**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

http://wrap.warwick.ac.uk/122281

**How to cite:**

Please refer to published version for the most recent bibliographic citation information.
If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

# Scaling Limits of Processes with Fast Nonlinear Mean Reversion[*]

Thomas Cayé[†]    Martin Herdegen[‡]    Johannes Muhle-Karbe[§]

June 3, 2019

## Abstract

We derive scaling limits for integral functionals of Itô processes with fast nonlinear mean-reversion speeds. In these limits, the fast mean-reverting process is "averaged out" by integrating against its invariant measure. The convergence is uniformly in probability and, under mild integrability conditions, also in $\mathcal{S}^p$. These results are a crucial building block for the analysis of portfolio choice models with small superlinear transaction costs, carried out in the companion paper of the present study [11].

**Mathematics Subject Classification (2010):** 60F25, 60H10.

**Keywords:** processes with fast nonlinear mean reversion; scaling limits

## 1 Introduction and Main Results

**Motivation**   Superlinear trading costs play an important role in financial engineering as reduced-form models for the adverse price impact generated by large trades, cf., e.g., [3, 2] as well as many more recent studies. In this context, optimal policies typically prescribe to track some "target portfolio" at a finite, absolutely-continuous rate [15, 16, 6, 9, 10, 4, 25, 17, 18]. If trading costs are quadratic, this trading speed is linear in the deviation from the target [15, 16, 25, 17], leading to Ornstein-Uhlenbeck dynamics of the deviation in the small-cost limit. The well-known properties of this process can in turn be used to analyze the asymptotic performance of the corresponding tracking portfolios [9, 10].

However, empirical studies suggest that actual trading costs are superlinear but also sub-quadratic, corresponding to the "square-root law" for price impact advocated by many practitioners [24, 5]. For such trading costs, optimal trading rates become nonlinear: compared to the quadratic case, trading slows down near the target, where costs are higher. Conversely, trading is sped up far away from the target, where costs are comparatively lower. In the limiting case where the trading costs become proportional, this leads to "reflecting" singular controls: no transactions at

all inside some "no-trade region" and instantaneous reflection by trading at an "infinite rate" once its boundaries are breached [12]. The controlled deviation from the target process in turn follows a reflected diffusion process. In the small-cost limit, its study boils down to the analysis of doubly-reflected Brownian motion. Whence, the asymptotic analysis of the corresponding singular tracking strategies can also be performed by appealing to well-known probabilitic results [20, 1, 9, 10]. In addition to these law of large numbers type results, even a central limit theorem has been derived for the asymptotic performance in this context [8].

However, such results are not available for the empirically most relevant transaction costs that fall between linear and quadratic. The corresponding controlled deviations then correspond to processes with *nonlinear* mean-reversions speeds [18, 11]. In this paper, we determine the scaling limits of such processes in the regime where the mean-reversion speed becomes large. This is a main building block for the derivation of asymptotically optimal trading strategies with small superlinear trading costs in the companion paper of the present study [11]. There, the results from the present paper are used to compute both the expected transaction costs incurred by a given tracking strategy, and its average squared displacement from the target. Trading off these two terms in an optimal manner in turn leads to the asymptotically optimal performance.

Our scaling limits also contribute to the classical literature on "averaging results", where a "fast variable" is averaged out appropriately as it oscillates faster and faster. Results of this kind were first developed by [34, 35, 23, 26]; textbook treatments can be found in [13, 33]. Our probabilistic approach allows us to extend existing results of, e.g, [27, 28] that require a Markovian or semimartingale structure on the drift and diffusion coefficients of the SDEs to general unbounded drift and diffusion coefficients.

