be estimated via simulation, in particular proposing a method for obtaining the desired distribution when the invariant density p for Z is not explicitly known. Section 4 provides a numerical example in a specific case where the joint distribution of (Z_0, X_0) is explicitly identifiable. Here, we compare the performance of the reversal method versus the direct method for obtaining the distribution of X_0. In particular, we show that for a given desired level of accuracy (see Sect. 4 for a more precise definition), the method of time reversal is approximately 175 to 300 times faster than the direct method. The remaining sections contain the proofs. Section 6 proves the statements regarding the finiteness of X_0; Sect. 7 proves the partial differential equation result; Sect. 8 obtains the dynamics for the time-reversed process (ζ, χ); Sect. 9 proves the (weak) ergodicity with the correct invariant distribution. Finally, a number of technical supporting results are included in the Appendix.

2 Problem setup

2.1 Well-posedness and ergodicity

The first order of business is to specify precise coefficient assumptions so that Z in (1.2) and D in (1.3) are well defined. As for Z, we work in the standard locally elliptic setup for diffusions; for more information, see [27, Chaps. 3.7, 4.1]. Let $E \subseteq \mathbb{R}^d$ be an open, connected region. We assume the existence of $\gamma \in (0, 1]$ such that

(A1) there exists a sequence of regions $(E_n)_{n \in \mathbb{N}}$ such that $E = \bigcup_{n=1}^{\infty} E_n$, with each E_n being open, connected, bounded, with ∂E_n being $C^{2,\gamma}$ and satisfying $\bar{E}_n \subset E_{n+1}$ for all $n \in \mathbb{N}$;

(A2) $m \in C^{1,\gamma}(E; \mathbb{R}^d)$ and $c \in C^{2,\gamma}(E; \mathbb{S}_{++}^d)$, where $\mathbb{S}_{++}^d$ is the space of symmetric and strictly positive definite $(d \times d)$-dimensional matrices.

With the provisos in (A1) and (A2), define L^Z as the generator associated to (m, c), i.e.,[2]

$$L^Z := \frac{1}{2} \sum_{i, j=1}^{d} c^{ij} \partial_{ij}^2 + \sum_{i=1}^{d} m^i \partial_i.$$

---

[2]In the sequel, the summation indices will be omitted using Einstein's convention; therefore, L^Z will be written as $L^Z = (1/2)c^{ij} \partial_{ij}^2 + m^i \partial_i$.

Under (A1) and (A2), one can infer the existence of a solution to the martingale problem for L^Z on E, with the possibility of explosion to the boundary of E; see [27, Chap. 1.13]. We wish for something stronger, namely, to construct a filtered probability space $(\Omega, \mathbf{F}, \mathbb{P})$ on which there is a strong, stationary, ergodic solution to the SDE in (1.2) with invariant density p. In (1.2), W is a d-dimensional Brownian motion and $\sigma = \sqrt{c}$, the unique positive definite symmetric matrix such that $\sigma^2 = c$. In order to achieve this, we ask that

(A3) the martingale problem for L^Z on E is well posed and the corresponding solution is recurrent. Furthermore, there exists a strictly positive $p \in C^{2,\gamma}(E, \mathbb{R})$ with $\int_E p(z)\, \mathrm{d}z = 1$ and satisfying $\tilde{L}^Z p = 0$, where $\tilde{L}^Z$ is the formal adjoint of L^Z given by

$$\tilde{L}^Z := \frac{1}{2} c^{ij} \partial_{ij}^2 - (m^i - \partial_j c^{ij}) \partial_i - \left(\partial_i m^i - \frac{1}{2} \partial_{ij}^2 c^{ij} \right). \tag{2.1}$$

We summarize the situation in the following result; the extra Brownian motion B in its statement will be used to define the process D via (1.3) later on.

Theorem 2.1 *Under Assumptions* (A1)–(A3), *there exists a filtered probability space* $(\Omega, \mathbf{F}, \mathbb{P})$ *satisfying the usual conditions and supporting two independent Brownian motions* W *and* B, d-*dimensional and* k-*dimensional, respectively, such that* Z *satisfies* (1.2) *and is stationary and ergodic with invariant density* p.

Remark 2.2 According to [27, Corollary 5.1.11], in the one-dimensional case where $E = (\alpha, \beta)$ for $-\infty \leq \alpha < \beta \leq \infty$, the above Assumption (A3) is true if and only if for some $z_0 \in E$,

$$\int_\alpha^{z_0} \exp\left(-2 \int_{z_0}^z \frac{m(s)}{c(s)}\, \mathrm{d}s \right) \mathrm{d}z = \infty,$$

$$\int_{z_0}^\beta \exp\left(-2 \int_{z_0}^z \frac{m(s)}{c(s)}\, \mathrm{d}s \right) \mathrm{d}z = \infty,$$

$$\int_\alpha^\beta \frac{1}{c(z)} \exp\left(2 \int_{z_0}^z \frac{m(s)}{c(s)}\, \mathrm{d}s \right) \mathrm{d}z < \infty.$$

In this case, it holds that

$$p(z) = K c^{-1}(z) \exp\left(2 \int_{z_0}^z \frac{m(s)}{c(s)}\, \mathrm{d}s \right), \quad z \in (\alpha, \beta),$$

where $K > 0$ is a normalizing constant.

In the multidimensional case, suppose that there exists a function $H : E \to \mathbb{R}$ with the property that $c^{-1}(2m - \operatorname{div} c) = \nabla H$, where $\operatorname{div} c$ is the (matrix) divergence defined by[3] $(\operatorname{div} c)^i = \partial_j c^{ij}, i = 1, \ldots, d$. Then Z from Theorem 2.1 is a reversing

---

[3]This definition is equivalent to the standard definition of divergence for matrices, where the divergence operator is applied to the columns, by the symmetry of c. Also, to differentiate the matrix divergence from its vector counterpart, we write $\operatorname{div} A$ for symmetric matrices A and $\nabla \cdot v$ for vector-valued functions v.

Markov process in the sense that the time-reversed process on any interval $[0, T]$ has the same dynamics as Z; see [19]. Furthermore, Assumption (A3) follows if it can be shown that Z does not explode to the boundary of E and $K := \int_E \exp(H(z)) \, dz < \infty$. Indeed, by construction, $p = e^H / K$ satisfies $\tilde{L}^Z p = 0$ and $\int_E p(z) \, dz = 1$. Thus if Z does not explode, it follows from [27, Theorem 2.8.1, Corollary 4.9.4] that Z is recurrent. In fact, Z is ergodic, as shown in [27, Theorems 4.3.3, 4.9.5]. If we are not in the reversing case, there are many known techniques for checking ergodicity; see [5, 27]. For example, if there exist a smooth function $u : E \to \mathbb{R}$, an integer N and constants $\varepsilon > 0$ and $C > 0$ such that $L^Z u \le -\varepsilon$ and $u \ge -C$ on $E \setminus E_N$, then (A3) holds.

In order to ensure that D in (1.3) is well defined, we assume that
(A4) $a \in C^{1,\gamma}(E; \mathbb{R}_+)$, $\eta \in C^{2,\gamma}(E; \mathbb{R}^k)$ and $\theta \in C^{2,\gamma}(E; \mathbb{R}^d)$.

Given (A4) and all previous assumptions, it follows that (1.3) possesses a strong solution on $(\Omega, \mathbf{F}, \mathbb{P})$ from Theorem 2.1; in fact, defining $R := -\log D$, it holds that

$$R = \int_0^\cdot \left(a + \frac{1}{2}(\theta'c\theta + |\eta|^2) \right) (Z_t) \, dt + \int_0^\cdot \theta(Z_t)'\sigma(Z_t) \, dW_t + \int_0^\cdot \eta(Z_t)' \, dB_t. \tag{2.2}$$

2.2 Finiteness of X_0

Having the setup for the existence of Z and D, we proceed to X_0. For the time being, we just assume[4] that the function $f : E \to \mathbb{R}_+$ is in $\mathbb{L}^1(E, p)$. For the PDE results of Theorem 3.1 below, we require a slightly stronger regularity assumption on f, although the time-reversal results of Theorem 3.4 make no additional assumptions. Now, for f not necessarily in $\mathbb{L}^1(E, p)$, it is entirely possible that X_0 takes infinite values with positive probability. In this section, conditions are given under which $\mathbb{P}[X_0 < \infty] = 1$ or, conversely, when $\mathbb{P}[X_0 < \infty] = 0$.

Lemma 2.3 *Let* (A1)–(A4) *hold. For the invariant density p of Z, assume there exists $\varepsilon > 0$ such that*

$$\left(a + \frac{1-\varepsilon}{2}(\theta'c\theta + \eta'\eta) \right)^- \in \mathbb{L}^1(E, p),$$
$$\int_E \left(a + \frac{1-\varepsilon}{2}(\theta'c\theta + \eta'\eta) \right)(z) \, p(z) \, dz > 0. \tag{2.3}$$

Then the following hold:

(i) *There exists $\kappa > 0$ such that for all $z \in E$, $\mathbb{P}[\lim_{t\to\infty} e^{\kappa t} D_t = 0 \mid Z_0 = z] = 1$. In particular, $\lim_{t\to\infty} e^{\kappa t} D_t = 0$ $\mathbb{P}$-a.s.*
(ii) *For any $f \in \mathbb{L}^1(E, p)$, it holds that $\mathbb{P}[X_0 < \infty] = 1$.*

---

[4]We define $\mathbb{L}^1(E, p)$ to be those Borel-measurable functions g on E so that $\int_D |g(z)|p(z) \, dz < \infty$. Thus, Borel-measurability is implicitly assumed throughout.

Remark 2.4 Note that (2.3) holds if $a > 0$ on E. The more complicated form in (2.3) allows a to take (unbounded) negative values. Furthermore, in the case where $(\theta'c\theta + \eta'\eta) \in \mathbb{L}^1(E, p)$, (2.3) is equivalent to

$$\left(a + \frac{1}{2}(\theta'c\theta + \eta'\eta)\right)^{-} \in \mathbb{L}^1(E, p), \quad \int_E \left(a + \frac{1}{2}(\theta'c\theta + \eta'\eta)\right)(z)\, p(z)\, \mathrm{d}z > 0.$$

As a partial converse to Lemma 2.3, we have

Lemma 2.5 *Let* (A1)–(A4) *hold. For the invariant density p of Z, assume there exists $\varepsilon > 0$ such that*

$$\left(a + \frac{1+\varepsilon}{2}(\theta'c\theta + \eta'\eta)\right)^{+} \in \mathbb{L}^1(E, p),$$

$$\int_E \left(a + \frac{1+\varepsilon}{2}(\theta'c\theta + \eta'\eta)\right)(z)\, p(z)\, \mathrm{d}z \le 0.$$

(If $\theta'c\theta + \eta'\eta \equiv 0$, then assume that $a^+ \in \mathbb{L}^1(E, p)$ and $\int_E a(z)p(z)\, \mathrm{d}z < 0$.) If f is such that $\int_E f(z)p(z)\, \mathrm{d}z > 0$, then $\mathbb{P}[X_0 < \infty] = 0$.

Remark 2.6 Let (A1)–(A4) hold and assume that a is nonnegative. A combination of Lemmas 2.3 and 2.5 yields sharp conditions for the finiteness of X_0 that do not require knowledge of p, at least for bounded f:

– If $a + (1/2)(\theta'c\theta + \eta'\eta) \not\equiv 0$, then $\mathbb{P}[X_0 < \infty] = 1$ holds if $f \in \mathbb{L}^1(E, p)$.
– If $a + (1/2)(\theta'c\theta + \eta'\eta) \equiv 0$, then $\mathbb{P}[X_0 < \infty] = 0$ holds if $\int_E f(z)p(z)\, \mathrm{d}z > 0$.

In view of Lemma 2.3, we ask that

(A5) $f \in \mathbb{L}^1_+(E, p)$, $\int_E f(z)p(z)\, \mathrm{d}z > 0$ and there exists $\varepsilon > 0$ such that

$$\left(a + \frac{1-\varepsilon}{2}(\theta'c\theta + \eta'\eta)\right)^{-} \in \mathbb{L}^1(E, p),$$

$$\int_E \left(a + \frac{1-\varepsilon}{2}(\theta'c\theta + \eta'\eta)\right)(z)\, p(z)\, \mathrm{d}z > 0.$$

To recapitulate, **for the remainder of the article,** the following is assumed:

Assumption 2.7 We enforce throughout all the above Assumptions (A1)–(A5).

3 Main results

3.1 The distribution of X_0 via a partial differential equation

Define the cumulative distribution function g of X_0 given Z_0 by

$$g(z, x) := \mathbb{P}[X_0 \le x \mid Z_0 = z], \quad (z, x) \in F := E \times (0, \infty). \tag{3.1}$$

Next, recall that Assumption 2.7 implies that Z_0 has a density p, and define the joint distribution π of (Z_0, X_0) by

$$\pi(A) := \iint_A p(z)g(z, dx)\,dz, \qquad A \in \mathcal{B}(F). \tag{3.2}$$

Under Assumption 2.7, as well as an additional smoothness requirement on f and non-degeneracy requirement on η, the first main result (Theorem 3.1 below) shows that g solves a certain PDE on the state space F. This will imply that the joint distribution of (Z_0, X_0) has a density (still labeled π) and the law of X_0 charges all of $(0, \infty)$.

To motivate the result as well as to fix notation, for each $x \in (0, \infty)$, consider the process

$$Y^x := \frac{1}{D}\left(x - \int_0^{\cdot} D_t f(Z_t)\,dt\right).$$

Since Assumption 2.7 implies that $\mathbb{P}[\lim_{t \to \infty} D_t = 0 \mid Z_0 = z] = 1$ for all $z \in E$, it is clear that given $Z_0 = z$, the process Y^x tends to ∞ on $\{X_0 < x\}$. Alternatively, on $\{X_0 > x\}$, Y^x will hit 0 at some finite time. What happens on $\{X_0 = x\}$ is not immediately clear, but it will be shown under the given assumptions that there is probability zero of this occurring. For fixed $(z, x) \in F$, it follows that $1 - g(z, x)$ equals the probability that Y^x hits zero, given $Z_0 = z$. According to the Feynman–Kac formula, such probabilities "should" solve a PDE. To identify the PDE, note that the joint equations governing Z and Y^x are

$$Z = Z_0 + \int_0^{\cdot} m(Z_t)\,dt + \int_0^{\cdot} \sigma(Z_t)\,dW_t,$$

$$Y^x = x + \int_0^{\cdot}\left(-f(Z_t) + Y_t^x\big(a(Z_t) + \theta'c\theta(Z_t) + \eta'\eta(Z_t)\big)\right)dt$$

$$+ \int_0^{\cdot} Y_t^x\big(\theta'\sigma(Z_t)\,dW_u + \eta(Z_t)'\,dB_t\big).$$

Define $b : F \to \mathbb{R}^{d+1}$ and $A : F \to \mathbb{S}_{++}^{d+1}$ by

$$b(z, x) := \begin{pmatrix} m(z) \\ -f(z) + x(a + \theta'c\theta + \eta'\eta)(z) \end{pmatrix},$$

$$A(z, x) := \begin{pmatrix} c(z) & xc\theta(z) \\ x\theta'c(z) & x^2(\theta'c\theta + \eta'\eta)(z) \end{pmatrix}, \tag{3.3}$$

for all $(z, x) \in F$. Note that if in addition to Assumption 2.7, $|\eta|(z) > 0$ for $z \in E$, then A is locally elliptic. Let L be the second order differential operator associated to (A, b), i.e.,

$$L := \frac{1}{2}A^{ij}\partial_{ij}^2 + b^i\,\partial_i. \tag{3.4}$$

Note that $L\phi = L^Z\phi$ for functions ϕ of $z \in E$ alone. With the previous notation, the first main result now follows.

Theorem 3.1 *Let Assumption* 2.7 *hold, and suppose further that*
(a) $f \in C^{1,\gamma}(E; \mathbb{R}_+)$,
(b) $|\eta(z)| > 0$ *for all* $z \in E$.
Then g *is in* $C^{2,\gamma}(F)$ *and satisfies* $Lg = 0$ *with the "locally uniform" boundary conditions*

$$\lim_{n \to \infty} \sup_{x \leq n^{-1}, z \in E_k} g(z, x) = 0, \qquad \lim_{n \to \infty} \inf_{x \geq n, z \in E_k} g(z, x) = 1, \quad \forall k \in \mathbb{N}. \tag{3.5}$$

Furthermore, g *is the unique function satisfying* $Lg = 0$ *with the above boundary conditions, in the set of functions* $\{\tilde{g} \in C^2(F) : 0 \leq \tilde{g} \leq 1\}$.

Remark 3.2 The non-degeneracy assumption on η is essential for the existence of a density; if $\eta \equiv 0$, it may be that the distribution of X_0 has an atom. Indeed, take $f \equiv 1$, $a \equiv 1$, $\eta \equiv 0$, $\theta \equiv 0$. Then $X_0 = \int_0^\infty e^{-t} \, dt = 1$ with probability one.

Remark 3.3 Theorem 3.1 implies that the law of X_0 charges all of $(0, \infty)$, even for those functions f which are bounded from above. Theorem 3.1 also implies that X_0 has a density without imposing Hörmander's condition [22, Chap. 2] on the coefficients in (3.3). Rather, the infinite horizon combined with the presence of the independent Brownian motion B "smooth out" the distribution of X_0.

Theorem 3.1 is certainly important from a theoretical viewpoint. However, it appears to be of limited practical use. Even under the extra non-degeneracy condition $|\eta| > 0$, it is unclear how to numerically solve the PDE $Lg = 0$ with the given boundary conditions (3.5), as there are no natural auxiliary boundary conditions in the spatial domain of $z \in E$. In Sect. 3.2 that follows, we provide an alternative, more useful method for estimating numerically the law of (Z_0, X_0).

3.2 The distribution of (Z_0, X_0) via diffusion time reversal

The goal here is to show that the distribution of (Z_0, X_0) coincides with the invariant distribution of a positive recurrent process (ζ, χ). In order to see the connection, extend X_0 to a process $(X_t)_{t \in \mathbb{R}_+}$ defined via

$$X := \frac{1}{D} \int_{\cdot}^{\infty} D_t f(Z_t) \, dt, \tag{3.6}$$

and note that $(Z_t, X_t)_{t \in \mathbb{R}_+}$ is a stationary process under $\mathbb{P}$. Fix $T > 0$ and define the process $(\zeta_t^T, \chi_t^T)_{t \in [0,T]}$ via time reversal, i.e.,

$$\zeta_t^T := Z_{T-t}, \qquad \chi_t^T := X_{T-t}, \qquad t \in [0, T]. \tag{3.7}$$

It still follows that (ζ^T, χ^T) is stationary under $\mathbb{P}$, with the same one-dimensional marginal distribution as (Z_0, X_0). Furthermore, stationarity of (Z, X) clearly implies that the law of the process (ζ^T, χ^T) does not depend on T (except for its time-domain of definition). Therefore, one may create a new process $(\zeta_t, \chi_t)_{t \in \mathbb{R}_+}$, on a potentially different probability space (e.g. the space of continuous functions), such that the law

of (ζ^T, χ^T) is the same as the law of $(\zeta_t, \chi_t)_{t \in [0,T]}$ for all $t \in T$. If one can establish that (ζ, χ) is ergodic, then the distribution of (Z_0, X_0) may be efficiently estimated via the ergodic theorem.

