Superdiffusive limits for Bessel-driven stochastic kinetics

Miha Brešar* Conrado da Costa† Aleksandar Mijatović*‡ Andrew Wade†

6th February 2025

Abstract

We prove anomalous-diffusion scaling for a one-dimensional stochastic kinetic dynamics, in which the stochastic drift is driven by an exogenous Bessel noise, and also includes endogenous volatility which is permitted to have arbitrary dependence with the exogenous noise. We identify the superdiffusive scaling exponent for the model, and prove strong and weak convergence results on the corresponding scale. We show how our result extends to admit, as exogenous noise processes, not only Bessel processes but more general processes satisfying certain asymptotic conditions.

Key words: Stochastic kinetic dynamics; anomalous diffusion; superdiffusivity; Bessel process; additive functional.

AMS Subject Classification: 60J60 (Primary), 60K50 (Secondary).

1 Introduction and main results

The subject of this work is the long-term behaviour of an Itô process $X = (X_t)_{t\in\mathbb{R}_+}$ in $\mathbb{R}_+ := [0,\infty)$ with a representation

$$X_t = X_0 + \int_0^t \frac{f(s,Y_s)}{X_s}\,\mathrm{d}s + B_t, \qquad X_0 > 0, \tag{1.1}$$

where the process $Y = (Y_t)_{t\in\mathbb{R}_+}$ is adapted to the same filtration as the Brownian motion (BM) $B = (B_t)_{t\in\mathbb{R}_+}$. Under suitable assumptions on the function $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ and the exogenous noise process Y, which drives the *stochastic drift* of X, the process X will exhibit *anomalous diffusion* (see YouTube [8] for a short overview of our results). Physical motivation includes dynamics of particles interacting with an external field or medium, or with an internal relaxation mechanism; see e.g. [6, 15, 16] and §2 below for further discussion of motivation and related literature.

We assume the function $f(t,y)$ that contributes to the drift via (1.1) has certain polynomial asymptotic growth behaviour for large t and y: rougly speaking, that $f(t,y) \sim \rho t^\gamma y^\alpha$ as both $t,y\to\infty$, for a constant $\rho>0$. In fact, the hypothesis is a little stronger than this (giving some comparable control for fixed t and large y, for example):

($\mathbf{A}_f$) Suppose that $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, and that for some constant $\rho \in (0,\infty)$ and exponents $\alpha\in\mathbb{R}_+$ and $\gamma\in(-\alpha,\infty)$ (i.e. $\alpha+\gamma>0$) satisfies the following. For every $\varepsilon>0$, there is some $r_\varepsilon\in\mathbb{R}_+$ such that

$$\sup_{(t,y)\in\mathbb{R}_+^2 :\, t+y\ge r_\varepsilon} \left| f(t,y)(1+t)^{-\gamma}(1+y)^{-\alpha} - \rho \right| \le \varepsilon. \tag{1.2}$$

*Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK.
Email: `miha.bresar@warwick.ac.uk`, `a.mijatovic@warwick.ac.uk`

†Department of Mathematical Sciences, Durham University, Upper Mountjoy, Durham, DH1 3LE, UK.
Email: `conrado.da-costa@durham.ac.uk`, `andrew.wade@durham.ac.uk`.

‡The Alan Turing Institute.

The aim of this paper is to give some natural and robust growth and stability hypotheses on the exogenous noise process Y so that process X satisfying (1.1) exhibits superdiffusive asymptotic behaviour, as quantified via (i) a distributional scaling limit, and (ii) an a.s.-quantification of the transient growth exponent. Our general result is presented in Theorem 1.3 below. First we present a prototypical example, in which Y is a squared-Bessel process; this is Theorem 1.1. First we need some notation.

Denote by $\mathrm{BESQ}^\delta(y)$ the law of a squared-Bessel process Y with "dimension" parameter $\delta \in (0, \infty)$ started at an arbitrary $Y_0 = y \in \mathbb{R}_+$; this law can be defined as that of the solution of the stochastic differential equation (SDE) in (1.6) below. Recall that $\delta > 0$ ensures Y is not absorbed at 0; for $\delta \in (0, 2)$, the process Y is point recurrent over $\mathbb{R}_+$, while for $\delta \geq 2$ it is point transient (see [31, Ch. XI] for further details).

The following is our general result as applied to the case of squared-Bessel exogenous noise. Because of the apparent singularity in the drift (1.1), in the same way as one does for the Bessel process, it is more convenient to formulate the dynamics via "X squared", which is essentially the process S in (1.3): see Remark 1.2(ii) after the statement of the theorem. We denote convergence in distribution by '$\xrightarrow{\mathrm{d}}$'.

Theorem 1.1. *Let $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfy (A_f) and assume $\delta \in (0, \infty)$. Suppose the adapted process (S, Y, B) consists of an $\mathbb{R}$-valued Brownian motion B, a squared-Bessel process Y with law $\mathrm{BESQ}^\delta(y)$ (defined via SDE (1.6)) and an $\mathbb{R}_+$-valued process $S = (S_t)_{t\in\mathbb{R}_+}$, satisfying*

$$S_t = S_0 + \int_0^t \big(2f(s, Y_s) + 1\big)\,\mathrm{d}s + 2\int_0^t \sqrt{S_s}\,\mathrm{d}B_s, \text{ for all } t \in \mathbb{R}_+, \tag{1.3}$$

started at a deterministic level $S_0 \in \mathbb{R}_+$. Then the $\mathbb{R}_+$-valued process $X = (X_t)_{t\in\mathbb{R}_+}$, given by $X_t := \sqrt{S_t}$, $t \in \mathbb{R}_+$, has the following asymptotic properties.

(a) There is superdiffusive transience, namely,

$$\lim_{t\to\infty} \frac{\log X_t}{\log t} = \frac{1+\gamma+\alpha}{2}, \text{ a.s.} \tag{1.4}$$

(b) There is a distributional limit, namely,

$$t^{-(1+\gamma+\alpha)/2} X_t \xrightarrow{\mathrm{d}} \left(2\rho \int_0^1 s^\gamma \widetilde{Y}_s^\alpha\,\mathrm{d}s\right)^{1/2}, \text{ as } t \to \infty, \tag{1.5}$$

where $\widetilde{Y}$ in the limit follows $\mathrm{BESQ}^\delta(0)$.

Remarks 1.2. (i) The result (1.5) implies an "in probability" version of (1.4), but not the a.s. version. On the other hand, the almost-sure asymptotic (1.4) combined with a martingale decomposition and the self-similarity of Bessel process yield (1.5), so in some sense it is the rough a.s.-asymptotic (1.4) that is the key to the result. Our proof of Theorem 1.1 goes by first proving a more general result, Theorem 1.3 below, that identifies the robust aspects of the Bessel process that are essential for this behaviour, and hence extends the framework to admit a considerably larger class of processes: see §2.2 below for further remarks on the proofs.

(ii) Under a mild additional hypothesis, X defined in Theorem 1.1 satisfies dynamics (1.1), see Appendix A. In Theorem 1.1 it is more convenient to take as the primitive $X^2 = S$, because SDE (1.3) fully specifies the process for a general continuous function f (cf. Bessel process defined as a square-root of BESQ in [31, Ch. XI]).

(iii) Since $\gamma + \alpha > 0$ by (A_f), the scaling exponent $(1 + \gamma + \alpha)/2$ in (1.4)–(1.5) exceeds 1/2, making the process X *superdiffusive* and transient. Moreover, by (A_f), the values of f on any compact set play no role. The log-scale convergence in (1.4) is a compact formulation of the statement that for every small $\varepsilon > 0$, it holds that for all $t \in \mathbb{R}_+$ sufficiently large, $t^{(1+\gamma+\alpha)/2-\varepsilon} < X_t < t^{(1+\gamma+\alpha)/2+\varepsilon}$; put differently, $X_t = t^{(1+\gamma+\alpha)/2+o(1)}$ as $t \to \infty$, and, in particular, $\liminf_{t\to\infty} t^{\varepsilon} X_t / t^{(1+\gamma+\alpha)/2} = \infty$ for all $\varepsilon > 0$ (cf. Assumption (A_Y)(c) below).

(iv) The dimension parameter $\delta \in (0,\infty)$ of the squared-Bessel process Y does not appear in the scaling exponent in (1.5). The condition $\delta > 0$ ensures that the Bessel process $\sqrt{Y}$ is *diffusive*, which determines how the parameter α in (A_f) enters the scaling exponent in (1.4)–(1.5). In particular, X in Theorem 1.1 exhibits superdiffusive transience for negative $\gamma \in (-\alpha, 0)$, even if Y is topologically recurrent (i.e. $\delta \in (0,2]$).

(v) We emphasize that the hypotheses in Theorem 1.1 permit arbitrary dependence between the process Y and the driving Brownian motion B in (1.3). Since Y with law $\mathrm{BESQ}^{\delta}(y)$ is the unique strong solution of the stochastic differential equation

$$\mathrm{d}Y_t = \delta\,\mathrm{d}t + 2|Y_t|^{1/2}\,\mathrm{d}W_t, \text{ for } Y_0 = y \in \mathbb{R}_+, \tag{1.6}$$

for some Brownian motion $W = (W_t)_{t\in\mathbb{R}_+}$, the limit in (1.5) requires only that the Brownian motions W and B are adapted to the same filtration. In particular, W and B may be equal, independent or have arbitrary stochastically evolving (adapted) covariation.

(vi) In the case $\alpha = 0$, so that $f(t,y) \sim \rho t^{\gamma}$ as $t \to \infty$, the impact of the process Y on the large-scale dynamics of X vanishes as $t \to \infty$. In particular, the limit in Theorem 1.1 is deterministic, and, in fact, the convergence in distribution can in that case be strengthened to almost sure convergence. Theorem 1.1 can thus be viewed as a generalisation of certain results in [18] to non-polynomial time-inhomogeneous drift; see §2 below for some elaboration on this connection.

(vii) In contrast to Remark 1.2(iii), if we had $\gamma = \alpha = 0$ (and hence f asymptotically constant), then X would be (essentially) a Bessel process of dimension $1 + 2\rho > 1$. In that case, the statement (1.5) is false, because the distributional limit of $t^{-1/2}X_t$ is *random*, while the right-hand side of (1.5) is deterministic when $\alpha = 0$. This suggests that the condition $\gamma + \alpha > 0$ in (A_f) cannot be omitted, essentially because Theorem 1.1 requires the (exogenous) Y-driven stochastic drift to dominate the (endogenous) Brownian noise in the evolution of X.