**Setting**    Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. For a (small) parameter $\varepsilon > 0$, we consider the following family of mean-reverting stochastic differential equations (SDEs):

$$dX_t^\varepsilon = \left(b_t - \frac{c_t L_t}{\varepsilon} g\left(\frac{M_t X_t^\varepsilon}{\varepsilon}\right)\right) dt + \sqrt{c_t} dW_t, \quad X_0^\varepsilon = x_0^\varepsilon, \quad t \in [0, T]. \tag{1.1}$$

Here, $W$ is a standard Brownian motion, the processes $b, c, L, M$ are adapted and continuous and $c, L, M$ are positive. The function $g$ describes the nonlinear nature of the mean reversion. It is locally Lipschitz, odd, non-decreasing and nonnegative on $\mathbb{R}_+$ (so that $X^\varepsilon$ is indeed always steered back towards zero), and of superlinear polynomial growth at infinity:[1]

$$\liminf_{x\to\infty} \frac{g(x)}{x^q} > 0, \quad \text{for some } q \geqslant 1. \tag{1.2}$$

(Larger values of $q$ correspond to faster mean-reversion for large deviations, which allow to weaken the integrability requirements that need to be imposed on the other primitives of the model in Assumption 2.) The antiderivative of $g$ is denoted by $G(x) = \int_0^x g(y)dy$. The processes $b, c, L, M$ and the function $g$ are all independent of the scaling parameter $\varepsilon$. In contrast, the initial value $x_0^\varepsilon \in \mathbb{R}$ of $X^\varepsilon$ may depend on $\varepsilon$ as long as $\limsup_{\varepsilon\to0} |x_0^\varepsilon|/\varepsilon < \infty$.

**Interpretation**    In the context of portfolio optimization with small nonlinear trading costs [11], $X^\varepsilon$ corresponds to the deviation of the frictional portfolio process from its frictionless counterpart.

---

[1] In sufficiently regular Markovian settings, weaker growth conditions suffice to ensure the recurrence of the diffusion $X^\varepsilon$, cf. [27]). Weaker versions of our results under such conditions are discussed in Remark 1.5.

$b$ and $c$ correspond to the drift and diffusion coefficients of this target position, whereas the mean-reverting part of the drift of (1.1) is the absolutely continuous control applied to steer the actual position in its direction. Up to rescaling, the asymptotic parameter $\varepsilon$ corresponds to the size of the trading cost. As it decreases, the mean-reversion becomes faster and faster and the frictional positions eventually converge to their frictionless counterparts.

In order to determine asymptotically optimal portfolios, the average squared values of the deviations (1.1) need to be traded off against the corresponding trading costs (a nonlinear functional of the control that is applied). In the present paper, we develop limit theorems that allow, in particular, to compute both of these terms in closed form at the leading-order for small $\varepsilon$.

**Results**  We first establish that the SDE (1.1) is well posed despite the superlinear growth of its drift rate at infinity; for better readability, the proof of this result is deferred to Section 3.

**Proposition 1.1.** *For each $\varepsilon > 0$, and $x_0^\varepsilon \in \mathbb{R}$, there exists a unique strong solution of the SDE (1.1).*

In order to formulate our scaling limits for the quickly mean-reverting processes (1.1), we fix a function $f : \mathbb{R} \to \mathbb{R}$ that is even,[2] of finite variation on compacts and satisfies the following polynomial growth condition:

$$|f(x)| \leqslant C_f(|x|^{q'} + 1), \quad \text{for some } q' \geqslant 0 \text{ and some } C_f > 0. \tag{1.3}$$

We can now formulate our first scaling limit, whose proof is delegated to Section 4. Since this first result only asserts convergence in probability, it does not require any integrability assumptions on the primitives of the SDE (1.1).

**Theorem 1.2.** *Let $(H_t)_{t \in [0,T]}$ and $(K_t)_{t \in [0,T]}$ be nonnegative, continuous, adapted processes. Then, as $\varepsilon \to 0$, the following limit holds uniformly in probability:*

$$\int_0^\cdot H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds \longrightarrow \int_0^\cdot H_s \frac{\int_\mathbb{R} f\left(\frac{K_s}{M_s} y\right) \exp\left(-2\frac{L_s}{M_s} G(y)\right) dy}{\int_\mathbb{R} \exp\left(-2\frac{L_s}{M_s} G(y)\right) dy} ds. \tag{1.4}$$