Towards this end, one needs to understand the behavior of (ζ, χ). Standard results (e.g. [19]) in the theory of time reversal imply that ζ is a diffusion in its own filtration, and identify the corresponding coefficients. In order to deal with χ, we return to the definition of χ^T and define yet one more process $(\Delta_t^T)_{t \in [0,T]}$ via

$$\Delta_t^T = \frac{D_T}{D_{T-t}}, \quad t \in [0, T]. \tag{3.8}$$

Using all previous definitions, we obtain that

$$\begin{aligned}
\chi_t^T = X_{T-t} &= \frac{1}{D_{T-t}} \int_{T-t}^{\infty} D_u f(Z_u) \, du \\
&= \frac{D_T}{D_{T-t}} \left(X_T + \int_{T-t}^{T} \frac{D_u}{D_T} f(Z_u) \, du \right) \\
&= \Delta_t^T \left(\chi_0^T + \int_0^t \frac{1}{\Delta_u^T} f(\zeta_u^T) \, du \right), \quad t \in [0, T]. \tag{3.9}
\end{aligned}$$

As it turns out, one can describe the joint dynamics of (ζ^T, Δ^T) in appropriate filtrations (and these dynamics do not depend on T, as expected). To ease the presentation, recall from Sect. 2 that for any $\mathbb{S}_{++}^d$-valued smooth function A on E, the (matrix) divergence is defined by $(\text{div } A)^i = \partial_j A^{ij}$ for $i = 1, \ldots, d$. It is then shown in Sect. 8 that (ζ^T, Δ^T) is such that

$$\begin{aligned}
\zeta^T &= \zeta_0^T + \int_0^{\cdot} \left(c \frac{\nabla p}{p} + \text{div } c - m \right) (\zeta_t^T) \, dt + \int_0^{\cdot} \sigma(\zeta_t^T) \, dW_t^T, \\
\Delta^T &= 1 + \int_0^{\cdot} \Delta_t^T \left(\theta' c \frac{\nabla p}{p} + \nabla \cdot (c\theta) - a \right) (\zeta_t^T) \, dt \\
&\quad + \int_0^{\cdot} \Delta_t^T \left(\eta(\zeta_t^T)' \, dB_t^T + \theta' \sigma(\zeta_t^T) \, dW_t^T \right) \\
&= 1 + \int_0^{\cdot} \Delta_t^T \left(\theta'(m - \text{div } c) + \nabla \cdot (c\theta) - a \right) (\zeta_t^T) \, dt \\
&\quad + \int_0^{\cdot} \Delta_t^T \left(\eta(\zeta_t^T)' \, dB_t^T + \theta(\zeta_t^T)' \, d\zeta_t^T \right)
\end{aligned}$$

for independent Brownian motions (W^T, B^T) in an appropriate filtration.

From the joint dynamics of (ζ^T, Δ^T), one obtains the joint dynamics of (ζ^T, χ^T), which again do not depend on T. In particular, since Δ^T is a semimartingale,

(3.9) yields that

$$\zeta^T = \zeta_0^T + \int_0^{\cdot} \left(c\frac{\nabla p}{p} + \operatorname{div} c - m \right)(\zeta_t^T)\, dt + \int_0^{\cdot} \sigma(\zeta_t^T)\, dW_t^T,$$

$$\chi^T = \chi_0^T + \int_0^{\cdot} \left(f(\zeta_t^T) - \chi_t^T \left(a - \theta' c\frac{\nabla p}{p} - \nabla \cdot (c\theta) \right)(\zeta_t^T) \right) dt$$

$$+ \int_0^{\cdot} \chi_t^T \left(\eta(\zeta_t^T)'\, dB_t^T + \theta' c(\zeta_t^T)'\, dW_t^T \right).$$

For a generic version (ζ, χ) with the same generator (which does not depend upon time) as (ζ^T, χ^T) above, ergodicity of Z implies ergodicity of ζ (see Proposition 8.1 below). Furthermore, χ is "mean reverting" as can easily be seen when $\theta \equiv 0$ and $a > 0$, and as continues to be true in the general case. Thus, one expects the empirical laws of (ζ, χ) to satisfy a certain strong law of large numbers, an intuition that is made precise in the following result.

Theorem 3.4 *If Assumption 2.7 holds, there exists a filtered probability space $(\Omega, \mathbf{F}, \mathbb{Q})$ supporting independent d- and k-dimensional Brownian motions W and B as well as a process ζ satisfying*

$$\zeta = \zeta_0 + \int_0^{\cdot} \left(c\frac{\nabla p}{p} + \operatorname{div} c - m \right)(\zeta_t)\, dt + \int_0^{\cdot} \sigma(\zeta_t)\, dW_t,$$

where ζ_0 is an $\mathcal{F}_0$-measurable random variable with density p.

Define the process Δ as the solution to the linear differential equation

$$\Delta = 1 + \int_0^{\cdot} \Delta_t \left(\theta'(m - \operatorname{div} c) + \nabla \cdot (c\theta) - a \right)(\zeta_t)\, dt$$

$$+ \int_0^{\cdot} \Delta_t \left(\eta(\zeta_t)'\, dB_t + \theta(\zeta_t)'\, d\zeta_t \right), \tag{3.10}$$

and then for any $x \in (0, \infty)$, define χ^x as the solution to the linear differential equation

$$\chi^x = x + \int_0^{\cdot} \chi_t^x \frac{d\Delta_t}{\Delta_t} + \int_0^{\cdot} f(\zeta_t)\, dt. \tag{3.11}$$

Finally, let $x \in (0, \infty)$, $T \in (0, \infty)$ and set $F = E \times (0, \infty)$ as in (3.1). Define the (random) empirical measure $\widehat{\pi}_T^x$ on $\mathcal{B}(F)$ by

$$\widehat{\pi}_T^x [A] := \frac{1}{T} \int_0^T \mathbb{I}_A(\zeta_t, \chi_t^x)\, dt, \quad A \in \mathcal{B}(F). \tag{3.12}$$

With the above notation, there exists a set $\Omega_0 \in \mathcal{F}_{\infty}$ with $\mathbb{Q}[\Omega_0] = 1$ such that

$$\lim_{T \to \infty} \widehat{\pi}_T^x(\omega) = \pi \text{ weakly}, \quad \text{for all } x \in (0, \infty) \text{ and } \omega \in \Omega_0, \tag{3.13}$$

where π is the joint distribution of (Z_0, X_0) under $\mathbb{P}$ given in (3.2).

Remark 3.5 In the context of Theorem 3.4, note that the processes Δ and χ^x can be given in closed form in terms of ζ; indeed,

$$\Delta = \exp\left(\int_0^{\cdot} \left(\theta'(m - \operatorname{div} c) + \nabla \cdot (c\theta) - a\right)(\zeta_t)\, \mathrm{d}t\right)$$

$$\times \mathcal{E}\left(\int_0^{\cdot} \left(\eta(\zeta_t)'\, \mathrm{d}B_t + \theta(\zeta_t)'\, \mathrm{d}\zeta_t\right)\right),$$

$$\chi^x = \Delta\left(x + \int_0^{\cdot} \frac{1}{\Delta_t} f(\zeta_t)\, \mathrm{d}t\right), \quad x \in (0, \infty).$$

Theorem 3.4 provides a way to estimate the joint distribution of (Z_0, X_0) efficiently through Monte Carlo simulation. Indeed, one need only obtain a single path of the reversed process (ζ, χ^x) to recover the distribution π. However, the applicability of the result above depends heavily on whether or not the distribution p for Z_0 is known, as it (together with its gradient) appears in the dynamics of ζ. In the case where Z is one-dimensional, or more generally reversing, p can be expressed in closed form from the model coefficients m and c in the dynamics for Z. Furthermore, there are certain cases of non-reversing, multidimensional diffusions where p can be (semi-)explicitly computed, as the next example shows.

Example 3.6 Assume that Z is a multidimensional Ornstein–Uhlenbeck process with dynamics

$$\mathrm{d}Z_t = -\gamma(Z_t - \Theta)\, \mathrm{d}t + \sigma\, \mathrm{d}W_t, \quad t \in \mathbb{R}_+,$$

where $\gamma \in \mathbb{R}^{d \times d}$, $\Theta \in \mathbb{R}^d$ and $\sigma \in \mathbb{R}^{d \times d}$. Here, $E = \mathbb{R}^d$ and (A1) clearly holds. Furthermore, (A2) is satisfied when $c = \sigma\sigma'$ is (strictly) positive definite; in fact, we take σ as the unique positive definite square root of c. The process Z need not be reversing, as can clearly be seen when σ is the identity matrix, $\Theta = 0$ and γ is not symmetric. However, as will be argued below, the ergodicity assumption (A3) holds when all eigenvalues of γ have strictly positive real parts, and one may identify the invariant density "almost" explicitly. To see this, a direct calculation shows that if a symmetric matrix J satisfies the Riccati equation

$$JJ = \sigma\gamma'\sigma^{-1}J + J\sigma^{-1}\gamma\sigma, \tag{3.14}$$

then the function

$$p(z) = \exp\left(-\frac{1}{2}(z - \Theta)'\sigma^{-1}J\sigma^{-1}(z - \Theta)\right), \quad z \in \mathbb{R}^d,$$

satisfies $\tilde{L}^Z p = 0$ where $\tilde{L}^Z$ is as in (2.1). If J is additionally positive definite, then up to a normalizing constant, p is the density for a normal random variable with mean Θ and covariance matrix $\Sigma = \sigma J^{-1}\sigma$. Thus p is integrable on $\mathbb{R}^d$ and (A3) follows from [27, Corollary 4.9.4], which proves recurrence for Z.

It thus remains to construct a symmetric, positive definite solution to (3.14). From [1, Lemma 2.4.1, Theorem 2.4.25], such a solution (called the "stabilizing solution"

therein) exists if (a) the pair $(\sigma^{-1}\gamma\sigma, 1_d)$ is *stabilizable*, in the sense that there exists a matrix F such that $\sigma^{-1}\gamma\sigma - F$ has eigenvalues with strictly negative real parts, and (b) the eigenvalues of $\sigma^{-1}\gamma\sigma$ have strictly positive real parts. In the present case, each of these statements readily follows. Indeed, for the first statement, one can take $F = \sigma^{-1}\gamma\sigma + 1_d$; for the second statement, note that the eigenvalues of $\sigma^{-1}\gamma\sigma$ coincide with those of γ, which by assumption have strictly positive real parts. Therefore, even in this non-reversing case, one may still identify p.

The previous interesting Example 3.6 notwithstanding, for non-reversing, multi-dimensional diffusions, even after verifying the ergodicity of Z (and hence the existence of p), one typically does not know p explicitly. In such cases, the following simulation method is proposed. Fix a large enough T and first simulate $(Z_t)_{t\in[0,2T]}$ via (1.2), starting from *any* point Z_0 (since the invariant density is unknown). If the choice of T is large enough, the process $(Z_t)_{t\in[T,2T]}$ will behave as the stationary version in (1.2), since Z_T will have approximately density p. In that case, defining $(\zeta_t)_{t\in[0,T]}$ via $\zeta_t = Z_{2T-t}$ for $t \in [0, T]$, ζ should behave as it should in the dynamics (8.6), even with ζ_0 having (approximate) density p. Now, given ζ, χ^x may be defined via the formulas of Remark 3.5; therefore, for large enough T, the empirical measure $\widehat{\pi}_T^x$ should approximate in the weak sense the joint law π.

Note finally that when p is known and $|\eta| > 0$, and under certain mixing conditions (see [29, 28]), one can also obtain uniform estimates for the speed at which the above convergence takes place.

Remark 3.7 In the case where $\theta = \eta \equiv 0$ and $f \in C^{1,\gamma}(E; \mathbb{R}_+)$, one can explicitly identify the support of π. Such an identification follows from more general ergodic results on "stochastic differential systems" obtained in [4, Sects. IIIA, IIIB]. To identify the support, note that when $\theta = \eta \equiv 0$, it follows that $\Delta_t = \exp(-\int_0^T a(\zeta_u)\,\mathrm{d}u)$. A direct calculation using Remark 3.5 shows that χ^x has dynamics

$$\mathrm{d}\chi_t^x = \left(f(\zeta_t) - \chi_t^x a(\zeta_t)\right)\mathrm{d}t. \tag{3.15}$$

Hence, the paths of χ^x are of bounded variation. Now define

$$\hat{u} := \inf\left\{x : \sup_{z\in E}\left(f(z) - xa(z)\right) \le 0\right\},$$

$$\hat{\ell} := \sup\left\{x : \inf_{z\in E}\left(f(z) - xa(z)\right) \ge 0\right\}.$$

Assumption 2.7 implies $a(z_0) > 0$ for some $z_0 \in E$ and thus $0 \le \hat{\ell} \le \hat{u} \le \infty$, with $\hat{\ell} = \hat{u}$ if and only if for some constant c, $f(z) = ca(z)$ for all $z \in E$. In this case, $X = c$ $\mathbb{P}^z$-almost surely for all $z \in E$. With this notation, [4] proves the following result.

Proposition 3.8 ([4, Sects. IIIA, IIIB]) *Let Assumption 2.7 hold. Assume also that $f \in C^{1,\gamma}(E; \mathbb{R}_+)$ and $\eta, \theta \equiv 0$. Then the support of π is $\bar{E}\times[\hat{\ell}, \hat{u}]$ ($[\hat{\ell}, \infty)$ if $\hat{u} = \infty$).*

4 A numerical example

We now provide an example which highlights the superiority (in terms of computational efficiency) of the time-reversal method over the naive method for obtaining the distribution of X_0. Consider the case $E = \mathbb{R}$ and

$$dZ_t = -\gamma Z_t \, dt + dW_t, \qquad X_0 = \int_0^\infty Z_t e^{-at} \, dt, \qquad \gamma, a > 0. \qquad (4.1)$$

Note that the function $\mathbb{R} \ni z \mapsto f(z) = z$ fails to be nonnegative. However, as argued below, the results of Theorem 3.4 still hold. As Z is a mean-reverting Ornstein–Uhlenbeck process, it is straightforward to verify Assumption (A3) with $p(z) = \sqrt{\gamma/\pi} \, e^{-\gamma z^2}$, so that $Z_0 \sim N(0, 1/(2\gamma))$. We claim that (Z_0, X_0) is normally distributed with mean vector $(0, 0)$ and covariance matrix

$$\Sigma = \begin{pmatrix} \frac{1}{2\gamma} & \frac{1}{2\gamma(a+\gamma)} \\ \frac{1}{2\gamma(a+\gamma)} & \frac{1}{2\gamma a(a+\gamma)} \end{pmatrix}.$$

Indeed, integration by parts shows that for $T > 0$,

$$\int_0^T e^{-at} Z_t \, dt = \frac{Z_0}{a+\gamma} + \frac{1}{a+\gamma} \int_0^T e^{-at} \, dW_t - \frac{1}{a+\gamma} e^{-aT} Z_T.$$

The ergodicity of Z implies that $\lim_{T\to\infty}(Z_T/T) = -\gamma \int_{\mathbb{R}} z p(z) \, dz = 0$ almost surely; therefore, it follows that $\lim_{T\to\infty} e^{-aT} Z_T = 0$ holds almost surely. Next, note that $Y_T := \int_0^T e^{-at} \, dW_t$ is independent of Z_0 and normally distributed with mean 0 and variance $(1 - e^{-2aT})/(2a)$. Finally, as a process, $Y = (Y_T)_{T\geq 0}$ is an L^2-bounded martingale and hence $Y_\infty := \lim_{T\to\infty} Y_T$ almost surely exists, where Y_∞ is independent of Z_0 and normally distributed with mean 0 and variance $1/(2a)$. Thus $X_0 = \lim_{T\to\infty} \int_0^T e^{-at} Z_t \, dt$ exists almost surely and

$$X_0 = \frac{Z_0}{a+\gamma} + \frac{Y_\infty}{a+\gamma}, \quad Z_0 \perp\!\!\!\perp Y_\infty, \quad Z_0 \sim N\left(0, \frac{1}{2\gamma}\right), \quad Y_\infty \sim N\left(0, \frac{1}{2a}\right),$$

from which the joint distribution follows. Now, even though $f(z) = z$ can take negative values, the time-reversal dynamics in (3.15) still hold, taking the form

$$d\zeta_t = -\gamma \zeta_t \, dt + dW_t, \qquad d\chi_t = (a - \zeta_t \chi_t) \, dt.$$

Lastly, even though Theorem 3.4 no longer directly applies, it is shown in [4, Theorem 3.3, Sect. 3.D, Proposition 3.15] that (ζ, χ) is still ergodic,[5] in that (3.13) holds.

For these dynamics, we performed the following test. For a fixed terminal time T and mesh size δ, we estimated the distribution of X_0 in two ways. First ("Method A") by sampling $\zeta_0 \sim p$ and setting $\chi_0 = 1$, and second ("Method B") by running the forward process Z until $2T$ and then setting $\zeta_t = Z_{2T-t}$, $\chi_0 = 1$. For each simulation,

---

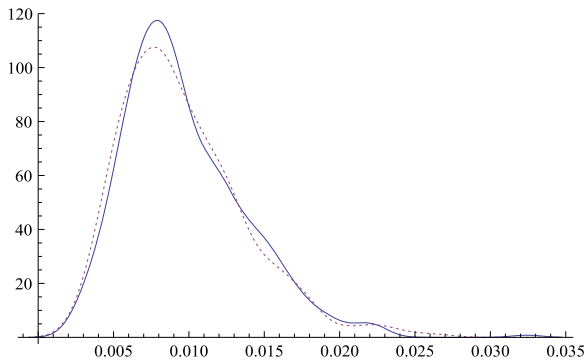[5]The tightness condition in [4, Proposition 3.15] is straightforward to verify.

Fig. 1 Kolmogorov–Smirnov distances between the empirical and the true distribution for X_0. The *solid line* is for the time-reversal method starting with $\zeta_0 \sim p$, and the *dashed line* for the time-reversal method running Z_0 to $2T$ and setting $\zeta_t = Z_{2T-t}$. Here, $T = 10'000$, $\delta = 1/24$, $\gamma = 2$ and $a = 1$. Computations were performed using *Mathematica* and the code can be found on the second author's website www.math.cmu.edu/users/scottrob/research

Table 1 Statistics on Kolmogorov–Smirnov distances between the empirical and the true distribution for X_0 using Methods A and B		Method A	Method B
	median distance	0.00887	0.00882
	standard deviation	0.00405	0.00413
	99th percentile	0.02168	0.02255
	1st percentile	0.00405	0.00290
	median time (seconds)	2.694	8.766

we computed the empirical distribution along a single path and then estimated the Kolmogorov–Smirnov distance ($d_{KS}(F, G) = \sup_x |F(x) - G(x)|$, for distribution functions F, G) between the empirical and the true distribution for X_0. The parameter values were $\gamma = 2$, $a = 1$, $T = 10'000$ and $\delta = 1/24$.

Figure 1 shows the resulting Kolmogorov–Smirnov distances for 500 sample paths. The plot gives a (smoothed) histogram comparing the distances using the two methods described above. As can be seen, the two methods give comparable results; this is not surprising given the rapid convergence of the distribution of ζ to its invariant distribution [8]. Table 1 provides summary statistics regarding the median distances and simulation times, as well as the standard deviation and tail data.