(viii) We do not consider here the case where the process Y in (1.1) is *ergodic*, because we anticipate rather different phenomena in that case. For example, if $f(s,y) = h(y)$ depends only on y, and if h is integrable with respect to the stationary measure of Y on $\mathbb{R}_+$, then it seems natural to suspect that X behaves similarly to a Bessel process with drift coefficient given via the stationary mean of $h(Y)$. While interesting, this case seems unlikely to generate the anomalous diffusion that is the focus of the present work; see §2 below for more details.

Remark 1.2(viii) suggests that for an ergodic process Y, the model (S, Y) in (1.3) does not exhibit superdiffusivity. But what properties of Y do guarantee anomalous diffusive behaviour in Theorem 1.1? By Theorem 1.3 below, Assumption (A_Y) on the additive functional

$$A_t := \int_1^t s^{\gamma} Y_s^{\alpha}\,\mathrm{d}s, \text{ for } t \in [1,\infty), \tag{1.7}$$

ensures such behaviour. Theorem 1.1 then follows by proving Assumption (A_Y) for squared-Bessel processes and identifying the weak limit.

(A_Y) Let α and γ be as in (A_f).

(a) It holds that $\int_0^1 \mathbb{E}\big[Y_t^{\alpha}\big]\,\mathrm{d}t < \infty$.

Assume that A_t in (1.7) satisfies the following hypotheses.

(b) For a random variable $\widetilde{A}$ in $\mathbb{R}_+$, $A_t/t^{1+\gamma+\alpha} \xrightarrow{\mathrm{d}} \widetilde{A}$ as $t \to \infty$.

(c) For every $\varepsilon > 0$, it holds that $\lim_{t\to\infty} t^{\varepsilon} A_t / t^{1+\gamma+\alpha} = \infty$, a.s.

(d) It holds that $\sup_{t\in[1,\infty)} \mathbb{E}\, A_t / t^{1+\gamma+\alpha} < \infty$.

The next limit theorem requires no assumption beyond (A_Y) on the dynamics of Y.

Theorem 1.3. *Assume that* (A_f) *holds and that the adapted process* (S, Y, B) *consists of an* $\mathbb{R}$*-valued Brownian motion* B*, a process* Y *on* $\mathbb{R}_+$ *satisfying hypothesis* (A_Y)*, and an* $\mathbb{R}_+$*-valued process* S *satisfying SDE* (1.3)*. Then, for* $X_t = \sqrt{S_t}$*, it holds that*

$$\lim_{t\to\infty} \frac{\log X_t}{\log t} = \frac{1+\gamma+\alpha}{2}, \text{ a.s.,} \tag{1.8}$$

and, as $t \to \infty$*, for* $\widetilde{A}$ *the random variable in hypothesis* (A_Y)*(b),*

$$t^{-(1+\gamma+\alpha)/2} X_t \xrightarrow{\mathrm{d}} \big(2\rho \widetilde{A}\big)^{1/2}. \tag{1.9}$$

The proof of Theorem 1.3 is given in §3. The proof of Theorem 1.1 is presented in §4, where we establish that the process Y with law $\mathrm{BESQ}^{\delta}(y)$ satisfies Assumption (A_Y). In §2.2 we give an overview of the proof strategy, after first (in §2.1) describing some motivation and relevant literature. In indication of possible extensions, we remark that self-similarity of the law $\mathrm{BESQ}^{\delta}(y)$ is only used in the final step of the proof of Theorem 1.1 to identify the law of the limit $\widetilde{A}$. This leads to a natural open problem: To find and/or characterise self-similar processes Y, such that X satisfies analogue of the weak limit in (1.5), appropriately adjusted for the self-similarity index of Y.

2 Motivation and discussion

2.1 Motivation and literature

<u>An example motivated by a self-interacting random walk.</u> A discrete-time relative of the model in Theorem 1.1 is studied in [11], motivated in part by a programme to study a certain self-interacting planar random walk. The parameters in the present model that correspond, heuristically, to the process described in [11] are $\gamma = 0$, $\alpha = 1/2$, and $\delta = 1$; informally for Brownian motions B, W on $\mathbb{R}$ with arbitrary dependence, $\mathrm{d}X_t = \rho |W_t| / X_t \,\mathrm{d}t + \mathrm{d}B_t$. Note that $|W|$ (reflected BM) has the same law as $\sqrt{Y}$ where Y follows $\mathrm{BESQ}^1(0)$. By Theorem 1.1, as $t \to \infty$, $t^{-3/4} X_t$ converges weakly to the law of $(2\rho \int_0^1 |W_s| \,\mathrm{d}s)^{1/2}$, which can be expressed in terms of Airy functions [5, p. 349]. Figure 1 plots the graphs of a path $t \mapsto X_t$ and $t \mapsto t^{3/4}$.

<u>Stochastic kinetic dynamics.</u> Various physical systems motivate *stochastic kinetic models* for processes $(X, V) \in \mathbb{R}^d \times \mathbb{R}^d$ with dynamics of the form

$$X_t = \int_0^t V_s \,\mathrm{d}s, \quad V_t = \int_0^t G(V_s) \,\mathrm{d}s + B_t, \tag{2.1}$$

where B is d-dimensional Brownian motion and $G : \mathbb{R}^d \to \mathbb{R}^d$. Here V is an autonomous *velocity process* which feeds into the stochastic drift of the d-dimensional process X. Associated to (2.1) is the so-called *kinetic Fokker–Planck* equation [15, 16]. A classical example is Paul Langevin's 1908 work on the *confining* case $G(v) = -v$ (see [29]), which was proposed to model physical Brownian motion. In that case, V is a d-dimensional Ornstein–Uhlenbeck process, and, for some mean-zero Gaussian random variable $\overline{V}$,

$$V_t \xrightarrow{\mathrm{d}} \overline{V}, \text{ as } t \to \infty, \tag{2.2}$$

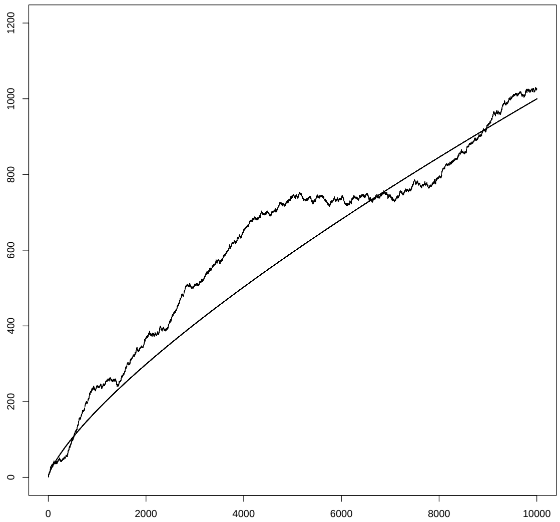

Figure 1: Simulated realizations (Euler scheme, step size of 1/10) of a trajectory $t \mapsto X_t$ of X (with $\gamma = 0$, $\alpha = 1/2$, i.e. $f(t, y) = \sqrt{y}$, and $\delta = 1$) and the graph $t \mapsto t^p$ for exponent $p = (1 + \alpha + \gamma)/2 = 3/4$ on the time interval $[0, T]$ with $T = 10^5$.

i.e., there is weak convergence to a stationary, isotropic velocity distribution. Since $\overline{V}$ is light-tailed and V is rapidly mixing, then X itself, defined through the additive functional in (2.1), will look like Brownian motion on large scales.

Renewed interest in the system (2.1), where $G(v)$ is of order $1/v$ for large v (so that V is a Bessel-like process), has been stimulated by both modelling of specific physical systems and realization that such processes can generate a variety of scaling behaviours corresponding to *anomalous diffusion*; we refer to [1, 6, 7, 15–17, 22, 24], and references therein.

The mechanism for anomalous diffusion explored in these works is different from ours. Indeed, these works maintain that G be confining, taking $G(v) = -\delta/(1 + v)$ (or variations on this) for $\delta > 0$. For appropriate δ, this leads to a *heavy-tailed* stationary distribution $\overline{V}$ for which (2.2) holds. In this case a competition between the mixing rate of V and the index of the domain of attraction in which $\overline{V}$ lives determines the asymptotics of the additive functional X defined through (2.1). In contrast, our model G is *positive* and we are interested in velocity processes V such that $V_t \to \infty$ in probability, as $t \to \infty$ (possibly recurrent, however). This gives a quite distinct mechanism for anomalous diffusion, and some different phenomena; for example, the range of scaling laws in [16, Thm 2] are on scale $t^{1/2}$ (diffusive) through to $t^{3/2}$ (super-ballistic), but our growth exponent in Theorem 1.1(a) has no upper bound in $(1/2, \infty)$.

A variation of (2.1) is to replace the dynamics of X by

$$\mathrm{d}X_t = \frac{f(t, Y_t)}{X_t} \, \mathrm{d}t, \tag{2.3}$$

for example, where Y is an autonomous process. Then $\mathrm{d}(X_t^2) = 2f(t, Y_t) \, \mathrm{d}t$, so, in some sense, the generalization that (2.3) provides to (2.1) is only in the form of the velocity process V_t. Our model (1.1) extends (2.3) further by introducing *exogenous noise* into the dynamics of X.

Other stochastic drift models. From an entirely different direction, a number of two-dimensional models of a similar structure to (2.1) and (2.3) have been studied by Lefebvre and collaborators [25–28], motivated by lifetime studies and models of wear, among other applications. Specifically [27, 28] formulate a version of (2.3) for the model (X, Y), where $\mathrm{d}X_t = -c(Y_t/X_t) \, \mathrm{d}t$

and Y is geometric Brownian motion. The focus of these papers is evaluating certain hitting distributions, rather than examining $t \to \infty$ asymptotics.

Long-time behaviour of stochastic interest-rate models. Models for the instantaneous interest rate (also known as the *short rate*) r typically use a mean-reverting diffusion as its stochastic representation of r. Deelstra & Delbaen [13] (see also [12]) extend this framework to a model where the mean-reversion level of the diffusion r is itself stochastic and has arbitrary dependence with the driving Brownian motion: the SDE for the short rate r has constant negative mean-reversion rate $\beta < 0$ and a stochastic mean-reversion level Y,

$$\mathrm{d}r_t = (\beta r_t + Y_t)\,\mathrm{d}t + \sqrt{r_t}\,\mathrm{d}B_t, \qquad \text{with } \frac{1}{t}\int_0^t Y_s\,\mathrm{d}s \xrightarrow{\text{a.s.}} \overline{Y} \text{ as } t \to \infty. \tag{2.4}$$

The main result in [13] shows that r is indeed mean-reverting, $\frac{1}{t}\int_0^t r_s\,\mathrm{d}s \xrightarrow{\text{a.s.}} -\overline{Y}/\beta$ as $t \to \infty$, allowing, as in our Theorem 1.3 above, arbitrary dependence between Y and the Brownian motion B. In contrast to our setting, the process $\sqrt{r}$ in [13] is diffusive. Moreover, as long as Y is ergodic as in (2.4) above, the process r is likely to remain diffusive even if $\beta = 0$, as it resembles the squared-Bessel process of positive dimension.