The intuition for the limit (1.4) is the following. As the mean-reversion speed of the processes (1.1) becomes faster and faster, one can essentially treat the "slow" processes $H$, $K$, $b$, $c$, $L$, $M$ as constant on each time step in a fine partition of $[0, T]$. In contrast, the rescaled process $X^\varepsilon/\varepsilon$ converges to a mean-reverting one-dimensional diffusion on each of these infinitesimal intervals. The limit (1.4) asserts that, as $\varepsilon \to 0$, one can replace $X^\varepsilon/\varepsilon$ in the integrand by an integral with respect to the stationary distribution of this limiting process, which is given in terms of its normalized and rescaled speed measure.[3]

Results similar to Theorem 1.2 have been developed under abstract assumptions and verified for the simplest case of linear mean-reversion speeds in [10, Section 3.2]. Here, we extend this to the nonlinear trading speeds arising naturally in the context of square-root price impact and show that convergence in probability remains valid under minimal assumptions also in this case.

---

[2]The assumption that $f$ is even is crucially used in the arguments leading to Lemma 4.2. In its proof, we bound the process $(X^\varepsilon)^2$ from above and below by positive processes, and the fact that $f$ is even allows us in turn to derive bounds for $f(X^\varepsilon)$.

[3]In the limiting case where the rescaled process is a reflected diffusion, the stationary law is uniform, as exploited in [1, 9, 10]. Conversely, if the rescaled process has linear mean reversion, it is Gaussian, compare [9, 10]. More general averaging results for Markovian settings are developed in [27, 28], for example.

Our second main result provides conditions under which this limit theorem can be lifted to convergence in expectation. (For better readability, the proof is delegated to Section 4.) This is needed to study the small-cost asymptotics of expected utility maximization problems rather than the pathwise, quadratic criteria of [9]. Unlike for singularly- or impulse-controlled deviations (where integrability is inherited directly from the corresponding trading boundaries [1, 14, 19]), this necessitates further delicate estimates that require the following integrability assumptions on the processes $b, c, L, M, H, K$ appearing in the SDE for $X^\varepsilon$ and the scaling limit (1.4):[4]

**Assumption 1.** $\mathbb{E}[\exp(8\int_0^T \frac{b_t^2}{c_t}dt)] < \infty$.

**Assumption 2.** *For $p \geqslant 1$, there exists $\eta > 0$ such that*

$$\mathbb{E}\left[\sup_{u\in[0,T]} (L_u c_u \wedge M_u)^{-\frac{2(q+1)(4pq'(1+\eta)\vee 2)}{q-1}}\right] + \mathbb{E}\left[\sup_{u\in[0,T]} c_u^{8pq'(1+\eta)\vee 4}\right]$$

$$+ \mathbb{E}\left[\int_0^T H_t^{p(1+\eta)}dt\right] + \mathbb{E}\left[\int_0^T \left(H_t K_t^{q'}\right)^{2p(1+\eta)}dt\right] + \mathbb{E}\left[\int_0^T \left(\frac{1}{L_s \wedge M_s}\right)^{4pq'(1+\eta)}dt\right]$$

$$+ \mathbb{E}\left[\int_0^T \left(\frac{M_t}{L_t}\right)^{4p(1+\eta)}dt\right] + \mathbb{E}\left[\int_0^T \exp\left(\eta\frac{L_t}{M_t}\right)dt\right] < \infty.$$

**Remark 1.3.** If $q = 1$ in (1.2), then the first moment condition of Assumption 2 is to be understood as $\operatorname{essinf}_{u\in[0,T]}(L_u c_u \wedge M_u) > 0$.

We now turn to our second main result, which is the main tool for the analysis of asymptotically optimal trading strategies with nonlinear trading costs in the companion paper of the present study [11]:

**Theorem 1.4.** *Suppose that Assumptions 1 and 2 are satisfied for some $p \geqslant 1$. Then the scaling limit (1.4) also holds in $\mathcal{S}^p([0,T])$.*[5]

**Remark 1.5.** (i) For sufficiently regular Markovian settings, Theorem 1.2 with convergence in probability for all $t \in [0,T]$ and Theorem 1.4 with convergence in $L^p$ instead of $\mathcal{S}^p$ are consequences of [28, Theorem 4].