Having obtained Kolmogorov–Smirnov distances using time-reversal methods, we next compared our results to a naive simulation of X_0, obtained by sampling $Z_0 \sim p$ and computing X_0 directly via (4.1). Here, for the median distance d using Method A from Table 1, we sampled X_0 stopping at the first instance N so that the Kolmogorov–Smirnov distance between the empirical and the true distribution for X_0 fell below d. As can be seen from Fig. 2 and the summary statistics in Table 2, the naive simulation performs significantly worse; at the median, it took 7'002 paths and a simulation time of 8.66 minutes to achieve the same level of accuracy as 1 path (2.94 seconds) of the time-reversed process. Further, the histogram shows the presence of a significant
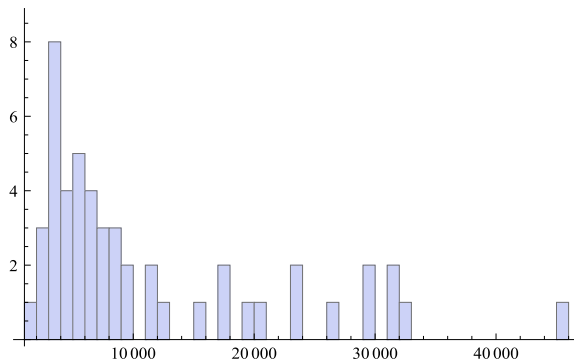
Fig. 2 Histogram for the number N of paths necessary so that using the naive simulation for X_0, the Kolmogorov–Smirnov distance between the empirical distribution and the true distribution for X_0 fell below the median distance d using Method A from Table 1. The integral was computed using $T = 100$ with mesh size of $\delta = 1/24$; furthermore, the values $\gamma = 2$ and $a = 1$ we used. Computations were performed using *Mathematica* and the code can be found on the second author's website www.math.cmu.edu/users/scottrob/research

Table 2 Summary statistics using the naive forward simulation method

Summary for the forward simulation	
median number of paths	7'002
mean number of paths	11'446
standard deviation	10'165
minimum number of paths	1'846
maximum number of paths	45'004
median simulation time (minutes)	8.66
mean simulation time (minutes)	14.34

number of trials where significantly more than the median number of paths were needed to achieve the given accuracy.

5 Conclusion

In this work, using the method of time reversal, an efficient method for simulating the joint distribution of (Z_0, X_0) for perpetuities of the form (1.1) is obtained. The joint distribution may be obtained by sampling a single path of the reversed process, as opposed to sampling numerous paths of X_0 using the naive method. However, the effectiveness of the proposed method depends on being able to obtain analytic representations for the distribution p of Z_0—an undertaking that, though always possible in the one-dimensional case, is often not possible for non-reversing multidimensional diffusions. Furthermore, results are presented for perpetuities with nonnegative underlying cash flow rates. More research is needed to identify an effective time-reversal

method for perpetuities of the form

$$X_0 = \int_0^\infty D_t \, dF_t$$

for general Markovian processes F (i.e., not just $dF_t = f(Z_t) \, dt$) containing both jumps and diffusive terms. Additionally, the performance of the method where Z is run until a large time $2T$ and then setting $\zeta_t = Z_{2T-t}$ must be thoroughly analyzed; in particular, how fast does the distribution of Z_{2T} converge to p given a fixed starting point? To answer these questions, one must first analyze the resultant backward dynamics and associated PDEs for the invariant density, obtaining rates of convergence.

6 Proofs for Sect. 2.2

We present here the proofs of Lemmas 2.3 and 2.5.

Proof of Lemma 2.3 Let $\varepsilon > 0$ be as in (2.3). We first treat the case $\theta' c\theta + \eta'\eta \equiv 0$. Then $R = \int_0^\cdot a(Z_t) \, dt$ and (2.3) specifies to $a^- \in \mathbb{L}^1(E, p)$ and $\int_E a(z) p(z) \, dz > 0$. Set $\kappa := (1/4) \int_E a(z) p(z) \, dz > 0$. Fix $z \in E$ and denote by $\mathbb{P}^z$ the probability obtained by conditioning upon $Z_0 = z$. The positive recurrence of Z implies ([27, Theorem 4.9.5]) that there exists a $\mathbb{P}^z$-a.s. finite random variable $T(z)$ such that $t \geq T(z)$ implies that $R_t = \int_0^t a(Z_u) \, du \geq 2\kappa t$, and hence the first conclusion of Lemma 2.3 holds. Furthermore, since Z is stationary and ergodic under $\mathbb{P}$, the ergodic theorem implies there is a $\mathbb{P}$-a.s. finite random variable T such that $t \geq T$ implies $R_t \geq 2\kappa t$. Now, let $n \in \mathbb{N}$ be such that $n > 1/(2\kappa)$. We have

$$\sup_{t \geq 0} (t/n - R_t) \leq \sup_{t \leq T} (t/n - R_t) < \infty,$$

where the last inequality follows by the regularity of a and the non-explosivity of Z. Thus

$$X_0 = \int_0^\infty e^{-R_t} f(Z_t) \, dt \leq e^{\sup_{t \leq T} (t/n - R_t)} \int_0^\infty e^{-t/n} f(Z_t) \, dt.$$

By the stationarity of Z,

$$\mathbb{E}\left[\int_0^\infty e^{-t/n} f(Z_t) \, dt\right] = \int_0^\infty e^{-t/n} \mathbb{E}[f(Z_t)] \, dt = n \int_E f(z) p(z) \, dz < \infty,$$

hence $\mathbb{P}[\int_0^\infty e^{-t/n} f(Z_t) \, dt < \infty] = 1$, which in turn implies that $\mathbb{P}[X_0 < \infty] = 1$.

Assume now that $\theta' c\theta + \eta'\eta \not\equiv 0$, which by continuity of all involved functions implies that $\int_E (\theta' c\theta + \eta'\eta)(z) p(z) \, dz > 0$. Fix $z \in E$. Positive recurrence of Z gives that $\lim_{t \to \infty} \int_0^t (\theta' c\theta + \eta'\eta)(Z_u) \, du = \infty$ with $\mathbb{P}^z$-probability one. On the event

$\{\int_0^t (\theta' c\theta + \eta'\eta)(Z_u)\,\mathrm{d}u > 0\}$, note that

$$-R_t = -\int_0^t a(Z_u)\,\mathrm{d}u$$
$$+ \int_0^t (\theta' c\theta + \eta'\eta)(Z_u)\,\mathrm{d}u\left(-\frac{1}{2} - \frac{\int_0^t \theta'\sigma(Z_u)\,\mathrm{d}W_u + \eta(Z_u)\,\mathrm{d}B_u}{\int_0^t (\theta' c\theta + \eta'\eta)(Z_u)\,\mathrm{d}u}\right).$$

By the Dambis/Dubins/Schwarz theorem and the strong law of large numbers for Brownian motion, it follows that there exists a $\mathbb{P}^z$-a.s. finite random variable $T(z)$ such that

$$t \geq T(z) \quad \Longrightarrow \quad -\frac{\int_0^t \theta'\sigma(Z_u)\,\mathrm{d}W_u + \eta(Z_u)\,\mathrm{d}B_u}{\int_0^t (\theta' c\theta + \eta'\eta)(Z_u)\,\mathrm{d}u} \leq \frac{\varepsilon}{2};$$

therefore,

$$t \geq T(z) \quad \Longrightarrow \quad -R_t \leq -\int_0^t \left(a + \frac{1-\varepsilon}{2}(\theta' c\theta + \eta'\eta)\right)(Z_u)\,\mathrm{d}u.$$

With $\kappa := (1/4)\int_E (a + (1-\varepsilon)(\theta' c\theta + \eta'\eta)/2)(z)p(z)\,\mathrm{d}z > 0$, and increasing $T(z)$ if necessary (still keeping it $\mathbb{P}^z$-a.s. finite), it follows that $t \geq T(z)$ implies $-R_t \leq -2\kappa t$. Hence the first part of Lemma 2.3 holds true again. Additionally, the ergodic theorem applied with $\mathbb{P}$ gives a $\mathbb{P}$-a.s. finite random variable T such that $t \geq T$ implies $-R_t \leq -2\kappa t$. Again, for $n \in \mathbb{N}$ such that $n > 1/(2\kappa)$, we have

$$X_0 = \int_0^\infty e^{-R_t} f(Z_t)\,\mathrm{d}t \leq e^{\sup_{t \leq T}(t/n - R_t)} \int_0^\infty e^{-t/n} f(Z_t)\,\mathrm{d}t,$$

from which $\mathbb{P}[X_0 < \infty] = 1$ follows by the same line of reasoning as above. $\qquad\square$

Proof of Lemma 2.5 The proof is nearly identical to that of Lemma 2.3. Namely, in each of the cases $\theta' c\theta + \eta'\eta \equiv 0$ and $\theta' c\theta + \eta'\eta \not\equiv 0$, there are under the given hypothesis a constant $\kappa \geq 0$ and a $\mathbb{P}$-a.s. finite random variable T such that $-R_t \geq \kappa t$ holds for $t \geq T$. This gives that

$$X_0 \geq \int_T^\infty e^{\kappa t} f(Z_t)\,\mathrm{d}t \geq \int_T^\infty e^{\kappa t}(f \wedge N)(Z_t)\,\mathrm{d}t,$$

where N is large enough so that $\int_E (f(z) \wedge N)p(z)\,\mathrm{d}z > 0$. We thus have

$$X_0 \geq \int_0^\infty e^{\kappa t}(f \wedge N)(Z_t)\,\mathrm{d}t - \frac{N}{\kappa}(e^{\kappa T} - 1).$$

Ergodicity of Z implies that $\mathbb{P}$-almost surely,

$$\lim_{u \to \infty} \frac{1}{u}\int_0^u (f \wedge N)(Z_t)\,\mathrm{d}t = \int_E \left(f(z) \wedge N\right)p(z)\,\mathrm{d}z > 0,$$

so that $\lim_{u\to\infty} \int_0^u e^{\kappa t}(f \wedge N)(Z_t)\,\mathrm{d}t = \infty$, proving the result. $\qquad\square$

7 Proof of Theorem 3.1

Under the given assumptions, there exists a unique solution $(\mathbb{P}^{z,x})_{(z,x)\in F}$ to the generalized martingale problem for L on F, where L is from (3.4). Here, the measure space is $(\widetilde{\Omega}, \widetilde{\mathbf{F}})$, where $\widetilde{\Omega} = (C[0,\infty); \hat{F})$, with $\hat{F}$ being the one-point compactification of F. The filtration $\widetilde{\mathbf{F}}$ is the right-continuous enlargement of the filtration generated by the coordinate process $(\widetilde{Z}, \widetilde{Y})$ on $\widetilde{\Omega}$.

Let $(F_n)_{n\in\mathbb{N}}$ be an increasing sequence of smooth, bounded, open, connected domains of F such that $F = \bigcup_n F_n$. Note that F_n can be obtained by smoothing out the boundary of $E_n \times (1/n, n)$. By uniqueness of solutions to the generalized martingale problem, for each n, the law of $(\widetilde{Z}, \widetilde{Y})$ is the same as the law of (Z, Y^x) under $\mathbb{P}[\cdot \mid Z_0 = z]$ (where the latter will always denote a version of the conditional probability) up until the first exit time of F_n. Furthermore, since the process Z is recurrent, with $(\mathbb{P}^z)_{z\in E}$ being the restriction of $(\mathbb{P}^{z,x})_{(z,x)\in F}$ to the first d coordinates, for $z \in E$, the law of $\widetilde{Z}$ under $\mathbb{P}^z$ is the same as the law of Z under $\mathbb{P}[\cdot \mid Z_0 = z]$. For these reasons, and in order to ease the reading, we abuse notation and still use (Z, Y) instead of $(\widetilde{Z}, \widetilde{Y})$ for the coordinate process on $\widetilde{\Omega}$. The underlying space we are working on will be clear from the context.

Denote by τ_n the first exit time of (Z, Y) from F_n. Assumption 2.7 implies that Z does not explode under $\mathbb{P}^{z,x}$, and Y cannot explode to infinity since D is strictly positive almost surely under $\mathbb{P}[\cdot \mid Z_0 = z]$ for all $z \in E$. Therefore, the explosion time $\tau := \lim_{n\to\infty} \tau_n$ for (Z, Y) is the first hitting time of Y to 0, and the law of τ under $\mathbb{P}^{z,x}$ is the same as the law of the first hitting of Y^x to 0 under $\mathbb{P}[\cdot \mid Z_0 = z]$.

Note that $Y_t^x = D_t^{-1}(x - X_0 + \int_t^\infty D_u f(Z_u)\,du)$. Assumption 2.7 implies[6]

$$\mathbb{P}\left[\int_t^\infty D_u f(Z_u)\,du > 0 \;\middle|\; Z_0 = z\right] = 1, \qquad z \in E, t \geq 0. \tag{7.1}$$

Therefore,

$$g(z,x) = \mathbb{P}[X_0 \leq x \mid Z_0 = z] = \mathbb{P}^{z,x}[Y_t^x > 0, \;\forall t \geq 0] = \mathbb{P}^{z,x}[\tau = \infty].$$

Define

$$h(z,x) := \mathbb{P}^{z,x}\left[\lim_{t\to\infty} Y_t = \infty\right], \quad (z,x) \in F. \tag{7.2}$$

Fix $(z,x) \in F$ and let $0 < \varepsilon < x$. Note that $Y_t^x = Y_t^{x-\varepsilon} + \varepsilon/D_t$. Since $\lim_{t\to\infty} D_t = 0$ holds $\mathbb{P}[\cdot \mid Z_0 = z]$-a.s., it follows that

$$\mathbb{P}^{z,x-\varepsilon}[\tau = \infty] = \mathbb{P}[Y_t^{x-\varepsilon} > 0, \;\forall t \geq 0 \mid Z_0 = z]$$

$$\leq \mathbb{P}[Y_t^x \geq \varepsilon/D_t, \;\forall t \geq 0 \mid Z_0 = z]$$

$$\leq \mathbb{P}\left[\lim_{t\to\infty} Y_t^x = \infty \;\middle|\; Z_0 = z\right]$$

$$= \mathbb{P}^{z,x}\left[\lim_{t\to\infty} Y_t = \infty\right] \leq \mathbb{P}^{z,x}[\tau = \infty]. \tag{7.3}$$

---

[6]This follows from the ergodic theorem by using that we have the set inclusion $\{\int_t^\infty f(Z_u)D_u\,du = 0\} \subset \{\lim_{k\to\infty}(1/k)\int_t^{t+k} f(Z_u)\,du = 0\}$.

Therefore, $g(z, x - \varepsilon) \leq h(z, x) \leq g(z, x)$. By definition, $g(z, x)$ is right-continuous in x for a fixed z, and so

$$g(z, x) \leq \liminf_{\varepsilon \to 0} h(z, x + \varepsilon) \leq \limsup_{\varepsilon \to 0} h(z, x + \varepsilon) \leq \limsup_{\varepsilon \to 0} g(z, x + \varepsilon) = g(z, x).$$

Therefore, if $h(z, x)$ is continuous, it follows that $h(z, x) = g(z, x)$. We now show that in fact h is in $C^{2,\gamma}(F)$ and satisfies $Lh = 0$. This gives the desired result for g since $g = h$.

Let $\psi : (0, \infty) \to (0, 1)$ be a smooth function satisfying $\lim_{x \to 0} \psi(x) = 0$ and $\lim_{x \to \infty} \psi(x) = 1$. By the classical Feynman–Kac formula,

$$u^n(z, x) := \mathbb{E}^{\mathbb{P}^{z,x}}[\psi(Y_{\tau_n})]$$

satisfies $Lu^n = 0$ in F_n with $u^n(z, x) = \psi(x)$ on ∂F_n. As $\mathbb{P}[X_0 < \infty \mid Z_0 = z] = 1$, there exists a pair $(z_0, x_0) \in F$ with $\mathbb{P}[X_0 < x_0 \mid Z_0 = z_0] > 0$. Using (7.3), this gives

$$h(z_0, x_0) \geq \mathbb{P}[X_0 < x_0 \mid Z_0 = z_0] > 0. \tag{7.4}$$

Therefore, $(\mathbb{P}^{z,x})_{(z,x) \in F}$ is transient [27, Chap. 2], and since $(\mathbb{P}^z)_{z \in E}$ is positive recurrent, this implies that for all (z, x), with $\mathbb{P}^{z,x}$-probability one, either $\lim_{t \to \tau} Y_t = 0$ or $\lim_{t \to \tau} Y_t = \infty$, where in the latter case, $\tau = \infty$ since Y cannot explode to ∞. This in turn yields that $Y_{\tau_n} \to 0$ or $Y_{\tau_n} \to \infty$ with $\mathbb{P}^{z,x}$-probability one and hence by the dominated convergence theorem,

$$\lim_{n \to \infty} u^n(z, x) = \mathbb{P}^{z,x}\left[\lim_{t \to \tau} Y_t = \infty\right] = \mathbb{P}^{z,x}\left[\lim_{t \to \infty} Y_t = \infty\right] = h(z, x). \tag{7.5}$$

For (z_0, x_0) from (7.4), $g(z_0, x_0) \geq h(z_0, x_0) > 0$ and hence $g(z, x) > 0$ for all $(z, x) \in F$ [27, Theorem 1.15.1]. But this implies $h(z, x) \geq g(z, x/2) > 0$, and so from (7.5), the u^n are converging pointwise to a strictly positive function. Thus, by the interior Schauder estimates and Harnack's inequality, it follows by "the standard compactness argument" (see [27, p. 147]) that there exists a strictly positive function u in $C^{2,\gamma}(F)$ such that u^n converges to u in the $C^{2,\gamma}(D)$-Hölder space for all compact $D \subset F$. Clearly, this function u satisfies $Lu = 0$ in F. In fact, since u^n converges to h pointwise, $h = u$ and hence $Lh = 0$.

We now consider the boundary conditions for g. Let the integer k be given. It suffices to show that for each $\varepsilon > 0$, there is some $n(\varepsilon)$ such that

$$\sup_{x \leq n(\varepsilon)^{-1}, z \in E_k} g(z, x) \leq \varepsilon, \qquad \inf_{x \geq n(\varepsilon), z \in E_k} g(z, x) \geq 1 - \varepsilon.$$

The condition near $x = 0$ is handled first. By way of contradiction, assume there exists some $\varepsilon > 0$ such that for all integers n, there exist $z_n \in E_k$, $x_n \leq 1/n$ such that $g(z_n, x_n) > \varepsilon$. Since the z_n are all contained within E_k, there is a subsequence (still labeled n) such that $z_n \to z$ for $z \in \bar{E}_k$. Let $\delta > 0$ and choose N_δ such that $n \geq N_\delta$ implies $n^{-1} \leq \delta$. Since g is increasing in x, $\varepsilon < g(z_n, \delta)$. Since g is continuous, $\varepsilon \leq g(z, \delta)$. Since this is true for all $\delta > 0$, $\lim_{x \to 0} g(z, x) \geq \varepsilon$. But this is a contradiction because $\lim_{x \to 0} g(z, x) = 0$ for each $z \in E$. To see this, let $\delta > 0$ and choose

$\beta > 0$ such that $\mathbb{P}[X_0 \geq \beta \mid Z_0 = z] \geq 1 - \delta$. This is possible in view of (7.1). Thus, for $x < \beta$, $g(z, x) \leq \mathbb{P}[X_0 < \beta \mid Z_0 = z] \leq \delta$, and hence $\lim\sup_{x \to 0} g(z, x) \leq \delta$. Taking $\delta \to 0$ gives the result.

The proof for $x \to \infty$ is very similar. Assume by contradiction that there is some $\varepsilon > 0$ such that for all integers n, there exist $z_n \in E_k$, $x_n \geq n$ such that $g(z_n, x_n) < 1 - \varepsilon$. Again, by taking subsequences, we can assume $z_n \to z \in \bar{E}_k$. Fix $M > 0$. For $n \geq M$, since g is increasing in x, $g(z_n, M) < 1 - \varepsilon$. Since g is continuous, $g(z, M) \leq 1 - \varepsilon$. Since this holds for all M, $\lim_{x \to \infty} g(z, x) \leq 1 - \varepsilon$. But this violates the condition that under $\mathbb{P}[\cdot \mid Z_0 = z]$, $X_0 < \infty$ almost surely.