Time-inhomogeneous one-dimensional diffusions. In the case where the driving function $f(t, y)$ in (1.1) does not depend on y and is polynomial in t, the process X falls into a class of one-dimensional diffusions with space-and-time-dependent drifts, studied in detail by Gradinaru & Offret [18] among a more general class of time-inhomogeneous diffusions on $\mathbb{R}_+$, following earlier discrete-time work of [30]; there are also structural links to the *elephant random walk* [2,3] and to the *noise-reinforced Bessel process* [4].

Specifically, a special case of dynamics (1.1) is an SDE $\mathrm{d}X_t = \rho t^\gamma X_t^{-1}\,\mathrm{d}t + \mathrm{d}B_t$, which is the model of [18] in the case $\alpha = -1$, $\beta = -\gamma$ (their parametrization). Here, for $\gamma > 0$, [18, Thm 4.10(i)] states $\lim_{t\to\infty} t^{-\frac{1+\gamma}{2}} X_t = (2\rho/(1+\gamma))^{1/(1+\gamma)}$, a.s., strengthening the log-scale asymptotic in (1.4) (where our $\alpha = 0$) in this special case.

2.2 Overview and discussion of the proofs

We outline our approach to Theorem 1.3. In what follows, the concrete case of Theorem 1.1 can also be kept in mind, where Y is a square-Bessel process. Recall that the adapted process (S, Y, B), taking values in $\mathbb{R}_+^2 \times \mathbb{R}$ satisfies the dynamics (1.3). In particular, defining

$$M_t := 2\int_0^t \sqrt{S_s}\,\mathrm{d}B_s \quad \text{and} \quad U_t := 2\int_0^t f(s, Y_s)\,\mathrm{d}s, \tag{2.5}$$

we can write, for all $t \in \mathbb{R}_+$,

$$S_t - S_0 = U_t + t + M_t. \tag{2.6}$$

The martingale decomposition (2.6) leads to a natural approach to the proofs, in which the core step for both the weak limit (1.9) and a.s.-asymptotic (1.8) is to show that M is a.s. asymptotically negligible compared to U. Hence U determines the asymptotics of X, both in distribution (via hypothesis A_Y(b)) and in a.s.-asymptotic sense (via hypotheses A_Y(c)–(d)).

The growth rate of martingale M is governed by its quadratic variation $\langle M \rangle = 4\int_0^\cdot S_s\,\mathrm{d}s$: by the Dambis–Dubins–Schwarz theorem, $M = W_{\langle M \rangle}$ for a one-dimensional Brownian motion W time-changed by $\langle M \rangle$. A circular argument can be avoided, since (2.6) allows one to express $\mathbb{E}\langle M \rangle_t$ in terms of $\mathbb{E}\,U_t$, which is controlled by hypotheses A_Y(a) and (d). An argument based on Doob's inequality then yields the desired a.s. upper bound on M (this is the content of Proposition 3.1 below).

While known results about Bessel processes give access to alternative reasoning if one is interested only in Theorem 1.1, the proof in §3 is not long, uses only basic martingale ideas, and is not confined to the Bessel case. We emphasize we do not know any proof that works without some a.s. upper bound, and hence without a hypothesis like A_Y(d). Combined with hypothesis A_Y(c) (which gives an a.s. lower bound on U), this ensures that U is genuinely dominant over M in (2.6). Hypothesis A_Y(b) is required for the distributional limit to exist.

We make some comments on the more apparently technical conditions, which are to some extent necessary but where some variations are possible (say relaxing one at the expense of tightening another). The finite-time integrability hypothesis A_Y(a) is mild (automatically satisfied in the Bessel case) and provides initial control over $\mathbb{E}\langle M\rangle_t$ that is not guaranteed by hypothesis A_Y(d), since the integral in the definition of A_t in (1.7) starts at time 1. In the integral in (1.7), lower limit strictly greater than 0 avoids any possible singularity when $\gamma < 0$; so would using $1+s$ in place of s in the integrand, at the expense of some complications elsewhere (the expression being less amenable to scaling and self-similarity). Finally, we comment on the hypothesis on f given at (1.2). It is important in the present formulation that (1.2) provides an upper bound for $f(t, y)$ in terms of y^α for *all* t (see (3.1) below), since it is used to reduce to the hypothesis A_Y(a) what would otherwise be an hypothesis on $\int_0^1 \mathbb{E}\big[f(t, Y_t)\big]\,\mathrm{d}t$, which becomes a more indirect condition on f. In the Bessel case, Y spends almost all its time on order t, and so it seems likely one could relax the hypothesis on f to demand precise asymptotics only in a smaller region of (t, y)-space, but we do not know any examples that would make the extra work needed worthwhile.

3 Strong law with stochastic drift

In this section we prove Theorem 1.3 under Assumptions (A_f) and (A_Y). We start with a simple observation about f under the hypothesis (A_f). By (1.2), there exists $r_1 > 0$ such that $f(t, y) \le (\rho+1)(1+t)^\gamma(1+y)^\alpha$ for all $(t, y) \in \mathbb{R}_+^2$ with $t + y \ge r_1$. Since f is continuous (and hence bounded on compacts), there exists $C_0 \in (0,\infty)$ such that

$$f(t, y) \le C_0(1 + (1+t)^\gamma y^\alpha) \quad \text{ for all } (t, y) \in \mathbb{R}_+^2. \tag{3.1}$$

By (3.1), there exists $C < \infty$ such that $\sup_{0\le t\le 1} f(t, y) \le C(1+y^\alpha)$. Hence the hypothesis that $\mathbb{E}\int_0^1 Y_t^\alpha\,\mathrm{d}t < \infty$ from (A_Y) implies that $\mathbb{E}\,U_1 < \infty$, where U is as defined in (2.5). We use this fact, as well as the bound (3.1), in several places in the proofs below.

Recall by (A_Y) that for $\alpha \in \mathbb{R}_+$ and $\gamma \in (-\alpha,\infty)$, we have $\beta := \alpha+\gamma > 0$.

Proposition 3.1. *Suppose $\alpha \in \mathbb{R}_+$ and $\beta = \gamma+\alpha > 0$. Assume (A_f) and (A_Y)(d) hold. Then $M = (M_t)_{t\in\mathbb{R}_+}$ in (2.5) is a martingale and there exists a constant $C \in (0,\infty)$ such that $\mathbb{E}(M_t^2) \le Ct^{2+\beta}$ for all $t \in [1,\infty)$. Moreover, for any $\varepsilon > 0$, we have*

$$\lim_{t\to\infty} \sup_{0\le s\le t} |M_s|/t^{1+\beta/2+\varepsilon} = 0, \ \textit{a.s.} \tag{3.2}$$

Proof. For any $N \in \mathbb{R}_+$, set $\tau_N := \inf\{t \in \mathbb{R}_+ \colon S_t \ge N\}$. The quadratic variation of the local martingale $M^{\tau_N} = (M_{t\wedge\tau_N})_{t\in\mathbb{R}_+}$ is bounded: a.s., for all $t \in \mathbb{R}_+$, $\langle M\rangle_{t\wedge\tau_N} \le 4\int_0^t S_{s\wedge\tau_N}\,\mathrm{d}s \le 4tN$. Hence, by [31, Prop. IV.1.23], M^{τ_N} is a uniformly integrable martingale started at zero. Thus $\mathbb{E}\,M_{t\wedge\tau_N} = 0$ for all $t \in \mathbb{R}_+$. Moreover, by definition (1.7), the bound in (3.1) and (A_Y)(d), we have, for all $t \ge 1$, $0 \le \mathbb{E}(U_t - U_1) \le 2C_0(t + \mathbb{E}\,A_t) \le C_0'(1+t)^{1+\beta}$ for a constant $C_0' \in \mathbb{R}_+$. Hence, since U is non-decreasing and $\mathbb{E}\,U_1 < \infty$ (see the comment after (3.1) above), there exists a constant $C_1 \in \mathbb{R}_+$ such that

$$\mathbb{E}\,S_{t\wedge\tau_N} \le \mathbb{E}\,S_0 + \mathbb{E}\,U_1 + \mathbb{E}(U_t - U_1) + t \le C_1(1+t)^{1+\beta}, \text{ for all } t, N \in \mathbb{R}_+.$$

Hence $\mathbb{E}\langle M\rangle_{t\wedge\tau_N} \le 4\int_0^t \mathbb{E}\,S_{s\wedge\tau_N}\,\mathrm{d}s \le 2C_1(1+t)^{2+\beta}$ for $t, N \in \mathbb{R}_+$. Since $\langle M\rangle_{t\wedge\tau_N} \uparrow \langle M\rangle_t$ as $N \to \infty$, monotone convergence implies $\mathbb{E}\langle M\rangle_t \le Ct^{2+\beta}$ for $t \in [1,\infty)$ and $C := 2^{3+\beta}C_1$. By [31, Cor. IV.1.25], M is a martingale and $\mathbb{E}(M_t^2) = \mathbb{E}\langle M\rangle_t \le Ct^{2+\beta}$ for $t \in [1,\infty)$.