(ii) [28, Theorem 4] requires weaker growth conditions for the function $g$ than (1.2). Under additional assumptions on the diffusion coefficient of $X^\varepsilon$, our results can be extended in this direction. (We are grateful to one of the referees for pointing this out.) To illustrate this, suppose that $c_t = \sigma^2 > 0$ is constant and that, instead of (1.2) for $g$, there exists $C, \tilde{C} > 0$ such that, for all $x > \tilde{C}$,

$$L_t g(M_t x) \geqslant C, \quad t \in [0,T], \quad \text{and} \quad G(x) \geqslant Cx. \tag{1.5}$$

Assume moreover that the integrability conditions from Assumption 2 are satisfied (except for the first two terms, which are not needed in this case). Then, the convergence result in Theorem 1.2 holds in probability for all $t \in [0,T]$ instead of uniformly on compacts in probability. Likewise, the convergence in Theorem 1.4 holds in $L^p$ instead of in $\mathcal{S}^p$.

---

[4]Assumptions 1 and 2 are required to apply Hölder's inequality in Lemmas A.1, A.3, B.4, and D.6.

[5]$\mathcal{S}^p([0,T])$ denotes the set of càdlàg adapted processes whose running supremum on $[0,T]$ has finite absolute $p$-th moments. A sequence of processes $(Y^{(n)})_{n\in\mathbb{N}}$ converges to $Y$ in $\mathcal{S}^p([0,T])$ if $\lim_{n\to\infty} \mathbb{E}[\sup_{t\in[0,T]} |Y_t^{(n)} - Y_t|^p] = 0$. Assumption 2 guarantees that the limit in Theorem 1.4 is finite, cf. Lemma D.6.

The above extension heavily relies on the comparison theorem for SDEs; cf. Remark 4.5. For more general diffusion coefficients, one would have to extend the arguments in Veretennikov [36] to non-constant volatilities. As the latter will typically depend on the asymptotic parameter $\varepsilon$, this is rather challenging.

The remainder of this article is organized as follows. Section 2 describes a localization argument that allows to reduce the analysis to the case of bounded coefficients. This crucially relies on a uniform integrability result (Lemma A.3) established in Appendix A. Section 3 contains the proof of Proposition 1.1 and introduces the fundamental time and measure changes used throughout the rest of the paper. Section 4 contains the proof of our main result, Theorem 1.4; due to the localization argument from Section 2, the proof of Theorem 1.2 turns out to be a simple corollary. The proof of Theorem 1.4 is rather delicate. First, in Section 4.1 we establish a local scaling limit on a small interval. This combines sandwiching and approximation arguments with ergodic theory for one-dimensional diffusions. A key difficulty for the ergodic result in Lemma 4.3 is that the diffusions under consideration also depend on the time horizon (via the small parameter $\varepsilon$). In a second step, we concatenate the local limit theorems to a global limit theorem in Section 4.2. Appendix A contains the already mentioned result on uniform integrability. Appendix B proves a maximal inequality in the spirit of Peskir [29] necessary for the concatenation argument in Section 4.2. Appendix C establishes some comparison and existence results for SDEs, and Appendix D concludes with some other auxiliary results.

## 2 Reduction to Bounded Coefficients

In this section, we show why – up to the results on uniform integrability in Appendix A – it suffices to establish all results for the case of bounded coefficients. For fixed $\kappa \in (0, 1)$, define the stopping time $\tau^\kappa$ by

$$\tau^\kappa := \inf \left\{ t \in [0, T] : \int_0^t \frac{b_u^2}{c_u} du > \frac{1}{\kappa}, H_t > \frac{1}{\kappa}, K_t > \frac{1}{\kappa}, \right.$$
$$\left. c_t \notin \left[ \kappa, \frac{1}{\kappa} \right], L_t \notin \left[ \kappa, \frac{1}{\kappa} \right], M_t \notin \left[ \kappa, \frac{1}{\kappa} \right] \right\} \wedge T. \qquad (2.1)$$