The uniqueness claim is now proved. Let $\tilde{g}$ be a $C^2(F)$ solution of $L\tilde{g} = 0$ such that $0 \leq \tilde{g} \leq 1$ and such that (3.5) holds. Define the stopping times

$$\sigma_k := \inf\{t \geq 0 : Z_t \notin E_k\}, \qquad \rho_k := \inf\{t \geq 0 : Y_t = k\}.$$

By Itô's formula, for any k, n, m,

$$\tilde{g}(z, x) = \mathbb{E}^{\mathbb{P}^{z,x}}\left[g\left(Z_{\sigma_k \wedge \rho_{1/n} \wedge \rho_m}, Y_{\sigma_k \wedge \rho_{1/n} \wedge \rho_m}\right) \right.$$
$$\left. \times \left(1_{\{\rho_{1/n} < \sigma_k \wedge \rho_m\}} + 1_{\{\rho_{1/n} \geq \sigma_k \wedge \rho_m\}}(1_{\{\tau < \infty\}} + 1_{\{\tau = \infty\}})\right)\right].$$

Since $\lim_{m \to \infty} \rho_m = \infty$ $\mathbb{P}^{z,x}$-almost surely, taking $m \to \infty$ yields

$$\tilde{g}(z, x) = \hat{\mathbb{E}}^{\mathbb{P}^{z,x}}\left[g\left(Z_{\sigma_k \wedge \rho_{1/n}}, Y_{\sigma_k \wedge \rho_{1/n}}\right)\left(1_{\{\rho_{1/n} < \sigma_k\}} + 1_{\{\rho_{1/n} \geq \sigma_k\}}(1_{\{\tau < \infty\}} + 1_{\{\tau = \infty\}})\right)\right].$$

On $\{\rho_{1/n} < \sigma_k\}$, we have $Z_{\rho_{1/n}} \in E_k$, $Y_{\rho_{1/n}} \leq 1/n$ and hence by $0 \leq \tilde{g} \leq 1$ and (3.5), for any $\varepsilon > 0$, there is an $n(\varepsilon)$ such that for $n \geq n(\varepsilon)$,

$$\tilde{g}(z, x) \leq \varepsilon + \mathbb{P}^{z,x}[\rho_{1/n} \geq \sigma_k, \tau < \infty] + \mathbb{P}^{z,x}[\rho_{1/n} \geq \sigma_k, \tau = \infty].$$

Taking $n \to \infty$ thus gives

$$\tilde{g}(z, x) \leq \varepsilon + \mathbb{P}^{z,x}[\tau \geq \sigma_k, \tau < \infty] + \mathbb{P}^{z,x}[\tau = \infty].$$

Taking $k \to \infty$ gives

$$\tilde{g}(z, x) \leq \varepsilon + \mathbb{P}^{z,x}[\tau = \infty],$$

and hence taking $\varepsilon \to 0$ gives $\tilde{g}(z, x) \leq \mathbb{P}^{z,x}[\tau = \infty] = g(z, x)$. Similarly, for k, n, m,

$$\tilde{g}(z, x) = \mathbb{E}^{\mathbb{P}^{z,x}}\left[g\left(Z_{\sigma_k \wedge \rho_{1/n} \wedge \rho_m}, Y_{\sigma_k \wedge \rho_{1/n} \wedge \rho_m}\right)(1_{\{\rho_m < \sigma_k \wedge \rho_{1/n}\}} + 1_{\{\rho_m \geq \sigma_k \wedge \rho_{1/n}\}})\right],$$
$$\geq (1 - \varepsilon)\hat{\mathbb{P}}^{z,x}\left[\rho_m < \sigma_k \wedge \rho_{1/n}, \lim_{t \to \infty} Y_t = \infty\right],$$

for all $\varepsilon > 0$ and $m \geq m(\varepsilon)$ for some $m(\varepsilon)$. Note that the set $\{\rho_m < \sigma_k \wedge \rho_{1/n}\}$ is restricted to include $\{\lim_{t \to \infty} Y_t = \infty\}$, but this is fine since lower bounds are considered. Now, on the event $\{\lim_{t \to \infty} Y_t = \infty\}$, it holds that $\rho_{1/n} \to \infty$. Thus, taking

$n \to \infty$,

$$\tilde{g}(z, x) \geq (1 - \varepsilon)\mathbb{P}^{z,x}\left[\rho_m < \sigma_k, \lim_{t\to\infty} Y_t = \infty\right].$$

Taking $k \to \infty$ gives

$$\tilde{g}(z, x) \geq (1 - \varepsilon)\mathbb{P}^{z,x}\left[\rho_m < \infty, \lim_{t\to\infty} Y_t = \infty\right].$$

Taking $m \to \infty$ and noting that for m large enough, $\rho_m < \infty$ on $\{\lim_{t\to\infty} Y_t = \infty\}$, it holds that

$$\tilde{g}(z, x) \geq (1 - \varepsilon)\mathbb{P}^{z,x}\left[\lim_{t\to\infty} Y_t = \infty\right] = (1 - \varepsilon)h(z, x),$$

where the last equality follows by the definition of h in (7.2). Now, in proving $Lg = 0$ it was shown that $g = h$ and hence $\tilde{g}(z, x) \geq (1 - \varepsilon)g(z, x)$. Taking $\varepsilon \to 0$ gives that $\tilde{g}(z, x) \geq g(z, x)$, finishing the proof. $\qquad\square$

8 Dynamics for the time-reversed process

The goal of the next two sections is to prove Theorem 3.4. We keep all notation from Sect. 3.2. We first identify the dynamics for ζ^T.

Proposition 8.1 *Suppose that Assumption 2.7 holds. Then for each $T > 0$, the law of ζ^T under $\mathbb{P}$ solves the martingale problem on E (for $t \leq T$) for the operator $L^\zeta := (1/2)c^{ij}\partial_{ij}^2 + \mu^i\partial_i$, where*

$$\mu := c\frac{\nabla p}{p} + \operatorname{div} c - m. \tag{8.1}$$

The operator L^ζ does not depend upon T. Thus, if $(\mathbb{Q}^z)_{z\in E}$ denotes the solution of the generalized martingale problem for L^ζ on E, then in fact $(\mathbb{Q}^z)_{\zeta\in E}$ solves the martingale problem for L^ζ on E and is positive recurrent.

Remark 8.2 If Z is reversing, then p satisfies $m = (1/2)(c\nabla p/p + \operatorname{div} c)$. Thus, in this instance, $\mu = m$ and as the name suggests, ζ^T has the same dynamics as Z.

Proof of Proposition 8.1 The first statement regarding the martingale problem is based on the argument in [19]. Since Z is positive recurrent with invariant measure p and Z_0 has initial distribution p under $\mathbb{P}$, Z is stationary with distribution p. Since $\tilde{L}^Z p = 0$, Eq. (2.5) in [19] holds, noting that p does not depend upon t.

For $0 \leq s \leq t$ and $g \in C_c^\infty(E)$, define the function $v(s, z) := \mathbb{E}[g(X_t) \,|\, Z_s = z]$. The Feynman–Kac formula implies that v satisfies $v_s + L^z v = 0$ for $0 < s < t$, $z \in E$, with $v(t, z) = g(z)$; see [20, 18] for an extension of the classical Feynman–Kac formula to the current setup. Therefore, the condition in Eq. (2.7) of [19] holds as well. Thus, the formal argument on p. 1191 of [19] is rigorous, and the law of ζ^T under $\mathbb{P}$ solves the martingale problem for L^ζ.

Turning to the statement regarding $(\mathbb{Q}^z)_{z \in E}$, set $\tilde{L}^\zeta$ as the formal adjoint to L^ζ. $\tilde{L}^\zeta$ is given by (2.1) with μ replacing m therein. Using the formula for μ in (8.1) and for $\tilde{L}^Z$ in (2.1), calculation shows that

$$\tilde{L}^\zeta f = \tilde{L}^Z f - 2\nabla \cdot \left(\frac{f}{p} \left(\frac{1}{2}(c\nabla p + p \operatorname{div} c) - pm \right) \right).$$

Since

$$0 = \tilde{L}^Z p = \nabla \cdot \left(\frac{1}{2}(c\nabla p + p \operatorname{div} c) - pm \right), \qquad (8.2)$$

it follows by considering $f = p$ above that $\tilde{L}^\zeta p = 0$. Therefore, p is an invariant density for L^ζ if and only if the diffusion corresponding to the operator $\tilde{L}^{\zeta,p}$ does not explode, where $\tilde{L}^{\zeta,p}$ is the h-transform of $\tilde{L}^\zeta$ [27, Theorem 4.8.5]. But by the definition of the h-transform [27, Sect. 4.1] and (2.1) with μ replacing m,

$$\tilde{L}^{\zeta,p} f := \frac{1}{p} \tilde{L}^\zeta (fp) = \frac{1}{2} c^{ij} \partial_{ij}^2 f - \left(\mu^i - \operatorname{div} c^i - \left(c \frac{\nabla p}{p} \right)^i \right) \partial_i f + \frac{f}{p} \tilde{L}^\zeta p$$

$$= \frac{1}{2} c^{ij} \partial_{ij}^2 f + m^i \partial_i f = L^Z f,$$

where the third equality follows from (8.1). Thus Assumption 2.7 (specifically the fact that Z is ergodic and $\int_E p(z)\,dz = 1$) implies that the diffusion for $\tilde{L}^{\zeta,p}$ not only does not explode, but is also positive recurrent, finishing the proof. $\qquad\square$

In preparation for the proof of the main result of this section, which is Proposition 8.5, we first need to define a certain "backward" filtration $\mathbf{G}^T$ and present two lemmas. Fix $T \in (0,\infty)$, $t \in [0,T]$ and let $\tilde{\mathcal{G}}_t^T$ be the σ-field generated by X_T, $(Z_{T-u})_{u\in[0,t]}$, $(W_T - W_{T-u})_{u\in[0,t]}$ and $(B_T - B_{T-u})_{u\in[0,t]}$. Then denote by $\mathbf{G}^T := (\mathcal{G}_t^T)_{t\in[0,T]}$ the usual augmentation of $(\tilde{\mathcal{G}}_t^T)_{t\in[0,T]}$. It is easy to check that (χ^T, ζ^T) is $\mathbf{G}^T$-adapted for all $T \in \mathbb{R}_+$, as well as that the process B^T defined via $B_t^T := B_{T-t} - B_T$ is a k-dimensional Brownian motion on $(\Omega, \mathbf{G}^T, \mathbb{P})$, independent of $(\chi_0^T, \zeta_0^T) = (X_T, Z_T)$. However, the $\mathbf{G}^T$-adapted process $(W_{T-t} - W_T)_{t\in[0,T]}$ is not necessarily a Brownian motion on $(\Omega, \mathbf{G}^T, \mathbb{P})$.

With this notation, the following two lemmas are essential for proving Proposition 8.5.

Lemma 8.3 *If Assumption 2.7 holds, then for any locally bounded Borel function $\eta : E \to \mathbb{R}$ and $0 \le s \le t \le T$, it holds that*

$$-\int_{T-t}^{T-s} \eta(Z_u)'\,dB_u = \int_s^t \eta(\zeta_u^T)'\,dB_u^T. \qquad (8.3)$$

Furthermore, if $\theta : E \to \mathbb{R}^d$ is continuously differentiable, then

$$-\int_{T-t}^{T-s} \theta'(Z_u)\,dZ_u = \int_s^t \theta'(\zeta_u^T)\,d\zeta_u^T + \int_s^t \left(\nabla \cdot (c\theta) - \theta'\operatorname{div} c \right)(\zeta_u^T)\,du. \qquad (8.4)$$

Proof Fix $0 \leq s \leq t \leq T$. For each $n \in \mathbb{N}$ and $i \in \{0, \ldots, n\}$, let

$$u_i^n := T - t + i(t - s)/n.$$

First, assume that η is twice continuously differentiable. The standard convergence theorem for stochastic integrals implies that (the following limit is to be understood in measure for $\mathbb{P}$)

$$\int_s^t \eta(\zeta_u^T)' \, \mathrm{d}B_u^T + \int_{T-t}^{T-s} \eta(Z_u)' \, \mathrm{d}B_u$$
$$= -\lim_{n \to \infty} \left(\sum_{i=1}^n \left(\eta(Z_{u_i^n}) - \eta(Z_{u_{i-1}^n}) \right)' (B_{u_i^n} - B_{u_{i-1}^n}) \right).$$

Since B and Z are independent, Itô's formula implies that the last quadratic covariation is zero. Therefore, (8.3) holds for twice continuously differentiable η. The fact that (8.3) holds whenever η is locally bounded follows from a monotone class argument.

In a similar manner, assume that θ is twice continuously differentiable. The standard convergence theorem for stochastic integrals implies that

$$\int_s^t \theta'(\zeta_u^T) \, \mathrm{d}\zeta_u^T + \int_{T-t}^{T-s} \theta'(Z_u) \, \mathrm{d}Z_u$$
$$= -\lim_{n \to \infty} \left(\sum_{i=1}^n \left(\theta(Z_{u_i^n}) - \theta(Z_{u_{i-1}^n}) \right)' (Z_{u_i^n} - Z_{u_{i-1}^n}) \right).$$

The last quadratic covariation process (without the minus sign) is equal to

$$\int_{T-t}^{T-s} \tilde{F}(c, \theta)(Z_u) \, \mathrm{d}u = \int_s^t \tilde{F}(c, \theta)(\zeta_u^T) \, \mathrm{d}u,$$

where $\tilde{F}(c, \theta) : E \to \mathbb{R}$ is given by

$$\tilde{F}(c, \theta) = \sum_{i,j=1}^d c^{ij} \partial_{z_i} \theta^j = \sum_{i,j=1}^d \left(\partial_{z_i}(c^{ij}\theta^j) - \theta^j \partial_{z_i}((c')^{ji}) \right) = \nabla \cdot (c\theta) - \theta' \operatorname{div} c,$$

since $c' = c$. Thus, (8.4) is established in the case where θ is twice continuously differentiable. The fact that (8.4) holds whenever θ is continuously differentiable follows from a density argument, noting that there exists a sequence $(\theta_n)_{n \in \mathbb{N}}$ of polynomials such that $\lim_{n \to \infty} \theta_n = \theta$ and $\lim_{n \to \infty} \nabla \theta_n = \nabla \theta$ both hold, where the convergence is uniform on compact subsets of E. $\qquad \square$

Lemma 8.4 *Let Assumption* 2.7 *hold. For each* $T \in \mathbb{R}_+$, *define the* $\mathbf{G}^T$*-adapted continuous-path* Δ^T *as in* (3.8). *Then* Δ^T *is a semimartingale on* $(\Omega, \mathbf{G}^T, \mathbb{P})$. *More*

precisely, for $t \in [0, T]$,

$$\Delta_t^T = 1 + \int_0^t \Delta_u^T \left(\theta' c \frac{\nabla p}{p} + \nabla \cdot (c\theta) - a \right) (\zeta_u^T) \, du$$

$$+ \int_0^t \Delta_u^T \left(\eta(\zeta_u^T)' \, dB_u^T + \theta' \sigma(\zeta_u^T) \, dW_u^T \right). \tag{8.5}$$

Proof Define $(\rho_t^T)_{t \in [0,T]}$ by $\rho_t^T := R_T - R_{T-t}$ for $t \in [0, T]$. In view of (1.2), (2.2), (8.1) and Lemma 8.3,

$$\rho^T = \int_{T-\cdot}^T \left(a + \frac{1}{2}(\theta' c \theta + \eta' \eta) \right) (Z_t) \, dt + \int_{T-\cdot}^T \left(\eta(Z_t)' \, dB_t + \theta' \sigma(Z_t) \, dW_t \right)$$

$$= \int_{T-\cdot}^T \left(a - \theta' m + \frac{1}{2}(\theta' c \theta + \eta' \eta) \right) (Z_t) \, dt + \int_{T-\cdot}^T \left(\eta(Z_t)' \, dB_t + \theta'(Z_t) \, dZ_t \right)$$

$$= \int_0^{\cdot} \left(a - \theta' m + \theta' \operatorname{div} c - \nabla \cdot (c\theta) + \frac{1}{2}(\theta' c \theta + \eta' \eta) \right) (\zeta_t^T) \, dt$$

$$- \int_0^{\cdot} \left(\eta(\zeta_t^T)' \, dB_t^T + \theta'(\zeta_t^T) \, d\zeta_t^T \right),$$

$$= \int_0^{\cdot} \left(a - \theta' c \frac{\nabla p}{p} - \nabla \cdot (c\theta) + \frac{1}{2}(\theta' c \theta + \eta' \eta) \right) (\zeta_t^T) \, dt$$

$$- \int_0^{\cdot} \left(\eta(\zeta_t^T)' \, dB_t^T + \theta' \sigma(\zeta_t^T) \, dW_t^T \right).$$

The fact that $D = \exp(-R)$ gives $\Delta^T = \exp(-\rho^T)$. Then the dynamics for Δ^T follow from the dynamics of ρ^T. $\qquad\qquad\square$

Proposition 8.5 *Let Assumption 2.7 hold. Then for each $T > 0$, there are a filtration $\mathbf{G}^T$ satisfying the usual conditions and d- and k-dimensional independent $(\mathbb{P}, \mathbf{G}^T)$-Brownian motions W^T, B^T on $[0, T]$ so that the pair (ζ^T, χ^T) has dynamics*

$$\zeta_t^T = \zeta_0^T + \int_0^T \left(c \frac{\nabla p}{p} + \operatorname{div} c - m \right) (\zeta_u^T) \, du + \int_0^T \sigma(\zeta_u^T) \, dW_u^T,$$

$$\chi_t^T = \chi_0^T + \int_0^T \left(f(\zeta_u^T) - \chi_u^T \left(a - \theta' c \frac{\nabla p}{p} - \nabla \cdot (c\theta) \right) (\zeta_u^T) \right) du \tag{8.6}$$

$$+ \int_0^T \chi_u^T \left(\theta' \sigma(\zeta_u^T) \, dW_u^T + \eta(\zeta_u^T)' \, dB_u^T \right).$$

Proof Proposition 8.1 immediately implies that under $\mathbb{P}$, ζ^T has the dynamics

$$\zeta_t^T = \zeta_0^T + \int_0^t \left(c \frac{\nabla p}{p} + \operatorname{div} c - m \right) (\zeta_u^T) \, du + \int_0^t \sigma(\zeta_u^T) \, dW_u^T, \qquad t \in [0, T],$$

where $(W_t^T)_{t \in [0,T]}$ is a Brownian motion on $(\Omega, \mathbf{G}^T, \mathbb{P})$. In order to specify the dynamics for χ^T, recall the definition of Δ^T from (3.8). Observe that

$$X_{T-t} = \frac{1}{D_{T-t}} \int_{T-t}^{\infty} D_u f(Z_u) \, \mathrm{d}u = \frac{D_T}{D_{T-t}} \left(X_T + \int_{T-t}^{T} \frac{D_u}{D_T} f(Z_u) \, \mathrm{d}u \right), t \in [0, T].$$

Then, using the definitions of χ^T, ζ^T and Δ^T, the above is rewritten as

$$\chi_t^T = \Delta_t^T \left(\chi_0^T + \int_0^t \frac{1}{\Delta_u^T} f(\zeta_u^T) \, \mathrm{d}u \right), \qquad t \in [0, T]. \tag{8.7}$$

Lemma 8.4 implies that Δ^T is a semimartingale, and hence (8.7) yields

$$\chi_t^T = \chi_0^T + \int_0^t \chi_u^T \frac{\mathrm{d}\Delta_u^T}{\Delta_u^T} + \int_0^t f(\zeta_u^T) \, \mathrm{d}u, \qquad t \in [0, T].$$

The result now follows by plugging in for $\mathrm{d}\Delta_u^T / \Delta_u^T$ from (8.5). $\qquad \square$

9 Proof of Theorem 3.4

9.1 Preliminaries

We first prove two technical results. The first asserts the existence of a probability space and stationary processes (ζ, χ) consistent with (ζ, χ^x) from Theorem 3.4 in the sense that given $\chi_0 = x$, it holds that $\chi_t = \chi_t^x, t \geq 0$. The second proposition shows that under the non-degeneracy assumption $|\eta|(z) > 0, z \in E$, and the regularity assumption $f \in C^2(E; \mathbb{R}_+)$, it follows that (ζ, χ) is ergodic.