The limit in (3.2) will follow by a Borel–Cantelli argument. Doob's maximal inequality [31, Thm II.1.7] and the L^2-bound $\mathbb{E}\, M_t^2 \le Ct^{2+\beta}$ (for all $t \in [1,\infty)$) yield

$$\mathbb{P}\left(\sup_{0\le s\le t} |M_s| > a\right) \le \mathbb{E}\, M_t^2/a^2 \le Ct^{2+\beta}/a^2, \text{ for all } t \ge 1 \text{ and } a > 0. \tag{3.3}$$

Fix arbitrary $\varepsilon > 0$. Set $t_n := 2^n$ and $a_n := t_n^{1+\beta/2+\varepsilon}$ for $n \in \mathbb{Z}_+ := \mathbb{N} \cup \{0\}$. By the inequality in (3.3), applied with $t = t_n$ and $a = a_n$, the probabilities of the events

$$E_n := \left\{\sup_{0\le s\le t_n} |M_s| > a_n\right\} \quad \text{are summable:} \quad \sum_{n=0}^{\infty} \mathbb{P}(E_n) \le C \sum_{n=0}^{\infty} t_n^{-2\varepsilon} < \infty.$$

Hence, by the Borel–Cantelli lemma, there exists a (random) $n_0 \in \mathbb{N}$, a.s., such that E_n occurs for no $n \ge n_0$, i.e., $\sup_{0\le s\le t_n} |M_s| \le a_n$ for all $n \ge n_0$. Let $T := 2^{n_0} < \infty$ a.s. For every $t \in [1,\infty)$, there exists a unique $k \in \mathbb{Z}_+$, such that $t_k = 2^k \le t < 2^{k+1} = t_{k+1}$. Thus, for all $t \ge T$, we have

$$\sup_{0\le s\le t} |M_s| \le \sup_{0\le s\le t_{k+1}} |M_s| \le t_{k+1}^{1+\beta/2+\varepsilon} = 2^{1+\beta/2+\varepsilon}\left(2^k\right)^{1+\beta/2+\varepsilon} \le 2^{1+\beta/2+\varepsilon} t^{1+\beta/2+\varepsilon}.$$

Since $\varepsilon > 0$ was arbitrary, the limit in (3.2) holds almost surely. □

Remark 3.2 (Squared-Bessel process). A solution Y of SDE (1.6) has a decomposition analogous to (2.6), $Y_t = y + \delta t + \overline{M}_t$, with the local martingale $\overline{M}_t := 2\int_0^t \sqrt{Y_s}\,\mathrm{d}W_s$. Since $\delta > 0$ is a constant, a simpler version of the argument in the first paragraph of the previous proof implies $\mathbb{E}\,\overline{M}_t^2 \le Ct^2$ for $t \in [1,\infty)$ and a constant $C > 0$. Doob's maximal inequality [31, Thm II.1.7] yields the tail bound $\mathbb{P}(\sup_{0\le s\le t} |\overline{M}_s| > a) \le Ct^2/a^2$ for all $t, a > 1$. As in the proof of Proposition 3.1, the Borel–Cantelli lemma and a subsequence argument implies that, for any $\varepsilon > 0$, a.s. $\sup_{0\le s\le t} |\overline{M}_s| \le t^{1+(\varepsilon/2)}$ for all sufficiently large $t \in \mathbb{R}_+$, and thus

$$\lim_{t\to\infty} Y_t/t^{1+\varepsilon} = 0, \text{ a.s.} \tag{3.4}$$

We now establish the following key result giving the almost-sure rate of escape for the additive functional U. In particular, by (2.6), the almost-sure behaviour (as $t \to \infty$) of the process S is dominated by U.

Proposition 3.3. *Suppose that $\alpha \ge 0$ and $\beta = \gamma + \alpha > 0$, and assume that* (A_f) *and* (A_Y)*(c),* (A_Y)*(d) hold. Then, for every $\varepsilon > 0$, the process U in* (2.6) *a.s. satisfies*

$$t^{-\varepsilon} < U_t/t^{1+\beta} < t^{\varepsilon} \textit{ for all sufficiently large } t \in \mathbb{R}_+. \tag{3.5}$$

Proof. By (1.2) in (A_f), there exists a large constant $r_\rho > 1$ such that $f(t,y) \ge (\rho/2)t^\gamma y^\alpha$ for all $(t,y) \in [r_\rho,\infty) \times \mathbb{R}_+$. By the definition of U_t in (2.6), for all $t > r_\rho$, we have

$$\rho(A_t - A_{r_\rho}) = \rho \int_{r_\rho}^t s^\gamma Y_s^\alpha \,\mathrm{d}s \le 2\int_0^t f(s,Y_s)\,\mathrm{d}s = U_t,$$

where A_t is defined in (1.7). Hypothesis (A_Y)(c) on A_t implies $\lim_{t\to\infty} t^{\varepsilon-(1+\beta)} A_t = \infty$ a.s. for every $\varepsilon > 0$, yielding the lower bound in (3.5).

To prove the upper bound in (3.5), note that the definition of U_t in (2.6) and the upper bound on f in (3.1) imply $U_t - U_1 \le 2C_0(t + A_t)$ a.s. Hence, by (A_Y)(d), for a constant $C_2 < \infty$,

$$\mathbb{E}\, U_t = \mathbb{E}[U_t - U_1] + \mathbb{E}\, U_1 \le 2C_0(t + \mathbb{E}\, A_t) + \mathbb{E}\, U_1 \le C_2 t^{1+\beta} \quad \text{for all } t \in [1,\infty).$$

Pick any $\varepsilon > 0$. The bound in the previous display and Markov's inequality yield

$$\mathbb{P}(U_t \geq t^{1+\beta+\varepsilon}) \leq \mathbb{E}\, U_t/t^{1+\beta+\varepsilon} \leq C_2/t^\varepsilon.$$

Since $\sum_{n\in\mathbb{Z}_+} 1/t_n^\varepsilon < \infty$, where $t_n := 2^n$, the Borel–Cantelli lemma implies $U_{t_n} \leq t_n^{1+\beta+\varepsilon}$ for all but finitely many $n \in \mathbb{Z}_+$, a.s. As in the proof of Proposition 3.1, for every $t \in [1,\infty)$, there exists a unique $k \in \mathbb{Z}_+$, such that $t_k = 2^k \leq t < 2^{k+1} = t_{k+1}$. As U is increasing (since $f \geq 0$)

$$U_t \leq U_{t_{k+1}} \leq (2^{k+1})^{1+\beta+\varepsilon} = 2^{1+\beta+\varepsilon} \cdot t_k^{1+\beta+\varepsilon} \leq 2^{1+\beta+\varepsilon} t^{1+\beta+\varepsilon}.$$

Since $\varepsilon > 0$ is arbitrary, the upper bound in (3.5) holds for all large $t \in \mathbb{R}_+$. $\square$

The next result establishes weak convergence of U_t, defined in (2.6) and suitably scaled, to a random variable proportional to the limit $\widetilde{A}$ in hypothesis (A_Y)(b).

Lemma 3.4. *Let $\alpha \geq 0$ and $\beta = \gamma + \alpha > 0$. Assume (A_f) and (A_Y)(b) hold. Then,*

$$U_t/t^{1+\beta} \xrightarrow{\mathrm{d}} 2\rho\widetilde{A}, \text{ as } t \to \infty.$$

Proof. Recall the definitions of A_t in (1.7) and of U_t in (2.6). Define also

$$D_t := U_t/2 - \rho A_t = U_1/2 + \int_1^t f(s, Y_s)\,\mathrm{d}s - \rho A_t \quad \text{for any } t \in [1,\infty). \tag{3.6}$$

Since $U_t/t^{1+\beta} = 2(D_t + \rho A_t)/t^{1+\beta}$, by Slutsky's lemma (see e.g. [14, p. 105]) and the weak limit in (A_Y)(b), it is sufficient to prove that $D_t/t^{1+\beta}$ converges to zero in probability. In particular, it suffices to show that for every $\varepsilon', \varepsilon'' > 0$ we have

$$\limsup_{t\to\infty} \mathbb{P}\big(|D_t|/t^{1+\beta} > 2\varepsilon'\big) \leq \varepsilon''. \tag{3.7}$$

For every $\varepsilon > 0$, by (1.2), there exists $r_\varepsilon \in (1,\infty)$ such that, for every $s \in [r_\varepsilon, \infty)$ and $y \in \mathbb{R}_+$, we have $|f(s,y) - \rho s^\gamma y^\alpha| \leq \varepsilon s^\gamma y^\alpha$. Define $Z_\varepsilon := U_1/2 + \int_1^{r_\varepsilon} f(s, Y_s)\,\mathrm{d}s + \rho A_{r_\varepsilon}$ and note that, by (3.6), $|D_{r_\varepsilon}| \leq Z_\varepsilon$, a.s. Hence, for any $t > r_\varepsilon$, by (1.7) and (3.6) again, we have

$$\begin{aligned} |D_t| &\leq |D_{r_\varepsilon}| + |D_t - D_{r_\varepsilon}| \\ &\leq Z_\varepsilon + \int_{r_\varepsilon}^t |f(s,Y_s) - \rho s^\gamma Y_s^\alpha|\,\mathrm{d}s \leq Z_\varepsilon + \varepsilon(A_t - A_{r_\varepsilon}) \leq Z_\varepsilon + \varepsilon A_t. \end{aligned} \tag{3.8}$$

Pick any $\varepsilon', \varepsilon'' > 0$. Fix small $\varepsilon > 0$, such that $\lim_{t\to\infty} \mathbb{P}(\varepsilon A_t/t^{1+\beta} > \varepsilon') = \mathbb{P}(\widetilde{A} > \varepsilon'/\varepsilon)$ (by (A_Y)(b), limit holds for all but countably many ε) and $\mathbb{P}(\widetilde{A} > \varepsilon'/\varepsilon) < \varepsilon''/2$. Thus, by (3.8),

$$\mathbb{P}\big(|D_t|/t^{1+\beta} > 2\varepsilon'\big) \leq \mathbb{P}\left(Z_\varepsilon/t^{1+\beta} > \varepsilon'\right) + \mathbb{P}(A_t/t^{1+\beta} > \varepsilon'/\varepsilon) \leq \varepsilon''/2 + \varepsilon''/2$$

for all large $t \in \mathbb{R}_+$. Since $\varepsilon', \varepsilon'' > 0$ were arbitrary, (3.7) holds and the lemma follows. $\square$

We can now complete the proof of Theorem 1.3.

Proof of Theorem 1.3. Let $\beta := \gamma+\alpha$. Since $S_t = S_0+U_t+t+M_t$ and $(S_0+t+M_t)/t^{1+\beta} \xrightarrow{\text{a.s.}} 0$ as $t \to \infty$ by (3.2) of Prop. 3.1, the weak limit in Lemma 3.4 and Slutsky's lemma (see [14, p. 105]) yield $S_t/t^{1+\beta} \xrightarrow{\mathrm{d}} 2\rho\widetilde{A}$ as $t \to \infty$. Then (1.9) follows by continuous mapping. By (3.5) in Prop. 3.3, for any $\varepsilon > 0$ a.s. $\log S_t/\log t = 1 + \beta + \log\big((S_0 + t + M_t)/t^{1+\beta} + U_t/t^{1+\beta}\big)/\log t \leq 1 + \beta + \varepsilon$ for all large t. Similarly, for any $\varepsilon \in (0, \beta/2)$, (3.2) of Proposition 3.1 and the lower bound in (3.5) imply $1 + \beta - \varepsilon \leq \log S_t/\log t$ and limit (1.8) follows. $\square$

4 Asymptotics of certain squared-Bessel functionals

The aim of this section is to verify Assumption (A_Y) for the squared-Bessel process Y (with law $\mathrm{BESQ}^\delta(y)$) of positive "dimension" $\delta > 0$. This will, by Theorem 1.3, imply our main result Theorem 1.1. The focus in this section is the additive functional A_t, defined in (1.7), when Y follows SDE (1.6).