Note that the stopping times $\tau^\kappa$ are non-increasing in $\kappa$, and for every $\omega$ outside of a null set, there is a $\kappa(\omega) > 0$ such that $\tau^\kappa = T$ for $0 < \kappa \leqslant \kappa(\omega)$, by continuity of $b, c, H, K, L, M$ and positivity of $c, L, M$. This in turn implies that

$$\lim_{\kappa \to 0} \mathbb{P}[\tau^\kappa = T] = 1. \qquad (2.2)$$

For $\Upsilon \in \{b, c, H, K, L, M\}$, introduce the stopped processes

$$\Upsilon_t^\kappa := \Upsilon_t^{\tau^\kappa}, \quad t \in [0, T],$$

and consider for fixed $\varepsilon > 0$ the corresponding SDE

$$dX_t^{\varepsilon,\kappa} = \left( b_t^\kappa - \frac{c_t^\kappa L_t^\kappa}{\varepsilon} g \left( \frac{M_t^\kappa X_t^{\varepsilon,\kappa}}{\varepsilon} \right) \right) dt + \sqrt{c_t^\kappa} dW_t, \quad X_0^{\varepsilon,\kappa} = x_0^\varepsilon, \quad t \in [0, T]. \qquad (2.3)$$

Note that the SDE for $X^{\varepsilon,\kappa}$ coincides with the SDE (1.1) for $X^\varepsilon$ on $[\![0, \tau^\kappa]\!]$.

Now suppose that Proposition 1.1 as well as Theorems 1.2 and 1.4 have been established for $X^{\varepsilon,\kappa}$, $H^\kappa$, and $K^\kappa$ for each fixed $\kappa \in (0, 1)$. Then Proposition 1.1 for $X^\varepsilon$ follows from (2.2) and pathwise uniqueness of strong solutions. In particular, we have

$$X^\varepsilon = X^{\varepsilon,\kappa} \text{ on } [\![0, \tau^\kappa]\!]. \qquad (2.4)$$

5

Next, to establish Theorem 1.2 for $X^\varepsilon$, set

$$w_t := \frac{\int_\mathbb{R} f\left(\frac{K_t}{M_t} y\right) \exp\left(-2\frac{L_t}{M_t} G(y)\right) dy}{\int_\mathbb{R} \exp\left(-2\frac{L_t}{M_t} G(y)\right) dy}, \quad w_t^\kappa := \frac{\int_\mathbb{R} f\left(\frac{K_t^\kappa}{M_t^\kappa} y\right) \exp\left(-2\frac{L_t^\kappa}{M_t^\kappa} G(y)\right) dy}{\int_\mathbb{R} \exp\left(-2\frac{L_t^\kappa}{M_t^\kappa} G(y)\right) dy},$$

for $t \in [0, T]$ and $\kappa \in (0, 1)$. Then (2.4) and Theorem 1.2 for each $X^{\varepsilon,\kappa}$ give

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\sup_{t \in [0,T]} \left| \int_0^t H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds - \int_0^t H_s w_s ds \right| \wedge 1 \right]$$

$$\leqslant \lim_{\varepsilon \to 0} \mathbb{E}\left[ \left( \sup_{t \in [0,T]} \left| \int_0^t H_s^\kappa f\left(\frac{K_s^\kappa X_s^{\varepsilon,\kappa}}{\varepsilon}\right) ds - \int_0^t H_s^\kappa w_s^\kappa ds \right| \wedge 1 \right) \mathbb{1}_{\{\tau^\kappa = T\}} \right] + \mathbb{P}[\tau^\kappa < T]$$

$$\leqslant \mathbb{P}[\tau^\kappa < T].$$

Now Theorem 1.2 for $X^\varepsilon$ follows from one of the equivalent characterizations of the convergence in probability (cf. [21, p. 63]) by letting $\kappa \to 0$ and using (2.2).