Lemma 9.1 *If Assumption 2.7 holds, there is a filtered probability space $(\Omega, \mathbf{F}, \mathbb{Q})$ supporting independent d- and k-dimensional Brownian motions W and B, $\mathcal{F}_0$-measurable random variables ζ_0, χ_0 with joint distribution π, as well as a stationary process ζ with dynamics*

$$\zeta = \zeta_0 + \int_0^{\cdot} \left(c \frac{\nabla p}{p} + \operatorname{div} c - m \right)(\zeta_t) \, \mathrm{d}t + \int_0^{\cdot} \sigma(\zeta_t) \, \mathrm{d}W_t. \tag{9.1}$$

Furthermore, with Δ, χ^x defined as in (3.10), (3.11), if the process χ is defined by $\chi_t := \chi_t^{\chi_0}$ (see Remark 3.5), then (ζ, χ) is stationary with invariant measure π and joint dynamics

$$\mathrm{d}\zeta_t = \left(c \frac{\nabla p}{p} + \operatorname{div} c - m \right)(\zeta_t) \, \mathrm{d}t + \sigma(\zeta_t) \, \mathrm{d}W_t, \quad t \in \mathbb{R}_+,$$

$$\mathrm{d}\chi_t = \left(f(\zeta_t) - \chi_t \left(a - \theta' c \frac{\nabla p}{p} - \nabla \cdot (c\theta) \right)(\zeta_t) \right) \mathrm{d}t$$

$$+ \chi_t \left(\theta' \sigma(\zeta_t) \, \mathrm{d}W_t + \eta(\zeta_t)' \, \mathrm{d}B_t \right), \quad t \in \mathbb{R}_+. \tag{9.2}$$

Proof This follows from Proposition 8.1. Indeed, one can start with a probability space $(\Omega, \mathbf{F}, \mathbb{Q})$ supporting independent d- and k-dimensional Brownian motions W and B, respectively, as well as an $\mathcal{F}_0$-measurable random variable $(\zeta_0, \chi_0) \sim \pi$ (hence independent of W and B). Under the given regularity assumptions, Proposition 8.1 yields a strong, stationary solution ζ satisfying (9.1). Then, defining Δ as in (3.8) and, for $x > 0$, χ^x as in (3.11), it follows that (ζ, χ^x) and hence (ζ, χ) satisfy the SDE in (9.2). Under the given regularity assumptions, the law under $\mathbb{P}$ of (ζ^T, χ^T) given $\zeta_0^T = z, \chi_0^T = x$ coincides with the law under $\mathbb{Q}$ of (ζ, χ^x) given that $\zeta_0 = z$. Since by construction, π is an invariant measure for (ζ^T, χ^T), it follows from the Markov property that π is invariant for (ζ, χ) under $\mathbb{Q}$ and hence (ζ, χ) is stationary with invariant measure π. $\qquad\square$

Define the measures $\mathbb{Q}^{z,x}$ for $(z, x) \in F$ via

$$\mathbb{Q}^{z,x}[A] = \mathbb{Q}[A \mid \zeta_0 = z, \chi_0 = x], \qquad A \in \mathcal{F}_\infty.$$

We now consider when $|\eta| > 0$ on E and $f \in C^2(E; \mathbb{R}_+)$. According to Theorem 3.1, $g \in C^{2,\gamma}(F)$ and hence π possesses a density satisfying

$$\pi(z, x) = p(z)\partial_x g(z, x); \qquad (z, x) \in F. \tag{9.3}$$

Additionally, we have the following result.

Proposition 9.2 *Let Assumption 2.7 hold, and additionally suppose that $|\eta|(z) > 0$ for $z \in E$ and that $f \in C^2(E; \mathbb{R}_+)$. Then the process (ζ, χ) from Lemma 9.1 is ergodic. Thus, for all bounded measurable functions h on F and all $(z, x) \in F$,*

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T h(\zeta_t, \chi_t)\, \mathrm{d}t = \int_F h\, \mathrm{d}\pi \quad \mathbb{Q}^{z,x}\text{-a.s.} \tag{9.4}$$

Proof Recall A from (3.3) and define $b^R : F \to \mathbb{R}^{d+1}$ by

$$b^R(z, x) := \begin{pmatrix} (c(\nabla p/p) + \operatorname{div} c - m)(z) \\ f(z) - x(a - \theta' c\,(\nabla p/p) - \nabla \cdot (c\theta))(z) \end{pmatrix}. \tag{9.5}$$

From (9.2), it is clear that the generator for (ζ, χ) is $L^R := (1/2)A^{ij}\partial_{ij}^2 + (b^R)^i \partial_i$. In an abuse of notation, let $(\mathbb{Q}^{z,x})_{(z,x) \in F}$ also denote the solution to the generalized martingale problem for L^R on F. Using Theorem 3.1 and the fact that under the given coefficient regularity assumptions, $g \in C^3(F)$ (see [15, Theorem 6.17]), a lengthy calculation performed in Lemma A.1 below shows that the density π from (9.3) solves $\tilde{L}^R\pi = 0$, where $\tilde{L}^R$ is the formal adjoint to L. Since by construction, $\iint_F \pi(z, x)\, \mathrm{d}z\, \mathrm{d}x = 1$, positive recurrence will follow once it is shown that $(\mathbb{Q}^{z,x})_{(z,x) \in F}$ is recurrent. By Proposition 8.1, the restriction of $\mathbb{Q}^{z,x}$ to the first d coordinates (i.e., the part for ζ) is positive recurrent. Since by (3.11) it is evident that χ does not hit 0 in finite time, it follows that χ does not explode under $\mathbb{Q}^{z,x}$. Thus, [27, Corollary 4.9.4] shows that (ζ, χ) is recurrent. Now (9.4) follows from [27, Theorem 4.9.5]. $\qquad\square$

9.2 Proof of Theorem 3.4

The proof of Theorem 3.4 uses a number of approximation arguments. To make these precise, we first enlarge the original probability space $(\Omega, \mathbf{F}, \mathbb{P})$ so that it contains a one-dimensional Brownian motion $\hat{B}$ which is independent of Z_0, W and B. Let D be as in (1.3), and for $\varepsilon > 0$, define $D^\varepsilon := D\mathcal{E}(\sqrt{\varepsilon}\hat{B})$. Similarly to (1.1), define

$$X_0^\varepsilon := \int_0^\infty D_t^\varepsilon f(Z_t)\, \mathrm{d}t. \tag{9.6}$$

Note that D^ε takes the form (1.3) for $\eta^\varepsilon(z) = (\eta(z), \sqrt{\varepsilon})$ and when the Brownian motion B therein is the $(k+1)$-dimensional Brownian motion $(B, \hat{B})$. Note that $|\eta^\varepsilon|^2 = |\eta|^2 + \varepsilon > 0$. Denote by π^ε the joint distribution of (Z_0, X_0^ε) under $\mathbb{P}$ and by g^ε the conditional cumulative distribution function of X_0^ε given $Z_0 = z$. By Theorem 3.1, it follows that $g^\varepsilon \in C^{2,\gamma}(F)$ and hence π^ε admits a density.

In a similar manner, by enlarging the probability space $(\Omega, \mathbf{F}, \mathbb{Q})$ of Lemma 9.1 to include a Brownian motion (still labeled $\hat{B}$) which is independent of ζ_0, χ_0, W and B, and defining the family of processes $(\Delta^\varepsilon)_{\varepsilon>0}$ and $(\chi^{\varepsilon,x})_{\varepsilon>0}$ for $x > 0$ according to

$$\Delta_t^\varepsilon := \Delta_t \mathcal{E}(\sqrt{\varepsilon}\hat{B})_t, \qquad t \geq 0,$$
$$\chi_t^{\varepsilon,x} := \Delta_t^\varepsilon \left(x + \int_0^t \frac{1}{\Delta_u^\varepsilon} f(\zeta_u)\, \mathrm{d}u \right), \qquad t \geq 0, \tag{9.7}$$

it follows that $(\zeta, \chi^{x,\varepsilon})$ solves the SDE

$$\mathrm{d}\zeta_t = (m + 2\xi)(\zeta_t)\, \mathrm{d}t + \sigma(\zeta_t)\, \mathrm{d}W_t,$$
$$\mathrm{d}\chi_t^{\varepsilon,x} = \left(f(\zeta_t) - \chi_t^\varepsilon \left(a - \theta' c \frac{\nabla p}{p} - \nabla \cdot (c\theta) \right)(\zeta_t) \right) \mathrm{d}t$$
$$+ \chi_t^\varepsilon \left(\theta' \sigma(\zeta_t)\, \mathrm{d}W_u + \eta^\varepsilon(\zeta_t)'(\mathrm{d}B_t, \mathrm{d}\hat{B}_t) \right). \tag{9.8}$$

Since $|\eta^\varepsilon| \geq \sqrt{\varepsilon} > 0$, Proposition 9.2 shows that for $f \in C^2(E; \mathbb{R}_+)$, the generator $L^{\varepsilon,R}$ associated to (9.8) is positive recurrent with invariant density π^ε and thus for all $(z, x) \in F$ and all bounded measurable functions h on F (note that conditioned upon $\chi_0 = x$, we have $\chi_0^{\varepsilon,x} = \chi_0^x = x = \chi_0$),

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T h(\zeta_t, \chi_t^{x,\varepsilon})\, \mathrm{d}t = \int_F h\, \mathrm{d}\pi^\varepsilon, \qquad \mathbb{Q}^{z,x}\text{-a.s.} \tag{9.9}$$

With all the notation in place, Theorem 3.4 is the culmination of a number of lemmas, which are now presented. The first lemma implies that π^ε converges weakly to π as $\varepsilon \downarrow 0$.

Lemma 9.3 *Let Assumption 2.7 hold. Define X_0^ε as in (9.6). Then X_0^ε converges to X in $\mathbb{P}$-measure as $\varepsilon \to 0$.*

Proof Denote by $\mathcal{G}$ the sigma-field generated by Z_0, W and B and define the process δ^ε by $\delta_t^\varepsilon := D_t^\varepsilon/D_t = \mathcal{E}(\sqrt{\varepsilon}\hat{B}_t)$. By the independence of δ^ε and $\mathcal{G}$,

$$\mathbb{E}\left[|X_0^\varepsilon - X_0| \mid \mathcal{G}\right] \leq \int_0^\infty \mathbb{E}\left[|\delta_t^\epsilon - 1| \mid \mathcal{G}\right] D_t f(Z_t)\, dt = \int_0^\infty \mathbb{E}[|\delta_t^\epsilon - 1|]D_t f(Z_t)\, dt.$$

Now set $h_t^\varepsilon := \sqrt{e^{\varepsilon t} - 1}$. Note that h^ε is increasing in ε with $\lim_{\varepsilon \to 0} h^\varepsilon = 0$. Furthermore,

$$\mathbb{E}[|\delta_t^\varepsilon - 1|] \leq \left(\mathbb{E}\left[|\delta_t^\varepsilon - 1|^2\right]\right)^{1/2} = \sqrt{\exp(\varepsilon t) - 1} = h_t^\varepsilon.$$

By assumption, $\mathbb{P}[X_0 < \infty] = 1$. Since for any $\varepsilon > 0$, $\sup_{t \geq 0} \delta_t^\varepsilon < \infty$ $\mathbb{P}$-a.s., it thus follows that $\mathbb{P}[X_0^\varepsilon < \infty] = 1$. The dominated convergence theorem applied pathwise (recall that there exists a $\kappa > 0$ so that $e^{\kappa t} D_t \to 0$ $\mathbb{P}$-a.s.) then gives that $\lim_{\varepsilon \to 0} \mathbb{E}[|X_0^\varepsilon - X_0| \mid \mathcal{G}] = 0$, which shows that the pair (Z_0, X_0^ε) converges in probability to (Z_0, X_0), finishing the proof. $\qquad\square$

Next, define $\mathcal{C}$ as the class of (Borel-measurable) functions h which are bounded and Lipschitz in x, uniformly in z; in other words,

$$\mathcal{C} := \left\{h \in \mathcal{B}(F; \mathbb{R}) : \sup_{z \in E} |h(z, x_1) - h(z, x_2)| \leq K(h)\,(1 \wedge |x_1 - x_2|)\right\} \tag{9.10}$$

for some $K(h) > 0$ (which may depend upon h) and all $x_1, x_2 > 0$. The next lemma gives a weak form of the convergence in Theorem 3.4 for regular f. Note that the notation $\mathbb{Q}\text{-}\lim_{T \to \infty}$ stands for the limit in $\mathbb{Q}$-probability as $T \to \infty$.

Lemma 9.4 *Let Assumption 2.7 hold. Assume additionally that $f \in C^2(E; \mathbb{R}_+)$. Then for all $x > 0$ and all $h \in \mathcal{C}$,*

$$\mathbb{Q}\text{-}\lim_{T \to \infty} \frac{1}{T} \int_0^T h(\zeta_t, \chi_t^x)\, dt = \int_F h\, d\pi. \tag{9.11}$$

Proof For ease of presentation, we adopt the following notational conventions. First, for any measurable function f and probability measure ν on F, set

$$\langle h, \nu \rangle := \int_F h\, d\nu. \tag{9.12}$$

Next, similarly to $\hat{\pi}_T^x$ in (3.12), we define $\hat{\pi}_T^{\varepsilon,x}$ to be the empirical measure of $(\zeta, \chi^{\varepsilon,x})$ on $[0, T]$ for $\chi^{\varepsilon,x}$ as in (9.7). Thus, we write

$$\frac{1}{T} \int_0^T h(\zeta_t, \chi_t^x)\, dt = \langle h, \hat{\pi}_T^x \rangle, \qquad \frac{1}{T} \int_0^T h(\zeta_t, \chi_t^{\varepsilon,x})\, dt = \langle h, \hat{\pi}_T^{\varepsilon,x} \rangle. \tag{9.13}$$

Proposition 9.2 implies for all $x > 0$ and $\varepsilon > 0$ that

$$\mathbb{Q}\text{-}\lim_{T \to \infty} \langle h, \hat{\pi}_T^{\varepsilon,x} \rangle = \langle h, \pi^\varepsilon \rangle.$$

Indeed, (9.9) gives for all $(z, x) \in F$ that

$$\lim_{T \to \infty} \langle h, \hat{\pi}_T^{\varepsilon, x} \rangle = \langle h, \pi^{\varepsilon} \rangle \qquad \mathbb{Q}^{z, x}\text{-a.s.}$$

Thus, the above limit holds $\mathbb{Q}$-almost surely and hence in probability.

To prove (9.11), we need to show that for any increasing $\mathbb{R}_+$-valued sequence $(T_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} T_n = \infty$, there is a subsequence $(T_{n_k})_{k \in \mathbb{N}}$ such that

$$\mathbb{Q}\text{-}\lim_{k \to \infty} \langle h, \hat{\pi}_{T_{n_k}}^x \rangle = \langle h, \pi \rangle,$$

as this implies (9.11) by considering double subsequences. To this end, let $(\varepsilon_k)_{k \in \mathbb{N}}$ be any strictly positive sequence that converges to zero, and assume that $\varepsilon_1 < \kappa$, where $\kappa > 0$ is from Assumption (A5). Next, pick T_{n_k} large enough so that $k / T_{n_k} \to 0$ and such that

$$\mathbb{Q}\left[|\langle h, \hat{\pi}_{T_{n_k}}^{\varepsilon_k, x} \rangle - \langle h, \pi^{\varepsilon_k} \rangle| > \frac{1}{k} \right] \leq \frac{1}{k}.$$

As argued above, this is possible since $\langle h, \hat{\pi}_T^{\varepsilon_k, x} \rangle$ converges to $\langle h, \pi^{\varepsilon_k} \rangle$ in $\mathbb{Q}$-probability. Since Lemma 9.3 implies $\lim_{\varepsilon \to 0} \langle h, \pi^{\varepsilon_k} \rangle = \langle h, \pi \rangle$, it follows that

$$\mathbb{Q}\text{-}\lim_{k \to \infty} \langle h, \hat{\pi}_{T_{n_k}}^{\varepsilon_k, x} \rangle = \langle h, \pi \rangle.$$

Since

$$|\langle h, \hat{\pi}_{T_{n_k}}^x \rangle - \langle h, \pi \rangle| \leq |\langle h, \hat{\pi}_{T_{n_k}}^x \rangle - \langle h, \hat{\pi}_{T_{n_k}}^{\varepsilon_k, x} \rangle| + |\langle h, \hat{\pi}_{T_{n_k}}^{\varepsilon_k, x} \rangle - \langle h, \pi \rangle|,$$

it suffices to show

$$\mathbb{Q}\text{-}\lim_{k \to \infty} |\langle h, \hat{\pi}_{T_{n_k}}^{\varepsilon_k, x} \rangle - \langle h, \hat{\pi}_{T_{n_k}}^x \rangle| = 0.$$

In fact, the claim is that

$$\lim_{k \to \infty} \mathbb{E}^{\mathbb{Q}} \big[|\langle h, \hat{\pi}_{T_{n_k}}^{\varepsilon_k, x} \rangle - \langle h, \hat{\pi}_{T_{n_k}}^x \rangle| \big] = 0,$$

or the even stronger (recall (9.13)) result

$$\lim_{k \to \infty} \left(\frac{1}{T_{n_k}} \int_0^{T_{n_k}} \mathbb{E}^{\mathbb{Q}} \big[|h(\zeta_t, \chi_t^{\varepsilon_k, x}) - h(\zeta_t, \chi_t^x)| \big] \, dt \right) = 0. \qquad (9.14)$$

From (9.10),

$$\frac{1}{T_{n_k}} \int_0^{T_{n_k}} \mathbb{E}^{\mathbb{Q}} \big[|h(\zeta_t, \chi_t^{\varepsilon_k, x}) - h(\zeta_t, \chi_t^x)| \big] \, dt \leq \frac{K}{T_{n_k}} \int_0^{T_{n_k}} \mathbb{E}^{\mathbb{Q}} \big[1 \wedge |\chi_t^{\varepsilon_k, x} - \chi_t^x| \big] \, dt. \tag{9.15}$$

Furthermore, recall that

$$\chi_t^x = \Delta_t \left(x + \int_0^t \frac{1}{\Delta_u} f(\zeta_u) \, du \right), \qquad \chi_t^{\varepsilon_k, x} = \Delta_t^{\varepsilon_k, x} \left(x + \int_0^t \frac{1}{\Delta_u^{\varepsilon_k}} f(\zeta_u) \, du \right),$$

where Δ^{ε_k} is from (9.7). With $\delta^{\varepsilon_k} := \mathcal{E}(\sqrt{\varepsilon_k}\hat{B})$, it follows that under $\mathbb{Q}$,

$$|\chi_t^{\varepsilon_k,x} - \chi_t^x| \leq x|\Delta_t^{\varepsilon_k} - \Delta_t| + \int_0^t \left|\frac{\Delta_t^{\varepsilon_k}}{\Delta_u^{\varepsilon_k}} - \frac{\Delta_t}{\Delta_u}\right| f(\zeta_u)\,\mathrm{d}u$$

$$= x\Delta_t|\delta_t^{\varepsilon_k} - 1| + \int_0^t \frac{\Delta_t}{\Delta_u}\left|\frac{\delta_t^{\varepsilon_k}}{\delta_u^{\varepsilon_k}} - 1\right| f(\zeta_u)\,\mathrm{d}u.$$

With $\mathcal{G}$ now denoting the σ-field generated by ζ_0, W and B, the independence of $\hat{B}$ and $\mathcal{G}$ implies that

$$\mathbb{E}^{\mathbb{Q}}\left[|\chi_t^{\varepsilon_k,x} - \chi_t^x| \mid \mathcal{G}\right] \leq x\Delta_t h_t^{\varepsilon_k} + \int_0^t \frac{\Delta_t}{\Delta_u} h_{t-u}^{\varepsilon_k} f(\zeta_u)\,\mathrm{d}u, \qquad (9.16)$$

where for any $\varepsilon > 0$, h^ε is from Lemma 9.3. Since ζ is stationary under $\mathbb{Q}$, it holds for all $t > 0$ that the distribution of Δ_t under $\mathbb{Q}$ coincides with the distribution of D_t under $\mathbb{P}$, and the distribution of $\int_0^t (\Delta_t/\Delta_u) h_{t-u}^{\varepsilon_k} f(\zeta_u)\,\mathrm{d}u$ under $\mathbb{Q}$ is the same as the distribution of $\int_0^t D_u h_u^{\varepsilon_k} f(Z_u)\,\mathrm{d}u$ under $\mathbb{P}$.