Theorem 4.1. *Suppose that Y has law $\mathrm{BESQ}^\delta(y)$ with parameter $\delta > 0$, started at arbitrary $y \in \mathbb{R}_+$. Let $\alpha \in \mathbb{R}_+$ and $\gamma \in (-\alpha, \infty)$. Recall $A_t = \int_1^t s^\gamma Y_s^\alpha \,\mathrm{d}s$, $t \in [1, \infty)$, defined in (1.7).*

(a) The following limit holds,

$$\lim_{t\to\infty} \frac{\log A_t}{\log t} = 1 + \gamma + \alpha, \ \ a.s. \tag{4.1}$$

(b) Let $\widetilde{Y}$ follow $\mathrm{BESQ}^\delta(0)$. Then, as $t \to \infty$,

$$A_t/t^{1+\gamma+\alpha} \xrightarrow{\mathrm{d}} \int_0^1 s^\gamma \widetilde{Y}_s^\alpha \,\mathrm{d}s. \tag{4.2}$$

(c) The growth of the expectation $\mathbb{E}\, A_t$ is at most polynomial, $\sup_{t\geq 1} \mathbb{E}\, A_t/t^{1+\gamma+\alpha} < \infty$.

Note that, since $\alpha \geq 0$, for all $t > 1$ and $Y_0 = y \in \mathbb{R}_+$ we have $0 < A_t < \infty$. The inequality $\limsup_{t\to\infty} \log A_t/\log t \leq 1+\gamma+\alpha$ in (4.1) of Theorem 4.1(a) follows directly from the a.s upper bound (3.4) above for the squared-Bessel process Y. Assumption (A_Y)(c) requires only $\liminf_{t\to\infty} \log A_t/\log t \geq 1 + \gamma + \alpha$ in (4.1), the most involved part of Theorem 4.1. Proposition 4.2 below states this as a stand-alone result, followed by an elementary, essentially self-contained proof. From the perspective of our main result (Theorem 1.1), the case $\delta \in (0, 2]$ is also the most interesting as it corresponds to the non-ergodic recurrent stochastic drift in (1.1), which nevertheless induces anomalous diffusive behaviour of $X = \sqrt{S}$. In the Brownian case ($\delta = 1$), the log-limit statement (4.1) can also be deduced, with a little work, from results of [23].

Parts (b)&(c) of Theorem 4.1 follow directly from self-similarity of the Bessel processes.

Proposition 4.2. *Suppose that $\delta > 0$, $\alpha \in \mathbb{R}_+$, and $\gamma > -\alpha$. Then, for every $\varepsilon > 0$, a.s.*

$$A_t > t^{1+\gamma+\alpha-\varepsilon} \qquad \text{for all large } t \in \mathbb{R}_+. \tag{4.3}$$

Remarks 4.3. (i) The intuition behind (4.3) is that, typically, $Y_t \approx t$, by Bessel scaling. However, the proof below reveals that the δ-independent exponent in (4.3) emerges from a rather delicate balance between the frequency and spatial extent of excursions, which both depend critically on δ; in [9] bounds of a similar type yield lower bounds on the rates of convergence of ergodic reflected Brownian motion. A related approach in discrete time can be found in [19].

(ii) In the case $\gamma = 0$, the quantity A_{τ_ℓ} is studied in detail by Çetin in [10], where Y follows $\mathrm{BESQ}^\delta(y)$ and $\delta > 0$ (the hitting time τ_ℓ is defined below). The approach in [10] does not address directly the a.s. asymptotics for A_t.

Throughout this section, we write $\mathbb{P}_y$ for the probability measure inducing the law $\mathrm{BESQ}^\delta(y)$, and $\mathbb{E}_y$ for the corresponding expectation. Let $\tau_x := \inf\{t \in \mathbb{R}_+ : Y_t = x\}$ (with convention $\inf \emptyset = \infty$) for $x \in \mathbb{R}_+$. The next lemma says the squared-Bessel process started at y will typically spend time at least of order y at levels of order y.

Lemma 4.4. *Suppose that $\delta > 0$. Then there exists $\varepsilon > 0$ such that,*

$$\mathbb{P}_y(\tau_{y/2} \geq \varepsilon y) \geq 1/2, \ \ \textit{for all } y \in \mathbb{R}_+.$$

Proof. It suffices to suppose that $Y_0 = y \in (0,\infty)$. From (1.6), we obtain, for all $t \in \mathbb{R}_+$,

$$Y_{t\wedge\tau_{2y}} = y + \delta(t\wedge\tau_{2y}) + M_{t\wedge\tau_{2y}}, \text{ where } M_t := 2\int_0^t |Y_s|^{1/2}\,\mathrm{d}W_s. \tag{4.4}$$

The process $(M_{t\wedge\tau_{2y}})_{t\in\mathbb{R}_+}$ is a martingale since its quadratic variation is a.s. bounded at any finite time t. Hence, by (4.4), we get $\mathbb{E}_y(Y_{t\wedge\tau_{2y}}) \le y + \delta t$. By [31, Prop. IV.1.23],

$$\mathbb{E}_y\, M^2_{t\wedge\tau_{2y}} = \mathbb{E}_y \langle M\rangle_{t\wedge\tau_{2y}} = 4\int_0^t \mathbb{E}_y\, Y_{s\wedge\tau_{2y}}\,\mathrm{d}s \le 4yt + 2\delta t^2, \text{ for all } t\in\mathbb{R}_+.$$

Consider the event $E_{y,t} := \{\sup_{0\le s\le t}|M_{s\wedge\tau_{2y}}| \ge y/3\}$. Then, by Doob's maximal inequality [31, Thm II.1.7] applied to the non-negative submartingale $(M^2_{t\wedge\tau_{2y}})_{t\in\mathbb{R}_+}$,

$$\mathbb{P}_y(E_{y,t}) \le (9/y^2)\,\mathbb{E}_y\, M^2_{t\wedge\tau_{2y}} \le 36(t/y) + 18\delta(t/y)^2. \tag{4.5}$$

If $t < y/(2\delta)$, we have $E^{\mathrm{c}}_{y,t} \subset \{\tau_{2y} > t\}$ $\mathbb{P}_y$-a.s., implying $E^{\mathrm{c}}_{y,t} = \{\sup_{0\le s\le t}|M_s| < y/3\}$. By (4.4) we thus get $E^{\mathrm{c}}_{y,t} \subset \{\inf_{0\le s\le t} Y_s > y/2\} = \{\tau_{y/2} > t\}$ $\mathbb{P}_y$-a.s., yielding (with (4.5))

$$\mathbb{P}_y\left(\tau_{y/2} > t\right) \ge \mathbb{P}_y(E^{\mathrm{c}}_{y,t}) = 1 - \mathbb{P}_y(E_{y,t}) \ge 1 - 36(t/y) - 18\delta(t/y)^2. \tag{4.6}$$

For any $\varepsilon \in (0,(2\delta)^{-1})$, define $t := \varepsilon y < y/(2\delta)$, and $\mathbb{P}_y(\tau_{y/2} \ge \varepsilon y) \ge 1 - 36\varepsilon - 18\delta\varepsilon^2$ by (4.6). Choose $\varepsilon \in (0,(2\delta)^{-1})$, such that $1 - 36\varepsilon - 18\delta\varepsilon^2 \ge 1/2$, to finish the proof. $\square$

We now establish a lower bound on the tail of a functional of an excursion of Y.

Lemma 4.5. *Let $\delta \in (0,2)$, $\alpha \in \mathbb{R}_+$. Then there exists a constant $c_{\alpha,\delta} \in (0,\infty)$ satisfying*

$$\mathbb{P}_1\left(\int_0^{\tau_0} Y_s^\alpha\,\mathrm{d}s \ge z\right) \ge c_{\alpha,\delta} z^{-\frac{2-\delta}{2+2\alpha}}, \qquad \text{for all } z \in (1,\infty).$$

Proof. First observe that, since $Y_s \in \mathbb{R}_+$ for all $s \in \mathbb{R}_+$, for every $y \in (0,\infty)$ and $z \in \mathbb{R}_+$,

$$\begin{aligned}\mathbb{P}_1\left(\int_0^{\tau_0} Y_s^\alpha\,\mathrm{d}s \ge z\right) &\ge \mathbb{E}_1\left[\mathbb{1}_{\{\tau_y<\tau_0\}}\mathbb{P}_1\left(\int_{\tau_y}^{\tau_0} Y_s^\alpha\,\mathrm{d}s \ge z \,\middle|\, \mathcal{F}_{\tau_y}\right)\right]\\ &= \mathbb{P}_1(\tau_y<\tau_0)\,\mathbb{P}_y\left(\int_0^{\tau_0} Y_s^\alpha\,\mathrm{d}s \ge z\right),\end{aligned} \tag{4.7}$$

where $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is the filtration generated by the Brownian motion driving SDE (1.6). The equality in (4.7) follows from the strong Markov property at τ_y, and the facts: $\{\tau_y < \tau_0\} \in \mathcal{F}_{\tau_y}$ and $Y_{\tau_y} = y$ $\mathbb{P}_1$-a.s. By Lemma 4.4 there exists $\varepsilon > 0$, such that $\mathbb{P}_y(\tau_{y/2} \ge \varepsilon y) \ge 1/2$. Moreover, on the event $\{\tau_{y/2} \ge \varepsilon y\}$, the inequality

$$\int_0^{\tau_0} Y_s^\alpha\,\mathrm{d}s \ge \int_0^{\varepsilon y} (y/2)^\alpha\,\mathrm{d}s = \varepsilon y^{\alpha+1}/2^\alpha \quad \text{holds } \mathbb{P}_y\text{-a.s.},$$

implying $\mathbb{P}_y\left(\int_0^{\tau_0} Y_s^\alpha\,\mathrm{d}s \ge \varepsilon y^{\alpha+1}/2^\alpha\right) \ge \mathbb{P}_y(\tau_{y/2} \ge \varepsilon y) \ge 1/2$ for all $y \in (0,\infty)$. By (4.7) with $z = \varepsilon y^{\alpha+1}/2^\alpha$, we obtain

$$\mathbb{P}_1\left(\int_0^{\tau_0} Y_s^\alpha\,\mathrm{d}s \ge \varepsilon y^{\alpha+1}/2^\alpha\right) \ge \frac{1}{2}\,\mathbb{P}_1(\tau_y<\tau_0) = \frac{1}{2y^{(2-\delta)/2}} \text{ for all } y \ge 1. \tag{4.8}$$

The last equality follows from the optional sampling theorem since the process $(Y^{(2-\delta)/2}_{t\wedge\tau_0\wedge\tau_y})_{t\in\mathbb{R}_+}$ is (by Itô's formula and SDE (1.6)) a bounded martingale. The following change of variable $y = (2^\alpha z/\varepsilon)^{1/(1+\alpha)}$ in (4.8) yields the stated bound. $\square$

Proof of Proposition 4.2. Suppose that $\delta \in (0,2)$ and $\alpha \in \mathbb{R}_+$. We will first establish the case $\gamma = 0$ of (4.3), then $\delta \geq 2$ and subsequently deduce the general case $\gamma \in (-\alpha, \infty)$.