We next claim that there exists a sequence $\delta_k \to 0$ such that

$$\sup_{t \in [k,\infty)} \mathbb{P}\left[1 \wedge \left(x D_t h_t^{\varepsilon_k} + \int_0^t D_u h_u^{\varepsilon_k} f(Z_u)\,\mathrm{d}u\right) > \delta_k\right] \leq \delta_k, \quad \forall k \in \mathbb{N}. \qquad (9.17)$$

This is shown at the end of the proof. Admitting this, from

$$\mathbb{E}^{\mathbb{Q}}\left[1 \wedge |\chi_t^{\varepsilon_k,x} - \chi_t^x| \mid \mathcal{G}\right] \leq 1 \wedge \mathbb{E}^{\mathbb{Q}}\left[|\chi_t^{x,\varepsilon_k} - \chi_t^x| \mid \mathcal{G}\right],$$

it follows that

$$\lim_{k\to\infty}\left(\sup_{t\in[k,\infty)} \mathbb{E}^{\mathbb{Q}}\left[1 \wedge |\chi_t^{\varepsilon_k,x} - \chi_t^x|\right]\right)$$

$$= \lim_{k\to\infty}\left(\sup_{t\in[k,\infty)} \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[1 \wedge |\chi_t^{\varepsilon_k,x} - \chi_t^x| \mid \mathcal{G}\right]\right]\right)$$

$$\leq \lim_{k\to\infty}\left(\sup_{t\in[k,\infty)} \mathbb{E}\left[1 \wedge \left(x D_t h_t^{\varepsilon_k} + \int_0^t D_u h_u^{\varepsilon_k} f(Z_u)\,\mathrm{d}u\right)\right]\right)$$

$$\leq \lim_{k\to\infty} 2\delta_k = 0.$$

Above, the first inequality holds because of (9.16) and the second by (9.17) and the fact that for any random variable Y, $\mathbb{E}[1 \wedge Y] \leq \delta + \mathbb{P}[1 \wedge Y > \delta]$. The last equality follows by the construction of δ_k. Recall that T_{n_k} was chosen so that

$\lim_{k\to\infty}(k/T_{n_k}) = 0$; so it follows that

$$\limsup_{k\to\infty}\left(\frac{1}{T_{n_k}}\int_0^{T_{n_k}}\mathbb{E}^{\mathbb{Q}}\left[1\wedge|\chi_t^{\epsilon_k,x}-\chi_t^x|\right]dt\right)$$

$$\leq\limsup_{k\to\infty}\left(\frac{k}{T_{n_k}}+\frac{T_{n_k}-k}{T_{n_k}}\sup_{t\in[k,\infty)}\mathbb{E}^{\mathbb{Q}}\left[1\wedge|\chi_t^{\varepsilon_k,x}-\chi_t^x|\right]\right)$$

$$= 0,$$

which in view of (9.15) implies (9.14), finishing the proof.

It remains to show (9.17). Since $1\wedge(a+b)\leq 1\wedge a+1\wedge b$ for any $a,b > 0$, the two terms on the right-hand side of (9.17) are treated separately. Let $\delta_k > 0$. First we have

$$\mathbb{P}[1\wedge xD_th_t^{\varepsilon_k} > \delta_k]\leq\mathbb{P}[xD_th_t^{\varepsilon_k} > \delta_k]=\mathbb{P}[xD_te^{\kappa t} > \delta_k e^{\kappa t}/h_t^{\varepsilon_k}]$$

Now, $h_t^{\varepsilon_k}\leq e^{\varepsilon_k/2t}$ so that for $t\geq k$, $e^{\kappa t}/h_t^{\varepsilon_k}\geq e^{(\kappa-\varepsilon_k/2)t}\geq e^{(\kappa-\varepsilon_k/2)k}$ since $\varepsilon_k/2 < \kappa$. So, for any $\delta_k > e^{-(\kappa-\varepsilon_k/2)(k/2)}$, it follows that

$$\mathbb{P}[xD_th_t^{\varepsilon_k} > \delta_k]\leq\mathbb{P}\left[xD_te^{\kappa t}\geq e^{(\kappa-\varepsilon_k/2)(k/2)}\right]$$

Set $\tilde{\delta}_k := \sup_{t\geq k}\mathbb{P}[xD_te^{\kappa t}\geq e^{(\kappa-\varepsilon_k/2)(k/2)}]$. Since $D_te^{\kappa t}$ goes to 0 in $\mathbb{P}$-probability, it follows that $\tilde{\delta}_k\to 0$. Thus, taking δ_k to be the maximum of $\tilde{\delta}_k$ and $e^{-(\kappa-\varepsilon_k/2)(k/2)}$, it follows that

$$\mathbb{P}[1\wedge\chi D_th_t^{\varepsilon_k} > \delta_k]\leq\delta_k.$$

Turning to the second term in (9.17), it is clear that

$$1\wedge\int_0^t D_uh_u^{\varepsilon_k}f(Z_u)\,du\leq 1\wedge\int_0^\infty D_uh_u^{\varepsilon_k}f(Z_u)\,du.$$

As shown in the proof of Lemma 9.3, $\int_0^\infty D_uh_u^{\varepsilon_k}f(Z_u)\,du$ goes to 0 almost surely as $k\to\infty$. Thus by the bounded convergence theorem, $\mathbb{E}[1\wedge\int_0^\infty D_uh_u^{\varepsilon_k}f(Z_u)\,du]\to 0$ as $k\to\infty$. Since

$$\mathbb{P}\left[1\wedge\int_0^\infty D_uh_u^{\varepsilon_k}f(Z_u)\,du > \delta_k\right]\leq\frac{1}{\delta_k}\mathbb{E}\left[1\wedge\int_0^\infty D_uh_u^{\varepsilon_k}f(Z_u)\,du\right],$$

upon defining $\delta_k := \sqrt{\mathbb{E}[1\wedge\int_0^\infty D_uh_u^{\varepsilon_k}f(Z_u)\,du]}$, it follows that

$$\mathbb{P}\left[1\wedge\int_0^\infty D_uh_u^{\varepsilon_k}f(Z_u)\,du > \delta_k\right]\leq\delta_k,$$

and $\delta_k\to 0$. This concludes the proof since to combine the two terms one can take δ_k to be twice the maximum of the δ_k for the individual terms. $\square$

The next lemma proves the convergence in Lemma 9.4 for $f\in\mathbb{L}^1(E,p)$, not just $f\in C^2(E;\mathbb{R}_+)$.

Lemma 9.5 *Let Assumption 2.7 hold. Then for all $x > 0$ and all $h \in \mathcal{C}$,*

$$\mathbb{Q}\text{-}\lim_{T\to\infty} \frac{1}{T} \int_0^T h(\zeta_t, \chi_t^x)\, dt = \int_F h\, d\pi. \tag{9.18}$$

Proof By mollifying f, since p is tight in E, there exists a sequence of functions $f^n \in C^2(E) \cap \mathbb{L}^1(E, p)$ with $f^n \geq 0$ such that

$$\int_E |f^n(z) - f(z)| p(z)\, dz \leq n^{-2} 2^{-n}.$$

Note that

$$
\begin{aligned}
\mathbb{E}\left[\int_0^\infty n e^{-t/n} |f^n(Z_t) - f(Z_t)|\, dt\right] &= \int_0^\infty n e^{-t/n} \mathbb{E}[|f^n(Z_t) - f(Z_t)|]\, dt \\
&= \int_0^\infty n e^{-t/n} \left(\int_E |f^n(z) - f(z)| p(z)\, dz\right) dt \\
&\leq \int_0^\infty n^{-1} e^{-t/n} 2^{-n}\, dt \\
&= 2^{-n}.
\end{aligned}
$$

Thus, by the Borel–Cantelli lemma, it follows that $\mathbb{P}$-almost surely

$$\lim_{n\to\infty} \int_0^\infty n e^{-t/n} |f^n(Z_t) - f(Z_t)|\, dt = 0.$$

For $n > \kappa$ from Assumption 2.7, let $A_n = n^{-1} \sup_{t\in\mathbb{R}_+} (e^{t/n} D_t)$. Note that we have $\lim_{n\to\infty} A_n = 0$ almost surely since for each $\delta > 0$, we can find a $\mathbb{P}$-almost surely finite random variable $T = T(\delta)$ so that $D_t \leq \delta e^{-\kappa t}$ for $t \geq T$, and hence

$$A_n = \frac{1}{n} \sup_{t\in t\in\mathbb{R}_+} (e^{t/n} D_t) \leq \frac{1}{n} e^{T/n} \sup_{t\leq T} D_t + \frac{\delta}{n}.$$

Since

$$\int_0^\infty D_t |f^n(Z_t) - f(Z_t)|\, dt \leq A_n \int_0^\infty n e^{-t/n} |f^n(Z_t) - f(Z_t)|\, dt,$$

we see that

$$\lim_{n\to\infty} \int_0^\infty D_t |f^n(Z_t) - f(Z_t)|\, dt = 0 \qquad \mathbb{P}\text{-a.s.} \tag{9.19}$$

Thus, with $X_0^n := \int_0^\infty D_t f^n(Z_t)\, dt$, we get $\lim_{n\to\infty} X_0^n = X_0$ almost surely, and hence if π^n is the joint distribution of (Z_0, X_0^n), then π^n converges to π weakly as $n \to \infty$. Now, on the same probability space as in Lemma 9.1, define

$$\chi_t^{x,n} := \Delta_t \left(x + \int_0^t \Delta_t^{-1} f^n(\zeta_t)\, dt\right), \quad t \geq 0.$$

Note that

$$|\chi_t^{n,x} - \chi_t^x| \le \Delta_t \int_0^t \Delta_u^{-1} |f^n(\zeta_u) - f(\zeta_u)|\, du, \qquad \forall\, t \ge 0,$$

and by construction the law of the process on the right-hand side above under $\mathbb{Q}$ is the same as the law of $\int_0^{\cdot} D_u |f^n(Z_u) - f(Z_u)|\, du$ under $\mathbb{P}$. It thus follows that for $\delta > 0$,

$$\sup_{t \in \mathbb{R}_+} \mathbb{Q}[|\chi_t^{n,x} - \chi_t^x| > \delta] \le \mathbb{P}\left[\int_0^\infty D_u |f^n(Z_u) - f(Z_u)|\, du > \delta\right] =: \phi^n(\delta).$$

By (9.19), we can find a nonnegative sequence $\delta_n \to 0$ with $\lim_{n\to\infty} \phi^n(\delta_n) = 0$. Now, for $h \in \mathcal{C}$, we have almost surely for $t \ge 0$ that

$$|h(\zeta_t, \chi_t^{n,x}) - h(\zeta_t, \chi_t^x)| \le K(1 \wedge |\chi_t^{n,x} - \chi_t^x|).$$

Therefore, with $\hat{\pi}_T^{x,n}$ denoting the empirical law of $(\zeta, \chi^{n,x})$, we have

$$\mathbb{E}^{\mathbb{Q}}\big[|\langle h, \hat{\pi}_T^{x,n}\rangle - \langle h, \hat{\pi}_T^x\rangle|\big] \le \frac{K}{T} \int_0^T \mathbb{E}^{\mathbb{Q}}\big[1 \wedge |\chi_t^{n,x} - \chi_t^x|\big]\, dt.$$

Since for any $0 < \delta < 1$ and random variable Y, we have $\mathbb{E}[1 \wedge |Y|] \le \delta + \mathbb{P}[|Y| > \delta]$, it follows that for any n,

$$\sup_{T \in \mathbb{R}_+} \mathbb{E}^{\mathbb{Q}}\big[|\langle h, \hat{\pi}_T^{x,n}\rangle - \langle h, \hat{\pi}_T^x\rangle|\big] \le K\big(\phi^n(\delta) + \delta\big),$$

and hence for the given sequence (δ_n) that

$$\limsup_{n\to\infty} \sup_{T \in \mathbb{R}_+} \mathbb{E}^{\mathbb{Q}}\big[|\langle h, \hat{\pi}_T^{x,n}\rangle - \langle h, \hat{\pi}_T^x\rangle|\big] \le \limsup_{n\to\infty} K\big(\phi^n(\delta_n) + \delta_n\big) = 0. \qquad (9.20)$$

Now fix a sequence (T_k) such that $\lim_{k\to\infty} T_k = \infty$. Since Lemma 9.4 implies for each n that $\mathbb{Q}\text{-}\lim_{T\to\infty} |\langle h, \hat{\pi}_T^{x,n}\rangle - \langle h, \pi^n\rangle| = 0$ for each n, we can find a T_{k_n} so that

$$\mathbb{Q}\left[|\langle h, \hat{\pi}_{T_{k_n}}^{x,n}\rangle - \langle h, \pi^n\rangle| > \frac{1}{n}\right] < \frac{1}{n}.$$

It thus follows that

$$\mathbb{Q}\text{-}\lim_{n\to\infty} |\langle h, \hat{\pi}_{T_{k_n}}^{n,x}\rangle - \langle h, \pi^n\rangle| = 0.$$

Since $\lim_{n\to\infty} |\langle h, \pi^n \rangle - \langle h, \pi \rangle| = 0$, it follows by (9.20) for each $\gamma > 0$ that

$$
\mathbb{Q}\big[|\langle h, \hat{\pi}^x_{T_{k_n}} \rangle - \langle h, \pi \rangle| > \gamma\big]
$$

$$
\leq \mathbb{Q}\bigg[|\langle h, \hat{\pi}^x_{T_{k_n}} \rangle - \langle h, \hat{\pi}^{x,n}_{T_{k_n}} \rangle| > \frac{\gamma}{3}\bigg] + \mathbb{Q}\bigg[|\langle h, \hat{\pi}^{x,n}_{T_{k_n}} \rangle - \langle h, \pi^n \rangle| > \frac{\gamma}{3}\bigg]
$$

$$
+ 1_{\{|\langle h, \pi^n \rangle - \langle h, \pi \rangle| > \frac{\gamma}{3}\}}
$$

$$
\leq \frac{3}{\gamma} \sup_{T \in \mathbb{R}_+} \mathbb{E}^{\mathbb{Q}}\big[|\langle h, \hat{\pi}^x_T \rangle - \langle h, \hat{\pi}^{x,n}_T \rangle|\big] + \mathbb{Q}\bigg[|\langle h, \hat{\pi}^{x,n}_{T_{k_n}} \rangle - \langle h, \pi^n \rangle| > \frac{\gamma}{3}\bigg]
$$

$$
+ 1_{\{|\langle h, \pi^n \rangle - \langle h, \pi \rangle| > \frac{\gamma}{3}\}}
$$

$$
\longrightarrow 0 \quad \text{as } n \to \infty.
$$

We have just shown that for any sequence $((h, \hat{\pi}^x_{T_k}))$, there is a subsequence $((h, \hat{\pi}^x_{T_{k_n}}))$ which converges in $\mathbb{Q}$-probability to $\langle h, \pi \rangle$, which proves that $((h, \hat{\pi}^x_T))$ converges in $\mathbb{Q}$-probability to $\langle h, \pi \rangle$, proving (9.18). $\qquad\square$

The next lemma strengthens the convergence in Lemma 9.5 to almost sure convergence under $\mathbb{Q}$, but for π-almost every $x > 0$, for $h \in \mathcal{C}$ from (9.10).

Lemma 9.6 *Let Assumption 2.7 hold. Then for all $h \in \mathcal{C}$ and π-almost every $x > 0$,*

$$
\lim_{T \to \infty} \frac{1}{T} \int_0^T h(\zeta_t, \chi^x_t) \, dt = \int_F h \, d\pi, \qquad \mathbb{Q}\text{-a.s.} \tag{9.21}
$$

Proof We again use the notation in (9.12). Recall χ from Lemma 9.1 and define $\hat{\pi}_T$ as the empirical law of (ζ, χ) on $[0, T]$. Given that (ζ, χ) is stationary under $\mathbb{Q}$, the ergodic theorem implies for all bounded measurable functions h on F that there is a random variable Y such that

$$
\lim_{T \to \infty} \langle h, \hat{\pi}_T \rangle = Y \qquad \mathbb{Q}\text{-a.s.} \tag{9.22}
$$

By Lemma 9.5, it holds that for $h \in \mathcal{C}$, $Y = \langle h, \pi \rangle$ with $\mathbb{Q}$-probability one. Indeed, let $\delta > 0$ and note that

$$
\mathbb{Q}[|Y - \langle h, \pi \rangle| \geq \delta] \leq \mathbb{Q}[|Y - \langle h, \hat{\pi}_T \rangle| + |\langle h, \hat{\pi}_T \rangle - \langle h, \pi \rangle| \geq \delta]
$$

$$
\leq \mathbb{Q}\bigg[|Y - \langle h, \hat{\pi}_T \rangle| \geq \frac{\delta}{2}\bigg] + \mathbb{Q}\bigg[|\langle h, \hat{\pi}_T \rangle - \langle h, \pi \rangle| \geq \frac{\delta}{2}\bigg].
$$

The first of these two terms goes to 0 by (9.22). As for the second, denote by $\pi|_x$ the marginal of π with respect to χ. Then

$$
\mathbb{Q}\bigg[|\langle h, \hat{\pi}_T \rangle - \langle h, \pi \rangle| \geq \frac{\delta}{2}\bigg] = \int_0^\infty \pi|_x(dx) \mathbb{Q}\bigg[|\langle h, \hat{\pi}^x_T \rangle - \langle h, \pi \rangle| \geq \frac{\delta}{2}\bigg].
$$

By Lemma 9.4, the integrand goes to 0 as $T \to \infty$ for all $x > 0$, and thus the result follows by the bounded convergence theorem. Next, we have

$$1 = \mathbb{Q}\Big[\lim_{T \to \infty} \langle h, \hat{\pi}_T \rangle = \langle h, \pi \rangle \Big] = \int_0^\infty \pi|_x(\mathrm{d}x) \mathbb{Q}\Big[\lim_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle = \langle h, \pi \rangle \Big],$$

and thus (9.21) holds for π-a.e. $x > 0$, finishing the proof. $\qquad\square$

The last preparatory lemma strengthens Lemma 9.6 to show almost sure convergence for all starting points $x > 0$, not just π-almost every $x > 0$.

Lemma 9.7 *Let Assumption* 2.7 *hold. Then for all $h \in \mathcal{C}$ and all $x > 0$,*

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T h(\zeta_t, \chi_t^x) \, \mathrm{d}t = \int_F h \, \mathrm{d}\pi \qquad \mathbb{Q}\text{-}a.s. \tag{9.23}$$

Proof Recall from Remark 3.5 that χ^x takes the form

$$\chi_t^x = \Delta_t \left(x + \int_0^t \frac{1}{\Delta_t} f(\zeta_t) \, \mathrm{d}t \right), \qquad t \geq 0. \tag{9.24}$$

Let $h \in \mathcal{C}$. By Lemma 9.6, there is some $x_0 > 0$ such that (9.23) holds. Using the notation in (9.12) and (9.24), it easily follows for any $x > 0$ that

$$
\begin{aligned}
|\langle h, \hat{\pi}_T^x \rangle - \langle h, \hat{\pi}_T^{x_0} \rangle| &\leq \frac{1}{T} \int_0^T |h(\zeta_t, \chi_t^x) - h(\zeta_t, \chi_t^{x_0})| \, \mathrm{d}t \\
&\leq \frac{K}{T} \int_0^T (1 \wedge |\chi_t^x - \chi_t^{x_0}|) \, \mathrm{d}t \\
&= \frac{K}{T} \int_0^T (1 \wedge \Delta_t |x - x_0|) \, \mathrm{d}t \leq \frac{K|x - x_0|}{T} \int_0^\infty \Delta_t \, \mathrm{d}t.
\end{aligned}
$$

We show below that $\mathbb{Q}[\int_0^\infty \Delta_t \, \mathrm{d}t < \infty] = 1$. Admitting this, it holds that $\mathbb{Q}$-almost surely, $\lim_{T \to \infty} |\langle h, \hat{\pi}_T^x \rangle - \langle h, \hat{\pi}_T^{x_0} \rangle| = 0$, and hence the result follows since (9.23) holds for x_0.

It remains to prove that $\mathbb{Q}[\int_0^\infty \Delta_t \, \mathrm{d}t < \infty] = 1$. By way of contradiction, assume there is some $0 < \delta \leq 1$ so that $\mathbb{Q}[\int_0^\infty \Delta_t \, \mathrm{d}t = \infty] = \delta$. Then, for each N it holds that $\mathbb{Q}[\int_0^\infty \Delta_t \, \mathrm{d}t > N] \geq \delta$, which in turn implies $\lim_{T \to \infty} \mathbb{Q}[\int_0^T \Delta_t \, \mathrm{d}t > N] \geq \delta$. By construction, for any fixed $T > 0$, the law of Δ on $[0, T]$ under $\mathbb{Q}$ coincides with the law of D under $\mathbb{P}$ on $[0, T]$. It thus holds that $\lim_{T \to \infty} \mathbb{P}[\int_0^T D_t \, \mathrm{d}t > N] \geq \delta$. But this gives $\mathbb{P}[\int_0^\infty D_t \, \mathrm{d}t > N] \geq \delta$ for all N and hence $\mathbb{P}[\int_0^\infty D_t \, \mathrm{d}t = \infty] > 0$. But this violates Assumption 2.7 since $\lim_{t \to \infty} e^{\kappa t} D_t = 0$ $\mathbb{P}$-almost surely for some $\kappa > 0$. Thus $\mathbb{Q}[\int_0^\infty \Delta_t \, \mathrm{d}t < \infty] = 1$, finishing the proof. $\qquad\square$

With all the above lemmas, the proof of Theorem 3.4 is now given.

Proof of Theorem 3.4 We again adopt the notation in (9.12). In view of Lemma 9.1, the remaining statement in Theorem 3.4 which must be proved is that there is a set $\Omega_0 \in \mathcal{F}_\infty$ with $\mathbb{Q}[\Omega_0] = 1$ such that (3.13) holds, i.e.,

$$\omega \in \Omega_0 \quad \Longrightarrow \quad \lim_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle(\omega) = \langle h, \pi \rangle, \quad \text{for all } x > 0, h \in C_b(F; \mathbb{R}).$$

Recall the definition of $\mathcal{C}$ from (9.10) and let $h \in C_b(F; \mathbb{R}) \cap \mathcal{C}$. In view of Lemma 9.7, there is a set $\Omega_h \in \mathcal{F}_\infty$ such that $\mathbb{Q}[\Omega_h] = 1$ and

$$\omega \in \Omega_h \quad \Longrightarrow \quad \lim_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle(\omega) = \langle h, \pi \rangle, \quad \text{for all } x > 0.$$

Let the (countable) subset $\tilde{\mathcal{C}} \subset \mathcal{C}$ be as in the technical Lemma A.2 below and set $\Omega_0 = \bigcap_{h \in \tilde{\mathcal{C}}} \Omega_h$. Clearly, $\mathbb{Q}[\Omega_0] = 1$. Let $\omega \in \Omega_0$ and $h \in C_b(F; \mathbb{R})$ with $C = \sup_{y \in F} |h(y)|$. Let $\varepsilon > 0$ and for $n \geq 5$ take as in Lemma A.2 $^\uparrow\phi_{m,k}^n, {^\downarrow}\phi_{m,k}^n$ and θ^n such that (A.7) holds. In what follows, the ω will be suppressed, but all evaluations are understood to hold for this ω.

Let $x > 0$. With ν from (A.7) equal to $\hat{\pi}_T^x$, it follows that

$$\langle {^\uparrow}\phi_{m,k}^n, \hat{\pi}_T^x \rangle - 2C\langle 1 - \theta^{n-4}, \hat{\pi}_T^x \rangle - 2\varepsilon \leq \langle h, \hat{\pi}_T^x \rangle$$
$$\leq \langle {^\downarrow}\phi_{m,k}^n, \hat{\pi}_T^x \rangle + 2C\langle 1 - \theta^{n-4}, \hat{\pi}_T^x \rangle + 2\varepsilon.$$

With ν from (A.7) equal to π, one obtains

$$\langle {^\uparrow}\phi_{m,k}^n, \pi \rangle - 2C\langle 1 - \theta^{n-4}, \pi \rangle - 2\varepsilon \leq \langle h, \pi \rangle \leq \langle {^\downarrow}\phi_{m,k}^n, \pi \rangle + 2C\langle 1 - \theta^{n-4}, \pi \rangle + 2\varepsilon.$$

Putting these two together yields

$$\langle h, \hat{\pi}_T^x \rangle - \langle h, \pi \rangle \geq \langle {^\uparrow}\phi_{m,k}^n, \hat{\pi}_T^x \rangle - 2C\langle 1 - \theta^{n-4}, \hat{\pi}_T^x \rangle - 2\varepsilon$$
$$- \left(\langle {^\downarrow}\phi_{m,k}^n, \pi \rangle + 2C\langle 1 - \theta^{n-4}, \pi \rangle + 2\varepsilon \right)$$
$$= \langle {^\uparrow}\phi_{m,k}^n, \hat{\pi}_T^x \rangle - \langle {^\downarrow}\phi_{m,k}^n, \pi \rangle$$
$$- 2C\left(\langle 1 - \theta^{n-4}, \hat{\pi}_T^x \rangle + \langle 1 - \theta^{n-4}, \pi \rangle \right) - 4\varepsilon.$$

Since $\theta^{n-4}, {^\uparrow}\phi_{m,k}^n, {^\downarrow}\phi_{m,k}^n \in \tilde{\mathcal{C}} \subset \mathcal{C}$, taking $T \to \infty$ gives

$$\liminf_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle - \langle h, \pi \rangle \geq \langle {^\uparrow}\phi_{m,k}^n, \pi \rangle - \langle {^\downarrow}\phi_{m,k}^n, \pi \rangle - 4C\langle 1 - \theta^{n-4}, \pi \rangle - 4\varepsilon.$$

Now by Lemma A.2, for fixed m, n, the functions $^\uparrow\phi_{m,k}^n$ and $^\downarrow\phi_{m,k}^n$ are increasing and decreasing, respectively, in k and such that both a) $\lim_{k \to \infty} {^\downarrow}\phi_{m,k}^n(y) - {^\uparrow}\phi_{m,k}^n(y) = 0$ for $y \in \bar{F}_{n-2}$, and b) $|{^\uparrow}\phi_{m,k}^n(y) - {^\uparrow}\phi_{m,k}^n(y)| \leq 2C + 2\varepsilon$ for all $y \in F$ and n, m, k. Therefore, taking $k \to \infty$ in the above and using the monotone convergence theorem,

we obtain

$$\liminf_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle - \langle h, \pi \rangle \geq -2(C + \varepsilon)\pi[\bar{F}_{n-2}^c] - 4C\langle 1 - \theta^{n-4}, \pi \rangle - 4\varepsilon.$$

From Lemma A.2, we know that $0 \leq \theta^n(y) \leq 1$, $\lim_{n \to \infty} \theta^n(y) = 1$ for all $y \in F$. Thus, by the bounded convergence theorem and the fact that π is tight in F, it follows by taking $n \uparrow \infty$ that

$$\liminf_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle - \langle h, \pi \rangle \geq -4\varepsilon.$$

Taking $\varepsilon \downarrow 0$ gives that $\liminf_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle - \langle h, \pi \rangle \geq 0$. Thus, we have just shown for $\omega \in \Omega_0$, $x > 0$ and $h \in C_b(F; \mathbb{R})$ that

$$\liminf_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle(\omega) - \langle h, \pi \rangle \geq 0.$$

By applying the above to $\hat{h} = -h \in C_b(F; \mathbb{R})$, we see that

$$\limsup_{T \to \infty} \langle h, \hat{\pi}_T^x \rangle(\omega) - \langle h, \pi \rangle \leq 0,$$

which finishes the proof. $\qquad\square$

Appendix: Some technical results

Lemma A.1 *Let Assumption* 2.7 *hold, and additionally assume that* $|\eta| > 0$ *and* $f \in C^2(E; \mathbb{R}_+)$. *Recall F from* (3.1) *and the invariant density p for Z. Let* $h \in C^2(F)$ *be given and set*

$$\phi(z, x) := p(z)h(z, x), \qquad \psi(z, x) := \int_0^x h(z, y) \, \mathrm{d}y.$$

Let the operator L be as in (3.4) *and the operator* $L^R = A^{ij}\partial_{ij}^2 + (b^R)^i\partial_i$ *as in the proof of Proposition* 9.2, *where A is from* (3.3) *and* b^R *is from* (9.5). *Let* $\tilde{L}^R$ *be the formal adjoint of* L^R. *Then* $\tilde{L}^R\phi = p \, \partial_x(L\psi)$. *In particular, if* $L\psi = 0$, *then* $\tilde{L}^R\phi = 0$.

Proof For notational ease, the arguments will be suppressed when writing functions except for the x appearing in the drifts and volatilities of the operators. Now, recall

that the dynamics for the reversed process (ζ, χ) in (9.2) are

$$d\zeta_t = \left(c\frac{\nabla p}{p} + \operatorname{div} c - m\right)(\zeta_t)\,dt + \sigma(\zeta_t)\,dW_t,$$

$$d\chi_t = \left(f(\zeta_t) - \chi_t\left(a - \theta'c\frac{\nabla p}{p} - \nabla\cdot(c\theta)\right)(\zeta_t)\right)dt$$
$$+ \chi_t\left(\theta'c(\zeta_t)\,dW_t + \eta(\zeta_t)'\,dB_t\right),$$

and note, as mentioned in the proof of Proposition 9.2, that L^R is the generator for (ζ, χ). To further simplify the calculations, set

$$\xi := \frac{1}{2}\left(c\frac{\nabla p}{p} + \operatorname{div} c\right) - m \tag{A.1}$$

and

$$H(c, \theta) := \nabla\cdot(c\theta) - \theta'\operatorname{div} c.$$

Note that by (8.2), it follows that $0 = \nabla\cdot(p\xi)$. With this notation, we have that

$$d\zeta_t = (m + 2\xi)(\zeta_t)\,dt + \sigma(\zeta_t)\,dW_t,$$

$$d\chi_t = \left(f(\zeta_t) - \chi_t(a - 2\theta'(m + \xi) - H(c, \theta))(\zeta_t)\right)dt$$
$$+ \chi_t\left(\theta'c(\zeta_t)\,dW_t + \eta(\zeta_t)'\,dB_t\right),$$

which in turn yields that

$$A = \begin{pmatrix} c & xc\theta \\ x\theta'c & x^2(\theta'c\theta + \eta'\eta) \end{pmatrix},$$
$$b^R = \begin{pmatrix} m + 2\xi \\ f - x(a - 2\theta'(m + \xi) - H(c, \theta)) \end{pmatrix}, \tag{A.2}$$

along with

$$b = \begin{pmatrix} m \\ -f + x(a + \theta'c\theta + \eta'\eta) \end{pmatrix}.$$

Finally, multivariate notation will be used for derivatives with respect to z and univariate notation for derivatives with respect to x. Thus, for the given ϕ,

$$\nabla_{(z,x)}\phi = (\nabla\phi, \dot\phi), \qquad D^2_{(z,x)}\phi = \begin{pmatrix} D^2\phi & \nabla(\dot\phi) \\ \nabla(\dot\phi)' & \ddot\phi \end{pmatrix}.$$

Since $\phi = ph$ and p is not a function of x,

$$\nabla_{(z,x)}\phi = \begin{pmatrix} p\nabla h + h\nabla p \\ p\dot h \end{pmatrix}.$$

By definition, $\tilde{L}_R\phi = \nabla_{(z,x)} \cdot ((1/2)(A\nabla_{(z,x)}\phi + \phi\,\mathrm{div}_{(z,x)}A) - b^R\phi)$. Using (A.2),

$$A\nabla_{(z,x)}\phi = \begin{pmatrix} pc\nabla h + hc\nabla p + px\dot{h}c\theta \\ px\theta'c\nabla h + hx\theta'c\nabla p + px^2\dot{h}(\theta'c\theta + \eta'\eta) \end{pmatrix}.$$

Calculation shows that

$$\mathrm{div}_{(z,x)}A = \begin{pmatrix} \mathrm{div}\,c + c\theta \\ x\nabla \cdot (c\theta) + 2x(\theta'c\theta + \eta'\eta) \end{pmatrix},$$

so that

$$A\nabla_{(z,x)}\phi + \phi\,\mathrm{div}_{(z,x)}A$$

$$= \begin{pmatrix} pc\nabla h + hc\nabla p + px\dot{h}c\theta + ph\,\mathrm{div}\,c + phc\theta \\ px\theta'c\nabla h + hx\theta'c\nabla p + px^2\dot{h}(\theta'c\theta + \eta'\eta) + pxh\nabla \cdot (c\theta) + 2pxh(\theta'c\theta + \eta'\eta) \end{pmatrix}.$$

This gives $(1/2)(A\nabla_{(z,x)}\phi + \phi\,\mathrm{div}_{(z,x)}A) - b^R\phi = (\mathbf{A}, \mathbf{B})'$, where

$$\mathbf{A} = \frac{1}{2}(pc\nabla h + hc\nabla p + px\dot{h}c\theta + ph\,\mathrm{div}\,c + phc\theta) - phm - 2ph\xi,$$

$$\mathbf{B} = \frac{1}{2}\big(px\theta'c\nabla h + hx\theta'c\nabla p + px^2\dot{h}(\theta'c\theta + \eta'\eta)$$

$$+ pxh\nabla \cdot (c\theta) + 2pxh(\theta'c\theta + \eta'\eta)\big)$$

$$- phf + pxha - 2pxh\theta'(m + \xi) - pxhH(c, \theta). \tag{A.3}$$

Now, $\tilde{L}_R\phi = \nabla \cdot \mathbf{A} + \dot{\mathbf{B}}$. $\mathbf{A}$ is treated first. (A.1) gives $p\,\mathrm{div}\,c + c\nabla p = 2p(m + \xi)$ and hence

$$2\mathbf{A} = pc\nabla h + px\dot{h}c\theta + phc\theta - 2ph\xi.$$

For a scalar function f and $\mathbb{R}^d$-valued function g, $\nabla \cdot (fg) = f\nabla \cdot g + \nabla f'g$. Using this,

$$2\nabla \cdot \mathbf{A} = p\nabla \cdot (c\nabla h) + \nabla h'c\nabla p + px\dot{h}\nabla \cdot (c\theta) + x\nabla(p\dot{h})'c\theta$$

$$+ ph\nabla \cdot (c\theta) + \nabla(ph)'c\theta - 2h\nabla \cdot (p\xi) - 2p\nabla h'\xi$$

$$= p\nabla \cdot (c\nabla h) + \nabla h'c\nabla p + px\dot{h}\nabla \cdot (c\theta) + px\nabla(\dot{h})'c\theta + x\dot{h}\nabla p'c\theta$$

$$+ ph\nabla \cdot (c\theta) + p\nabla h'c\theta + h\nabla p'c\theta - 2h\nabla \cdot (p\xi) - 2p\nabla h'\xi.$$

Using that $\nabla \cdot (c\nabla h) = \mathrm{tr}(cD^2h) + \nabla h'\,\mathrm{div}\,c$ and collecting terms by derivatives of h gives

$$2\nabla \cdot \mathbf{A} = p\,\mathrm{tr}(cD^2h) + px\nabla(\dot{h})'c\theta + \nabla h'(p\,\mathrm{div}\,c + c\nabla p + pc\theta - 2p\xi)$$

$$+ \dot{h}\big(px\nabla \cdot (c\theta) + x\nabla p'c\theta\big) + h\big(p\nabla \cdot (c\theta) + \nabla p'c\theta - 2\nabla \cdot (p\xi)\big).$$

Since $p \operatorname{div} c + c \nabla p = 2p(m + \xi)$, $\nabla \cdot (p\xi) = 0$ and $\nabla \cdot (c\theta) = H(c, \theta) + \theta' \operatorname{div} c$,

$$p \operatorname{div} c + c \nabla p + pc\theta - 2p\xi = 2pm + pc\theta,$$

$$px \nabla \cdot (c\theta) + x \nabla p' c\theta = 2px\theta'(m + \xi) + px H(c, \theta),$$

$$p \nabla \cdot (c\theta) + \nabla p' c\theta - 2 \nabla \cdot (p\xi) = 2p\theta'(m + \xi) + p H(c, \theta).$$