Define stopping times $\vartheta_0 := 0$ and, for $n \in \mathbb{N}$,

$$\varphi_n := \inf\{t \geq \vartheta_{n-1} : Y_t = 1\}, \qquad \vartheta_n := \inf\{t \geq \varphi_n : Y_t = 0\}. \tag{4.9}$$

Since the squared-Bessel process is point-recurrent for $\delta \in (0,2)$, we have $0 < \vartheta_n < \varphi_{n+1} < \vartheta_{n+1} < \infty$, a.s., for all $n \in \mathbb{N}$ (and $0 = \vartheta_0 < \varphi_1$ as well, unless $Y_0 = 1$). Define

$$N_t := \sup\{n \in \mathbb{Z}_+ : \vartheta_n \leq t\},\ t \in \mathbb{R}_+, \qquad \text{and} \qquad I_{\alpha,n} := \int_{\varphi_n}^{\vartheta_n} Y_s^\alpha \,\mathrm{d}s,\ n \in \mathbb{N}. \tag{4.10}$$

Since $Y_s \geq 0$ for all $s \in \mathbb{R}_+$, and $0 \leq \vartheta_{N_t} \leq t$, we can write

$$\int_0^t Y_s^\alpha \,\mathrm{d}s \geq \int_0^{\vartheta_{N_t}} Y_s^\alpha \,\mathrm{d}s \geq \sum_{n=1}^{N_t} I_{\alpha,n}, \text{ for all } t \in \mathbb{R}_+. \tag{4.11}$$

The strong Markov property and the fact that $\mathbb{P}(Y_{\varphi_n} = 1) = 1$ imply the random variables $(I_{\alpha,n})_{n\in\mathbb{N}}$ are i.i.d. By Lemma 4.5 we have

$$\mathbb{P}(I_{\alpha,1} \geq z) = \mathbb{P}_1\left(\int_0^{\tau_0} Y_s^\alpha \,\mathrm{d}s \geq z\right) \geq c_{\alpha,\delta} z^{-\frac{1-\delta/2}{1+\alpha}}, \quad \text{for all } z \in (1,\infty). \tag{4.12}$$

For any $\varepsilon > 0$, this bound and the Borel–Cantelli lemma yield: a.s.

$$\sum_{n=1}^{N} I_{\alpha,n} \geq N^{\frac{1+\alpha}{1-\delta/2}-\varepsilon} \quad \text{for all but finitely many } N \in \mathbb{N}. \tag{4.13}$$

Indeed, denoting $\theta := \frac{1-\delta/2}{1+\alpha} \in (0,1)$ (recall $0 < \delta < 2$ and $\alpha \in \mathbb{R}_+$) and picking $r > 0$ with $c_{\alpha,\delta} r > 1$, by (4.12) for all $N \geq 2$ we have $\mathbb{P}(I_{\alpha,1} \geq (N/(r\log N))^{1/\theta}) \geq c_{\alpha,\delta} r(\log N)/N$. Since the variables $(I_{\alpha,n})_{n\in\mathbb{N}}$ are i.i.d., the events $E_N := \{\max_{1\leq n\leq N} I_{\alpha,n} < (N/(r\log N))^{1/\theta}\}$ satisfy

$$\mathbb{P}(E_N) = (1 - \mathbb{P}(I_{\alpha,1} \geq N^{1/\theta}/(r\log N)^{1/\theta}))^N \leq (1 - c_{\alpha,\delta} r(\log N)/N)^N \leq N^{-\bar{c}},$$

for all $N \in \mathbb{N}$ and $\bar{c} := c_{\alpha,\delta} r > 1$ (the last inequality follows by taking logarithms on both sides and applying $\log(1-x) \leq -x$ for $x < 1$). Since $\bar{c} > 1$, the Borel–Cantelli lemma implies a.s. $\max_{1\leq n\leq N} I_{\alpha,n} \geq (N/(r\log N))^{1/\theta}$ for all sufficiently large N. As the sum of positive terms dominates the maximum and $(N/(r\log N))^{1/\theta} \geq N^{\frac{1+\alpha}{1-\delta/2}-\varepsilon}$ for every $\varepsilon > 0$ and all $N \in \mathbb{N}$ sufficiently large, (4.13) follows.

Let $\nu_n := \vartheta_n - \varphi_n$, for $n \in \mathbb{N}$, be the duration of the excursion (at the epoch φ_n) of Y from level 1 to level 0. We now prove (4.14) below that controls the tail of the duration ν_n. By the strong Markov property at φ_n we have $\mathbb{P}(\nu_n > t) = \mathbb{P}_1(\tau_0 > t)$ for all $t \in \mathbb{R}_+$ and $n \in \mathbb{N}$. For any $Y_0 = y \in \mathbb{R}_+$ and $t > 0$, by the scaling property, the process $(Y_{st}/t)_{s\in\mathbb{R}_+}$ follows $\mathrm{BESQ}^\delta(y/t)$. Hence $\mathbb{P}_y(\tau_0 > t) = \mathbb{P}_y(\inf_{s\in[0,t]} Y_s > 0) = \mathbb{P}_{y/t}(\inf_{s\in[0,1]} Y_s > 0) = G(y/t)$, for a measurable function $G : \mathbb{R}_+ \to [0,1]$ satisfying $G(0) = 0$ and $G(y) \to 1$ as $y \to \infty$.

The process $(Z_s)_{s\in[0,t]}$, $Z_s := \mathbb{P}(\tau_0 > t \mid \mathcal{F}_s) = \mathbb{1}_{\{\tau_0>s\}} G(Y_s/(t-s))$, is a martingale with respect to the filtration $(\mathcal{F}_s)_{s\in[0,t]}$ generated by the Brownian motion W in SDE (1.6). Assuming G is twice differentiable on $(0,\infty)$, Itô's formula and the infinitesimal drift of $(Z_s)_{s\in[0,t]}$ being equal to zero imply that G satisfies an ordinary differential equation (ODE) $2G''(x) + (1+\delta/x)G'(x) = 0$ for $x \in (0,\infty)$ with boundary conditions above. The solution $\bar{G}(x) := \int_0^{x/2} u^{-\delta/2}\mathrm{e}^{-u}\,\mathrm{d}u/\Gamma(1+\delta/2)$ of the ODE, where Γ denotes the gamma function, yields a martingale $\bar{Z}_s := \mathbb{1}_{\{\tau_0>s\}}\bar{G}(Y_s/(t-s))$, $s \in [0,t]$, satisfying $Z_t = \mathbb{1}_{\{\tau_0>t\}} = \bar{Z}_t$. Thus $Z_s = \bar{Z}_s$ for all $s \in [0,t]$ a.s. Since the support of Y_s (for $s > 0$) contains $(0,\infty)$, we have $\bar{G}(y/t) =$

$\bar{G}(y/t) = \mathbb{P}_y(\tau_0 > t)$ for all $y \in \mathbb{R}_+$ and $t > 0$. Using $\mathrm{e}^{-u} \le 1$ for $u \in \mathbb{R}_+$, we obtain the bound $\bar{G}(1/t) \le c_\delta t^{-1+\delta/2}$ for all $t > 0$, where $c_\delta := 2^{-1+\delta/2}/((1-\delta/2)\Gamma(1+\delta/2))$. Thus, since τ_0 has a density, we get the upper tail bound[1]

$$\mathbb{P}(\nu_n \ge t) = \mathbb{P}_1(\tau_0 > t) \le c_\delta t^{-1+\delta/2} \quad \text{for all } t \in (0,\infty) \text{ and any } n \in \mathbb{N}. \tag{4.14}$$

By (4.14), the i.i.d. sequence $(\nu_n)_{n\in\mathbb{N}}$ satisfies $\mathbb{E}\,\nu_n^{(1-\delta/2)/(1+\varepsilon)} < \infty$ for any $\varepsilon > 0$. Since $0 < 1-\delta/2 < 1$, the Marcinkiewicz–Zygmund strong law [21, Thm 4.23] yields: a.s.,

$$\sum_{n=1}^{N} \nu_n \le N^{\frac{1}{1-\delta/2}+\varepsilon} \quad \text{for all but finitely many } N \in \mathbb{N}. \tag{4.15}$$

Define $\upsilon_n := \varphi_n - \vartheta_{n-1}$, for $n \in \mathbb{N}$, to be the time taken to exit the interval $[0,1]$ started from 0, after time ϑ_{n-1}. Since $y \mapsto \mathbb{P}_y(Y_1 > 1)$ is positive and continuous on $[0,1]$, there exists a constant $c_0 > 0$ such that $\mathbb{P}_y(Y_1 > 1) > c_0$ for all $y \in [0,1]$. Thus, $\mathbb{P}(\upsilon_n \ge t) \le \mathbb{P}_0(\cap_{i=1}^{\lfloor t\rfloor}\{Y_i \le 1\}) \le (1-c_0)^{\lfloor t\rfloor}$ for $t \ge 1$ (here $\lfloor t\rfloor := \max\{i \in \mathbb{Z}_+ : i \le t\}$), implying $\mathbb{E}\,\upsilon_n < \infty$ for all $n \in \mathbb{N}$. Since the sequence $(\upsilon_n)_{n\in\mathbb{N}}$ is i.i.d., the strong law of large numbers says $N^{-1}\sum_{n=1}^N \upsilon_n \to \mathbb{E}\,\nu_1$ as $N \to \infty$ a.s. Recalling $\nu_n = \vartheta_n - \varphi_n$, by (4.15) and since $1 < 1/(1-\delta/2)$, for any $\varepsilon > 0$ we obtain a.s.