Plugging this in and factoring out the p yields

$$\frac{2}{p} \nabla \cdot \mathbf{A} = \operatorname{tr}(cD^2 h) + x \nabla(\dot{h})' c\theta + \nabla h'(2m + c\theta) + \dot{h}\big(2x\theta'(m + \xi) + x H(c, \theta)\big)$$

$$+ h\big(2\theta'(m + \xi) + H(c, \theta)\big). \tag{A.4}$$

We next turn to $\mathbf{B}$ in (A.3). Using the relations $p \operatorname{div} c + c \nabla p = 2p(m + \xi)$ and $\nabla \cdot (c\theta) = H(c, \theta) + \theta' \operatorname{div} c$ yields

$$2\mathbf{B} = px\theta' c \nabla h - 2pxh\theta'(m + \xi) + px^2 \dot{h}(\theta' c\theta + \eta' \eta) + 2pxh(\theta' c\theta + \eta' \eta)$$

$$- 2phf + 2pxha - pxh H(c, \theta).$$

Since only h depends upon x,

$$2\dot{\mathbf{B}} = p\theta' c \nabla h + px \nabla(\dot{h})' c\theta - 2ph\theta'(m + \xi) - 2pxh\theta'(m + \xi)$$

$$+ 2px\dot{h}(\theta' c\theta + \eta' \eta) + px^2 \ddot{h}(\theta' c\theta + \eta' \eta) + 2ph(\theta' c\theta + \eta' \eta)$$

$$+ 2px\dot{h}(\theta' c\theta + \eta' \eta) - 2p\dot{h} f$$

$$+ 2pha + 2px\dot{h}a - ph H(c, \theta) - px\dot{h} H(c, \theta).$$

Grouping terms by derivatives of h and factoring out the p yields

$$\frac{2}{p} \dot{\mathbf{B}} = x\ddot{h}(\theta' c\theta + \eta' \eta) + x \nabla(\dot{h})' c\theta$$

$$+ h\big(-2\theta'(m + \xi) + 2(\theta' c\theta + \eta' \eta) + 2a - h H(c, \theta)\big) + \nabla h' c\theta$$

$$+ \dot{h}\big(-2x\theta'(m + \xi) + 4x(\theta' c\theta + \eta' \eta) - 2f + 2xa - x H(c, \theta)\big). \tag{A.5}$$

Putting together (A.4) and (A.5) and using that $\tilde{L}_R \phi = \nabla \cdot \mathbf{A} + \dot{\mathbf{B}}$, we obtain

$$\frac{1}{p} \tilde{L}_R \phi = \frac{1}{2} \operatorname{tr}(cD^2 h) + x \nabla(\dot{h})' c\theta + \frac{1}{2} x^2 \ddot{h}(\theta' c\theta + \eta' \eta) + \nabla h'(m + c\theta)$$

$$+ \dot{h}\big(2x(\theta' c\theta + \eta' \eta) - f + xa\big) + h(\theta' c\theta + \eta' \eta + a). \tag{A.6}$$

Turning now to ψ, since

$$L\psi = \frac{1}{2} \operatorname{tr}(cD^2 \psi) + x \nabla(\dot{\psi})' c\theta + \frac{1}{2} x^2 \ddot{\psi}(\theta' c\theta + \eta' \eta) + \nabla \psi' m$$

$$+ \dot{\psi}\big(-f + xa + x(\theta' c\theta + \eta' \eta)\big),$$

it follows that (note that only ψ depends upon x and $\dot{\psi} = h$)

$$
\begin{aligned}
L\dot{\psi} = {} & \frac{1}{2}\mathrm{tr}(cD^2\dot{\psi}) + x\nabla(\ddot{\psi})'c\theta + \nabla(\dot{\psi})'c\theta + x\ddot{\psi}(\theta'c\theta + \eta'\eta) \\
& + \frac{1}{2}x^2\dddot{\psi}(\theta'c\theta + \eta'\eta) + \nabla(\dot{\psi})'m + \ddot{\psi}\bigl(-f + xa + x(\theta'c\theta + \eta'\eta)\bigr) \\
& + \dot{\psi}(a + \theta'c\theta + \eta'\eta) \\
= {} & \frac{1}{2}\mathrm{tr}(cD^2\dot{\psi}) + x\nabla(\ddot{\psi})'c\theta + \frac{1}{2}x^2\dddot{\psi}(\theta'c\theta + \eta'\eta) \\
& + \nabla(\dot{\psi})'(m + c\theta) + \ddot{\psi}\bigl(2x(\theta'c\theta + \eta'\eta) - f + xa\bigr) + \dot{\psi}(a + \theta'c\theta + \eta'\eta) \\
= {} & \frac{1}{2}\mathrm{tr}(cD^2 h) + x\nabla(\dot{h})'c\theta + \frac{1}{2}x^2\ddot{h}(\theta'c\theta + \eta'\eta) \\
& + \nabla h'(m + c\theta) + \dot{h}\bigl(2x(\theta'c\theta + \eta'\eta) - f + xa\bigr) + h(a + \theta'c\theta + \eta'\eta).
\end{aligned}
$$

But from (A.6), this last term is precisely $(1/p)\tilde{L}_R\phi$. $\qquad\square$

Lemma A.2 *Let Assumption 2.7 hold. Let $\mathcal{C}$ be as in (9.10). Recall that we have set $F = E \times (0, \infty)$ and let $(F_n)_{n\in\mathbb{N}}$ be a family of open, bounded, increasing subsets of F with smooth boundary such that $F = \bigcup_n F_n$. There exists a countable family of functions*

$$
\tilde{\mathcal{C}} := \bigl\{ {}^{\uparrow}\phi^n_{m,k}, {}^{\downarrow}\phi^n_{m,k}, \theta^n : n, m, k \in \mathbb{N}, n \geq 3 \bigr\} \subset \mathcal{C}
$$

such that

(i) *for each $n \geq 3$, $0 \leq \theta^n \leq 1$ with $\theta^n = 1$ on $\bar{F}_n$ and $\theta^n = 0$ on F_{n+1}^c;*

(ii) *for each $n \geq 3$ and m, the functions ${}^{\uparrow}\phi^n_{m,k}$ are increasing in k and the functions ${}^{\downarrow}\phi^n_{m,k}$ are decreasing in k. Furthermore, for any $n \geq 3$ and m, we have that $\lim_{k\to\infty} |{}^{\uparrow}\phi^n_{m,k}(y) - {}^{\downarrow}\phi^n_{m,k}(y)| = 0$ for $y \in \bar{F}_{n-2}$.*

Additionally, for any $h \in C_b(F; \mathbb{R})$, set $C = C(h) := \sup_{y\in F} |h(y)|$. Then, for any $\varepsilon > 0$ and any integer $n \geq 5$, there exists an integer $m = m(\varepsilon, n)$ such that for all $k \in \mathbb{N}$, $\sup_{y\in F} |{}^{\uparrow}\phi^n_{m,k}(y)| \leq C + \varepsilon$, $\sup_{y\in F} |{}^{\downarrow}\phi^n_{m,k}(y)| \leq C + \varepsilon$. Furthermore, for any Borel measure ν on F,

$$
\int_F {}^{\uparrow}\phi^n_{m,k}\, d\nu - 2C \int_F (1 - \theta^{n-4})\, d\nu - 2\varepsilon
$$

$$
\leq \int_F h\, d\nu
$$

$$
\leq \int_F {}^{\downarrow}\phi^n_{m,k}\, d\nu + 2C \int_F (1 - \theta^{n-4})\, d\nu + 2\varepsilon. \tag{A.7}
$$

Proof Fix $n \in \mathbb{N}$ and let $(\phi_m^n)_{m \in \mathbb{M}}$ be a countable dense (with respect to the supremum norm) subset of $C_b(\bar{F}_n; \mathbb{R})$. Now, let $k \in \mathbb{N}$ and define for $y \in \bar{F}_n$

$$
\begin{aligned}
{}^{\uparrow}\tilde{\phi}_{m,k}^n(y) &:= \inf_{y_0 \in \bar{F}_n} \left(\phi_m^n(y_0) + k|y - y_0| \right), \\
{}^{\downarrow}\tilde{\phi}_{m,k}^n(y) &:= \sup_{y_0 \in \bar{F}_n} \left(\phi_m^n(y_0) - k|y - y_0| \right).
\end{aligned} \tag{A.8}
$$

As shown in [2, Chap. 3.4], ${}^{\uparrow}\tilde{\phi}_{m,k}^n$ and ${}^{\downarrow}\tilde{\phi}_{m,k}^n$ are increasing and decreasing, respectively, in k, and Lipschitz-continuous in $\bar{F}_n$ with Lipschitz constant k. Furthermore, as $k \uparrow \infty$, ${}^{\uparrow}\tilde{\phi}_{m,k}^n \nearrow \phi_m^n$ and ${}^{\downarrow}\phi_{m,k}^n \searrow \phi_m^n$ on $\bar{F}_n$.

Next, let $\theta^n \in C^\infty(F; \mathbb{R})$ be such that $0 \leq \theta^n \leq 1$, $\theta^n(y) = 1$ on $\bar{F}_n$ and $\theta^n(y) = 0$ on F_{n+1}^c. Clearly, $\theta^n \in \mathcal{C}$ for each n. Now assume $n \geq 3$ and extend ${}^{\uparrow}\tilde{\phi}_{m,k}^n$ and ${}^{\downarrow}\tilde{\phi}_{m,k}^n$ from functions on $\bar{F}_n$ to all of F via

$$
{}^{\uparrow}\phi_{m,k}^n(y) = \begin{cases} {}^{\uparrow}\tilde{\phi}_{m,k}^n(y)\theta^{n-2}(y) & y \in \bar{F}_n, \\ 0 & \text{else}, \end{cases}
$$

$$
{}^{\downarrow}\phi_{m,k}^n(y) = \begin{cases} {}^{\downarrow}\tilde{\phi}_{m,k}^n(y)\theta^{n-2}(y) & y \in \bar{F}_n, \\ 0 & \text{else}. \end{cases}
$$

Clearly, ${}^{\uparrow}\phi_{m,k}^n$ and ${}^{\downarrow}\phi_{m,k}^n$ are Lipschitz on F, and since F_n is bounded, it also holds that ${}^{\uparrow}\phi_{m,k}^n$ and ${}^{\downarrow}\phi_{m,k}^n$ are in $\mathcal{C}$. Note also that ${}^{\uparrow}\phi_{m,k}^n$ and ${}^{\downarrow}\phi_{m,k}^n$ increase and decrease, respectively, as $k \uparrow \infty$ to a function which is equal to ϕ_m^n on $\bar{F}_{n-2}$ and that ${}^{\uparrow}\phi_{m,k}^n, {}^{\downarrow}\phi_{m,k}^n$ are bounded on all of F by $\sup_{y \in \bar{F}_n} |{}^{\uparrow}\tilde{\phi}_{m,k}^n(y)|$ and $\sup_{y \in \bar{F}_n} |{}^{\downarrow}\tilde{\phi}_{m,k}^n(y)|$, respectively. This proves (i) and (ii) above.

Now, let $h \in C_b(F; \mathbb{R})$ with $C = \sup_{y \in F} |h(y)|$. Let $\varepsilon > 0$ and for $n \geq 5$, choose $m = m(\varepsilon, n)$ so that $\sup_{y \in \bar{F}_n} |h(y) - \phi_m^n(y)| \leq \varepsilon$. By the construction of ${}^{\uparrow}\tilde{\phi}_{m,k}^n$ in (A.8), it follows for each k that

$$
-(C + \varepsilon) \leq \inf_{y_0 \in \bar{F}_n} \phi_m^n(y_0) \leq {}^{\uparrow}\tilde{\phi}_{m,k}^n(y) \leq \phi_m^n(y) \leq h(y) + \varepsilon \leq C + \varepsilon, \qquad y \in \bar{F}_n.
$$

By the definition of ${}^{\uparrow}\phi_{m,k}^n$, this gives $\sup_{y \in F} |{}^{\uparrow}\phi_{m,k}^n(y)| \leq C + \varepsilon$. Furthermore, as $\theta^{n-2}(y) = 1$ on $\bar{F}_{n-2}$, we have $h(y) \geq {}^{\uparrow}\phi_{m,k}^n(y) - \varepsilon$ on $\bar{F}_{n-2}$. Therefore, for any Borel measure ν, using the notation in (9.12),

$$
\begin{aligned}
\langle h, \nu \rangle &\geq \langle ({}^{\uparrow}\phi_{m,k}^n - \varepsilon) 1_{\bar{F}_{n-2}}, \nu \rangle - C\nu[\bar{F}_{n-2}^c] \\
&\geq \langle {}^{\uparrow}\phi_{m,k}^n, \nu \rangle - \langle {}^{\uparrow}\phi_{m,k}^n 1_{\bar{F}_{n-2}^c}, \nu \rangle - \varepsilon - C\nu[\bar{F}_{n-2}^c] \\
&\geq \langle {}^{\uparrow}\phi_{m,k}^n, \nu \rangle - (C + \varepsilon)\nu[\bar{F}_{n-2}^c] - \varepsilon - C\nu[\bar{F}_{n-2}^c] \\
&\geq \langle {}^{\uparrow}\phi_{m,k}^n, \nu \rangle - 2C\nu[\bar{F}_{n-2}^c] - 2\varepsilon \\
&\geq \langle {}^{\uparrow}\phi_{m,k}^n, \nu \rangle - 2C \int_F (1 - \theta^{n-4}) \, d\nu - 2\varepsilon,
\end{aligned}
$$

where the last inequality follows since $1_{\bar{F}^c_{n-2}(y)} \leq 1 - \theta^{n-4}(y)$. This gives the lower bound in (A.7). A similar calculation shows for all k that

$$-(C+\varepsilon) \leq h(y) - \varepsilon \leq \phi^n_m(y) \leq {}^\downarrow\tilde{\phi}^n_{m,k}(y) \leq \sup_{y_0 \in \bar{F}_n} \phi^n_m(y_0) \leq C+\varepsilon, \qquad y \in \bar{F}_n.$$

This gives $\sup_{y \in F} |{}^\downarrow\phi^n_{m,k}(y)| \leq C+\varepsilon$ and $h(y) \leq {}^\downarrow\phi^n_{m,k}(y) + \varepsilon$ on $\bar{F}_{n-2}$. Thus

$$\begin{aligned}
\langle h, \nu \rangle &\leq \left\langle \left({}^\downarrow\phi^n_{m,k} + \varepsilon\right)1_{\bar{F}_{n-2}}, \nu \right\rangle + C\nu[\bar{F}^c_{n-2}] \\
&\leq \left\langle {}^\downarrow\phi^n_{m,k}, \nu \right\rangle - \left\langle {}^\downarrow\phi^n_{m,k}1_{\bar{F}^c_{n-2}}, \nu \right\rangle + \varepsilon + C\nu[\bar{F}^c_{n-2}] \\
&\leq \left\langle {}^\downarrow\phi^n_{m,k}, \nu \right\rangle + (C+\varepsilon)\nu[\bar{F}^c_{n-2}] + \varepsilon + C\nu[\bar{F}^c_{n-2}] \\
&\leq \left\langle {}^\downarrow\phi^n_{m,k}, \nu \right\rangle + 2C\nu[\bar{F}^c_{n-2}] + 2\varepsilon \\
&\leq \left\langle {}^\downarrow\phi^n_{m,k}, \nu \right\rangle + 2C \int_F (1 - \theta^{n-4}) \, d\nu + 2\varepsilon.
\end{aligned}$$

Therefore, the upper bound in (A.7) is established. $\qquad\square$

References

1. Abou-Kandil, H., Freiling, G., Ionescu, V., Jank, G.: Matrix Riccati Equations in Control and Systems Theory. Birkhäuser Verlag, Basel (2003)
2. Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis: A Hitchhiker's Guide, 3rd edn. Springer, Berlin (2006)
3. Anderson, t.: Reverse-time diffusion equation models. Stoch. Process. Appl. **12**, 313–326 (1982)
4. Arnold, L., Kliemann, W.: Qualitative theory of stochastic systems. In: Bharucha-Reid, A.T. (ed.) Probabilistic Analysis and Related Topics, vol. 3, pp. 1–79. Academic Press, New York (1983)
5. Bhattacharya, R.N.: Criteria for recurrence and existence of invariant measures for multidimensional diffusions. Ann. Probab. **6**, 541–553 (1978)
6. Carmona, P., Petit, F., Yor, M.: Exponential functionals of Lévy processes. In: Barndorff-Nielsen, O.E., et al. (eds.) Lévy Processes, pp. 41–55. Birkhäuser Boston, Boston (2001)
7. Castañon, D.A.: Reverse-time diffusion processes. IEEE Trans. Inf. Theory **28**, 953–956 (1982)
8. Chen, M.F., Li, S.F.: Coupling methods for multidimensional diffusion processes. Ann. Probab. **17**, 151–177 (1989)
9. De Schepper, A., Goovaerts, M., Delbaen, F.: The Laplace transform of annuities certain with exponential time distribution. Insur. Math. Econ. **11**, 291–294 (1992)
10. Delbaen, F.: Consols in the CIR model. Math. Finance **3**, 125–134 (1993)
11. Dufresne, D.: The distribution of a perpetuity, with applications to risk theory and pension funding. Scand. Actuar. J. **1990**, 39–79 (1990)
12. Dufresne, D.: Distributions of discounted values. Actuarial Research Clearing House **1**, 11–24 (1992)
13. Elliott, R.J., Anderson, B.D.O.: Reverse time diffusions. Stoch. Process. Appl. **19**, 327–339 (1985)
14. Embrechts, P., Goldie, C.M.: Perpetuities and random equations. In: Mandl, P., Hušková, M. (eds.) Asymptotic Statistics, Prague, 1993. Contrib. Statist., pp. 75–86. Physica, Heidelberg (1994)
15. Gilbarg, D., Trudinger, N.S.: Elliptic Partial Differential Equations of Second Order. Classics in Mathematics. Springer, Berlin (2001). Reprint of the 1998 edition
16. Gjessing, H.K., Paulsen, J.: Present value distributions with applications to ruin theory and stochastic equations. Stoch. Process. Appl. **71**, 123–144 (1997)
17. Goldie, C.M., Maller, R.A.: Stability of perpetuities. Ann. Probab. **28**, 1195–1218 (2000)
18. Guasoni, P., Kardaras, C., Robertson, S., Xing, H.: Abstract, classic, and explicit turnpikes. Finance Stoch. **18**, 75–114 (2014)

19. Haussmann, U.G., Pardoux, É.: Time reversal of diffusions. Ann. Probab. **14**, 1188–1205 (1986)
20. Heath, D., Schweizer, M.: Martingales versus PDEs in finance: an equivalence result with examples. J. Appl. Probab. **37**, 947–957 (2000)
21. Nilsen, T., Paulsen, J.: On the distribution of a randomly discounted compound Poisson process. Stoch. Process. Appl. **61**, 305–310 (1996)
22. Nualart, D.: The Malliavin Calculus and Related Topics, 2nd edn. Springer, Berlin (2006)
23. Pardoux, É.: Smoothing of a diffusion process conditioned at final time. In: Kohlmann, M., Christopeit, N. (eds.) Stochastic Differential Systems, Bad Honnef, 1982. Lecture Notes in Control and Inform. Sci., vol. 43, pp. 187–196. Springer, Berlin (1982)
24. Paulsen, J.: Risk theory in a stochastic economic environment. Stoch. Process. Appl. **46**, 327–361 (1993)
25. Paulsen, J.: Present value of some insurance portfolios. Scand. Actuar. J. **1997**, 11–37 (1997)
26. Paulsen, J., Hove, A.: Markov chain Monte Carlo simulation of the distribution of some perpetuities. Adv. Appl. Probab. **31**, 112–134 (1999)
27. Pinsky, R.G.: Positive Harmonic Functions and Diffusion. Cambridge Studies in Advanced Mathematics, vol. 45. Cambridge University Press, Cambridge (1995)
28. Veretennikov, A.Y.: Bounds for the mixing rate in the theory of stochastic equations. Theory Probab. Appl. **32**, 273–281 (1984)
29. Veretennikov, A.Y.: Estimates for the mixing rate for Markov processes. Liet. Mat. Rink. **31**, 40–49 (1991)
30. Vervaat, W.: On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. Adv. Appl. Probab. **11**, 750–783 (1979)
31. Yor, M.: Exponential Functionals of Brownian Motion and Related Processes. Springer, Berlin (2001)