$$\vartheta_N = \sum_{n=1}^{N}(\vartheta_n - \vartheta_{n-1}) = \sum_{n=1}^{N}\nu_n + \sum_{n=1}^{N}\upsilon_n \le N^{\frac{1}{1-\delta/2}+\varepsilon} \quad \text{for all large } N \in \mathbb{N}. \tag{4.16}$$

By (4.10), for every $t \in \mathbb{R}_+$ we have $N_t < \infty$, $t < \vartheta_{N_t+1}$, a.s., and $N_t \xrightarrow{\text{a.s.}} \infty$ as $t \to \infty$. Thus, by (4.16), for any $\varepsilon \in (0,1)$, a.s. we have $N_t \ge t^{(1-\delta/2)/(1+\varepsilon)} \ge t^{(1-\delta/2)(1-\varepsilon)}$ for all $t \in \mathbb{R}_+$ sufficiently large. This inequality and the bound in (4.13), combined with (4.11), imply (4.3) when $\gamma = 0$: for any sufficiently small $\varepsilon > 0$, a.s., $\int_0^t Y_s^\alpha\,\mathrm{d}s \ge \sum_{n=1}^{N_t} I_{\alpha,n} \ge N_t^{\frac{1+\alpha}{1-\delta/2}-\varepsilon}$ and hence, for $t \in \mathbb{R}_+$ sufficiently large, $\int_0^t Y_s^\alpha\,\mathrm{d}s \ge t^{(1+\alpha)(1-2\varepsilon)}$. Thus, for any (possibly random) $T > 0$,

$$\text{a.s.} \quad \int_T^t Y_s^\alpha\,\mathrm{d}s \ge t^{(1+\alpha)(1-2\varepsilon)} \quad \text{for all sufficiently large } t \in \mathbb{R}_+. \tag{4.17}$$

If $\delta \ge 2$, since Y with law $\mathrm{BESQ}^\delta(y)$ is a continuous-state branching process with immigration, the process $\bar{Y}+Y'$, where $\bar{Y}$ and Y' are independent with laws $\mathrm{BESQ}^1(0)$ and $\mathrm{BESQ}^{\delta-1}(y)$, respectively, has the same law as Y [31, Thm XI.1.2]. Since $Y'_s \ge 0$ for $s \in \mathbb{R}_+$, by (4.17) a.s. $\int_0^t(\bar{Y}_s + Y'_s)^\alpha\,\mathrm{d}s \ge \int_0^t \bar{Y}_s^\alpha\,\mathrm{d}s \ge t^{(1+\alpha)(1-2\varepsilon)}$ for large t, implying (4.17) for all $\delta > 0$.

Suppose $\gamma > 0$. By (3.4), for every $\varepsilon \in (0,\gamma\wedge 1)$ we have $\lim_{t\to\infty} Y_t^{1-\varepsilon}/t = 0$. Hence, there exists (a random) $T \in \mathbb{R}_+$, such that $t^\gamma \ge Y_t^{\gamma-\varepsilon}$ and $\int_0^t s^\gamma Y_s^\alpha\,\mathrm{d}s \ge \int_T^t Y_s^{\gamma+\alpha-\varepsilon}\,\mathrm{d}s$ for $t \ge T$. Thus, by (4.17) (with $\alpha+\gamma-\varepsilon$ in place of α), (4.3) follows. If $\gamma \in (-\alpha, 0)$, (4.3) remains valid by (4.17) (with $T = 1$) and $\int_1^t s^\gamma Y_s^\alpha\,\mathrm{d}s \ge t^\gamma \int_1^t Y_s^\alpha\,\mathrm{d}s$, concluding the proof. □

Proof of Theorem 4.1. Part (a). By (3.4), for every $\varepsilon > 0$, $\lim_{t\to\infty} Y_t/t^{1+\varepsilon} = 0$ a.s. Hence, a.s., $Y_t \le t^{1+\varepsilon}$ for all sufficiently large $t \in \mathbb{R}_+$. Thus, by (1.7), $A_t \le t^{1+\gamma+\alpha+\varepsilon}$ for large t, implying

$$\limsup_{t\to\infty}\frac{\log A_t}{\log t} \le 1+\gamma+\alpha, \text{ a.s.} \tag{4.18}$$

The lower bound in (4.1) follows from (4.3) in Proposition 4.2, since, for any $\varepsilon > 0$,

$$\liminf_{t\to\infty}\frac{\log A_t}{\log t} \ge 1+\alpha+\gamma-\varepsilon + \liminf_{t\to\infty}\frac{\log(t^\varepsilon A_t/t^{1+\alpha+\gamma})}{\log t} \ge 1+\gamma+\alpha-\varepsilon \text{ a.s.}$$

[1]This upper bound is of the same order as the lower bound in the special case $\alpha = 0$ of the Lemma 4.5.

Part (b). By self-similarity of the Bessel process, it follows that, for every $t \in (0,\infty)$ and starting point $y \in \mathbb{R}_+$, the process $(t^{-1}Y_{ts})_{s\in\mathbb{R}_+}$ follows $\mathrm{BESQ}^\delta(y/t)$, where Y has the law $\mathrm{BESQ}^\delta(y)$. Moreover, for $\varepsilon_t := y/t \ge 0$ and $\widetilde{Y}$ with law $\mathrm{BESQ}^\delta(0)$, the process $(\widetilde{Y}_{s+\tilde{\tau}_{\varepsilon_t}})_{s\in\mathbb{R}_+}$ also follows $\mathrm{BESQ}^\delta(y/t)$ by the strong Markov property, where $\tilde{\tau}_{\varepsilon_t} = \inf\{s \in \mathbb{R}_+ : \widetilde{Y}_s = \varepsilon_t\}$. By substituting $s = tv$ in definition (1.7) of A_t, we obtain (4.2):

$$A_t/t^{1+\gamma+\alpha} = \int_{1/t}^1 v^\gamma (t^{-1}Y_{tv})^\alpha \,\mathrm{d}v \overset{\mathrm{d}}{=} \int_{1/t}^1 v^\gamma \widetilde{Y}^\alpha_{v+\tilde{\tau}_{\varepsilon_t}} \,\mathrm{d}v \xrightarrow{\text{a.s.}} \int_0^1 v^\gamma \widetilde{Y}^\alpha_v \,\mathrm{d}v \quad \text{as } t\to\infty. \tag{4.19}$$

If $\gamma \ge 0$, the a.s. convergence in (4.19) clearly holds. If $\gamma \in (-\alpha, 0)$, and hence $\alpha > 0$, time-reversal $(\widetilde{Y}_t)_{t>0} \overset{\mathrm{d}}{=} (t^2\widetilde{Y}_{1/t})_{t>0}$ and bound (3.4) imply a.s. $\widetilde{Y}_v \le v^{1-\varepsilon}$ for small $v > 0$ and any $\varepsilon \in (0, 1\wedge(1+\gamma/\alpha))$. Thus $\int_0^1 v^\gamma \widetilde{Y}^\alpha_v \,\mathrm{d}v < \infty$ and the a.s. limit holds.

Part (c). The scaling property implies $Y_t \overset{\mathrm{d}}{=} t\overline{Y}_1$, where $\overline{Y}$ follows $\mathrm{BESQ}^\delta(y/t)$. Thus, for $\alpha \ge 0$, there is a constant $C_{\delta,\alpha,y} < \infty$, such that $\mathbb{E}\, Y_t^\alpha = t^\alpha\, \mathbb{E}\,\overline{Y}_1^\alpha \le C_{\delta,\alpha,y} t^\alpha$ for all $t \in [1,\infty)$. Since $1+\gamma+\alpha > 0$, definition (1.7) implies $\mathbb{E}\, A_t = O(t^{1+\gamma+\alpha})$ as $t\to\infty$. □

Proof of Theorem 1.1. Suppose that Y follows $\mathrm{BESQ}^\delta(y)$, $\delta > 0$, for a fixed $y \in \mathbb{R}_+$. We check the hypotheses of Theorem 1.3. The fact that $\int_0^1 \mathbb{E}\big[Y_t^\alpha\big]\,\mathrm{d}t < \infty$ is immediate, since moments of all orders for Y_t are bounded over compact time intervals: this follows from continuity and the Gaussian-like tails of the explicit Bessel transition density (see e.g. [5, §IV.6]). Theorem 4.1 establishes that hypothesis (A_Y) is satisfied by Y following $\mathrm{BESQ}^\delta(y)$, provided that $\delta > 0$, $\alpha \in \mathbb{R}_+$, and $\gamma + \alpha > 0$. Indeed, Theorem 4.1(a) implies hypothesis (A_Y)(c), Theorem 4.1(b) implies (A_Y)(b), with limit $\widetilde{A} = \int_0^1 s^\gamma \widetilde{Y}^\alpha_s \,\mathrm{d}s$, and Theorem 4.1(c) implies (A_Y)(d). Hence Theorem 1.1 is a consequence of Theorem 1.3. □

A Existence, uniqueness and positivity for SDE (1.3)

Let (Y, B) be an adapted process, on a filtered probability space, where B is a scalar Brownian motion with respect to the given filtration and Y is a continuous adapted process in $\mathbb{R}_+$.

Existence and pathwise uniqueness. As noted in [12], the SDE in (1.3) for S is a special case of the Doléans-Dade and Protter equation, driven by the semimartingale B and an adapted continuous process $K = (K_t)_{t\in\mathbb{R}_+}$, defined by $K_t := \int_0^t (2f(s,Y_s)+1)\,\mathrm{d}s$ for some continuous function $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$. This equation is well-known [20] to possess a solution and satisfy pathwise uniqueness (making every solution strong). In [12], Deelstra & Delbaen construct the solution S of SDE (1.3) (which may *a priori* take values in $\mathbb{R}$ with the volatility coefficient given by $\sqrt{|S_s|}$) from an Euler scheme approximation, and prove that if $S_0 \in \mathbb{R}_+$ then $S_t \in \mathbb{R}_+$ for all $t \in \mathbb{R}_+$ a.s. We stress that, beyond Y and B being adapted to the same filtration, no assumption is made on the dependence between the two processes (or on the dynamics of Y).

Positivity. Consider now the strong solution S of (1.3), driven by B and K and started at $S_0 \in \mathbb{R}_+$. Assume in addition that $1+2f \ge \delta_0$ a.s. for all $t > 0$ and a constant $\delta_0 > 0$. Since the squared-Bessel SDE in (1.6) has pathwise uniqueness, we may construct a squared-Bessel process Z, satisfying SDE (1.6) with "dimension" parameter δ_0, started at $Z_0 = 0$ and driven by the Brownian motion B. Note that by (1.3) and (1.6) the quadratic variation of the semimartingale $S - Z$ equals $\langle S - Z\rangle_t = 4\int_0^t (\sqrt{S_s} - \sqrt{Z_s})^2\,\mathrm{d}s$. In particular, since $(\sqrt{y}-\sqrt{x})^2 \le |x-y|$ for all $x, y \in \mathbb{R}_+$, we have

$$\int_0^t \mathbb{1}_{\{Z_s - S_s > 0\}} \left|\sqrt{S_s} - \sqrt{Z_s}\right|^{-1} \mathrm{d}\langle S - Z\rangle_s \le 4t < \infty \quad \text{a.s. for all } t \in \mathbb{R}_+.$$

By [31, Lem. IX.3.3], the local time $L^0(S-Z)=0$ vanishes. Since $S_0 \geq Z_0 = 0$, the Tanaka formula [31, Thm VI.1.2] on $[0,\tau_N]$, where $\tau_N := \inf\{s \geq 0 : S_s \geq N\}$ for $N > 0$, yields $\mathbb{E}\max\{0, Z_{t\wedge\tau_N} - S_{t\wedge\tau_N}\} \leq 0$ (since $s\delta_0 \leq K_s$ for all $s \in \mathbb{R}_+$ and $x \wedge y := \min\{x,y\}$ for $x, y \in \mathbb{R}$). Thus $Z_{t\wedge\tau_N} \leq S_{t\wedge\tau_N}$, implying $S_t \geq Z_t$ a.s. for all $t \in \mathbb{R}_+$ (since $\lim_{N\uparrow\infty}\tau_N = \infty$).

Bessel-type representation (1.1) of $X = \sqrt{S}$. Since $f \geq 0$, we can always take $\delta_0 = 1$, implying that the modulus of BM (i.e. $\sqrt{Z}$ for $\delta_0 = 1$) bounds $X = \sqrt{S}$ from below, making the point 0 instantaneously reflecting for X. Moreover, if we had $f \geq \varepsilon/2$ for for some $\varepsilon > 0$, a Bessel process $\sqrt{Z}$ with "dimension" $\delta_0 = 1+\varepsilon$, which satisfies the Bessel SDE and, in particular, $\int_0^t \mathrm{d}s/\sqrt{Z_s} < \infty$ for all $t \in \mathbb{R}_+$, would bound X from below. Thus,

$$\int_0^t (f(s,Y_s)/X_s)\,\mathrm{d}s \leq \int_0^t (f(s,Y_s)/\sqrt{Z_s})\,\mathrm{d}s < \infty \quad \text{for all } t \in \mathbb{R}_+. \tag{A.1}$$

Once we know the integral in (A.1) is finite, by considering the excursions of X away from 0, it is not hard to see that the quadratic variation at t of the continuous process $X - \int_0^\cdot (f(s,Y_s)/X_s)\,\mathrm{d}s$ equals t, making it, by Lévy's characterisation, a Brownian motion and thus implying SDE (1.1) for X. However Assumption (A_f) is not consistent with $f \geq \varepsilon/2$ when $\alpha > 0$ (at large t, by (1.2), $f(t,y) \to 0$ as $y \downarrow 0$). But we may assume $f > 0$ on $(0,\infty)\times(0,\infty)$, implying (A.1) on the stochastic interval $[0,\tau_\varepsilon^{(f)})$, where $\tau_\varepsilon^{(f)} := \{t \in \mathbb{R}_+ : f(t,Y_t) = \varepsilon/2\}$ for small $\varepsilon > 0$. Since, for $\delta \geq 2$ and $y > 0$, the point 0 is polar for $Y \sim \mathrm{BESQ}^\delta(y)$, we have $\tau_\varepsilon^{(f)} \to \infty$ as $\varepsilon \downarrow 0$ implying SDE (1.1) for X in this case.

If $\delta \in (0,2)$, Y hits zero a.s., bringing SDE (1.1) into the realm of the Bessel process with $\delta = 1$ (which does not satisfy the corresponding SDE) at times $Y_t = 0$. This case would require the analysis of the joint zeros $S_t = Y_t = 0$, where the dependence in (Y,B) clearly matters.

Acknowledgements

The authors are grateful to Mikhail Menshikov and Vadim Shcherbakov for useful conversations. MB is supported by EPSRC grant EP/V009478/1. CdC and AW are supported by EPSRC grant EP/W00657X/1. AM is supported by EPSRC grants EP/V009478/1 and EP/W006227/1 and, through The Alan Turing Institute, by EP/X03870X/1. Some of the work on this paper was undertaken during the programme "Stochastic systems for anomalous diffusion" (July–December 2024) hosted by the Isaac Newton Institute, under EPSRC grant EP/Z000580/1.

References

[1] E. Barkai, E. Aghion, and D. A. Kessler. From the area under the Bessel excursion to anomalous diffusion of cold atoms. *Phys. Rev. X*, 4:021036, 2014.

[2] B. Bercu. A martingale approach for the elephant random walk. *J. Phys. A*, 51(1):015201, 16, 2018.

[3] J. Bertoin. Counting the zeros of an elephant random walk. *Trans. Amer. Math. Soc.*, 375(8):5539–5560, 2022.

[4] J. Bertoin. On the local times of noise reinforced Bessel processes. *Ann. H. Lebesgue*, 5:1277–1294, 2022.

[5] A. N. Borodin and P. Salminen. *Handbook of Brownian Motion—Facts and Formulae.* Probability and its Applications. Birkhäuser Verlag, Basel, 2nd edition, 2002.

[6] J.-P. Bouchaud and A. Georges. Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications. *Phys. Reports*, 195(4):127–293, 1990.

[7] E. Bouin, J. Dolbeault, and L. Lafleche. Fractional hypocoercivity. *Comm. Math. Phys.*, 390(3):1369–1411, 2022.

[8] M. Brešar, C. da Costa, A. Mijatović, and A. Wade. Anomalous diffusive behaviour for Bessel-driven stochastic kinetics. YouTube presentation, 2023. Published on Prob-AM YouTube channel.

[9] M. Brešar, A. Mijatović, and A. Wade. Brownian motion with asymptotically normal reflection in unbounded domains: from transience to stability. *Ann. Probab.*, (to appear), 2025. ArXiv: 2303.06916.

[10] U. Çetin. On certain integral functionals of squared Bessel processes. *Stochastics*, 87(6):1033–1060, 2015.

[11] C. da Costa, M. Menshikov, V. Shcherbakov, and A. Wade. Superdiffusive planar random walks with polynomial space-time drifts. *Stochastic Process. Appl.*, 176:Paper No. 104420, 17pp., 2024.

[12] G. Deelstra and F. Delbaen. Existence of solutions of stochastic differential equations related to the Bessel process. Unpublished manuscript; available at `https://people.math.ethz.ch/~delbaen/ftp/preprints/existentie_rev.pdf`, 1994.

[13] G. Deelstra and F. Delbaen. Long-term returns in stochastic interest rate models. *Insurance: Mathematics and Economics*, 17(2):163–169, 1995.

[14] R. Durrett. *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 4th edition, 2010.

[15] N. Fournier and C. Tardif. Anomalous diffusion for multi-dimensional critical kinetic Fokker-Planck equations. *Ann. Probab.*, 48(5):2359–2403, 2020.

[16] N. Fournier and C. Tardif. One dimensional critical kinetic Fokker-Planck equations, Bessel and stable processes. *Comm. Math. Phys.*, 381(1):143–173, 2021.

[17] M. Gradinaru and E. Luirard. Asymptotic behavior for a time-inhomogeneous Kolmogorov type diffusion. *ESAIM Probab. Stat.*, 27:1–18, 2023.

[18] M. Gradinaru and Y. Offret. Existence and asymptotic behaviour of some time-inhomogeneous diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(1):182–207, 2013.

[19] O. Hryniv, M. V. Menshikov, and A. R. Wade. Excursions and path functionals for stochastic processes with asymptotically zero drifts. *Stoc. Proc. Appl.*, 123(6):1891–1921, 2013.

[20] J. Jacod and J. Mémin. Weak and strong solutions of stochastic differential equations: existence and stability. In *Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980)*, volume 851 of *Lecture Notes in Math.*, pages 169–212. Springer, Berlin, 1981.

[21] O. Kallenberg. *Foundations of Modern Probability*, volume 99 of *Probability Theory and Stochastic Modelling.* Springer, 3rd edition, 2021.

[22] D. A. Kessler and E. Barkai. Infinite covariant density for diffusion in logarithmic potentials and optical lattices. *Phys. Rev. Lett.*, 105:120602, 2010.

[23] D. Khoshnevisan and Z. Shi. Chung's law for integrated Brownian motion. *Trans. Amer. Math. Soc.*, 350(10):4253–4264, 1998.

[24] G. Lebeau and M. Puel. Diffusion approximation for Fokker Planck with heavy tail equilibria: a spectral method in dimension 1. *Comm. Math. Phys.*, 366(2):709–735, 2019.

[25] M. Lefebvre. Exact solutions to two-dimensional homing problems. *Internat. J. Systems Sci.*, 33(11):879–884, 2002.

[26] M. Lefebvre. Integrated diffusion processes inside rectangles. *Bol. Soc. Mat. Mexicana (3)*, 8(1):171–183, 2002.

[27] M. Lefebvre. Mean first-passage time to zero for wear processes. *Stoch. Models*, 26(1):46–53, 2010.

[28] M. Lefebvre and D. Ait Aoudia. Two-dimensional diffusion processes as models in lifetime studies. *Internat. J. Systems Sci.*, 43(10):1943–1949, 2012.

[29] D. S. Lemons and A. Gythiel. Paul Langevin's 1908 paper "On the theory of Brownian motion". *American J. Physics*, 65(11):1079–1081, 1997.

[30] M. Menshikov and S. Volkov. Urn-related random walk with drift $\rho x^\alpha / t^\beta$. *Electron. J. Probab.*, 13:no. 31, 944–960, 2008.

[31] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, 3rd edition, 2005.