Finally, by (2.4), Theorem 1.4 for each $X^{\varepsilon,\kappa}$, and Hölder's inequality, we obtain

$$\lim_{\varepsilon \to 0} \mathbb{E}\left[\sup_{t \in [0,T]} \left| \int_0^t H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds - \int_0^t H_s w_s ds \right|^p \right]$$

$$\leqslant \lim_{\varepsilon \to 0} \mathbb{E}\left[ \sup_{t \in [0,T]} \left| \int_0^t H_s^\kappa f\left(\frac{K_s^\kappa X_s^{\varepsilon,\kappa}}{\varepsilon}\right) ds - \int_0^t H_s^\kappa w_s^\kappa ds \right|^p \mathbb{1}_{\{\tau^\kappa = T\}} \right]$$

$$\quad + \sup_{\varepsilon > 0} \mathbb{E}\left[ \sup_{t \in [0,T]} \left| \int_0^t H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds - \int_0^t H_s w_s ds \right|^p \mathbb{1}_{\{\tau^\kappa < T\}} \right]$$

$$\leqslant \sup_{\varepsilon > 0} \mathbb{E}\left[ \left( \int_0^T H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds + \int_0^T H_s w_s ds \right)^p \mathbb{1}_{\{\tau^\kappa < T\}} \right]$$

$$\leqslant \sup_{\varepsilon > 0} \mathbb{E}\left[ \left( \int_0^T H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds + \int_0^T H_s w_s ds \right)^{p(1+\eta)} \right]^{\frac{1}{1+\eta}} \mathbb{P}[\tau^\kappa < T]^{\frac{\eta}{1+\eta}},$$

with $\eta$ as in Assumption 2. Now Theorem 1.4 for $X^\varepsilon$ follows by letting $\kappa \to 0$, using (2.2) and noting that, under Assumption 1 and 2,

$$\sup_{\varepsilon > 0} \mathbb{E}\left[ \left( \int_0^T H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds + \int_0^T H_s w_s ds \right)^{p(1+\eta)} \right] < \infty. \tag{2.5}$$

Indeed, using the elementary inequality $(a + b)^{p(1+\eta)} \leqslant 2^{p(1+\eta)}(a^{p(1+\eta)} + b^{p(1+\eta)})$ for $a, b \geqslant 0$, (2.5) follows from

$$\sup_{\varepsilon > 0} \mathbb{E}\left[ \left( \int_0^T H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds \right)^{p(1+\eta)} \right] < \infty \quad \text{and} \quad \mathbb{E}\left[ \left( \int_0^T H_s w_s ds \right)^{p(1+\eta)} \right] < \infty.$$

Here, the first estimate follows from Jensen's inequality and Lemma A.3, the second estimate follows from Lemma D.6.

In summary, it therefore remains to establish Proposition 1.1 and Theorem 1.4, respectively, for uniformly bounded coefficients in order to prove Proposition 1.1 and Theorem 1.2. To establish Theorem 1.4 for general coefficients, it additionally remains to be shown that the integrability conditions from Assumptions 1 and 2 imply Lemmas A.3 and D.6.

6

# 3 Proof of Proposition 1.1

In this section, we establish that the SDE (1.1) has a unique strong solution on $[0, T]$ for each fixed $\varepsilon > 0$ and $x_0^\varepsilon \in \mathbb{R}$ given that the processes $b$, $c$, $L$ and $M$ are bounded from above by $1/\kappa$ and $c$, $L$ and $M$ are bounded from below by $\kappa$, for some $\kappa > 0$. By the localization argument from Section 2, this assumption is without loss of generality in the context of Proposition 1.1.

First, note that it suffices to show that the SDE (1.1) has a unique strong solution after a bijective time change, after which the SDE has constant volatility. To this end, for fixed $\varepsilon > 0$, set

$$\xi^\varepsilon = \int_0^T \varepsilon^{-2} c_t \, dt, \tag{3.1}$$

and make the following standard observation:

**Lemma 3.1.** *For each $\varepsilon > 0$, the family of stopping times (indexed by $\xi$)*

$$u_\xi^\varepsilon := \begin{cases} \inf\left\{ s \in \mathbb{R}_+ : \int_0^s \varepsilon^{-2} c_r \, dr > \xi \right\}, & \xi \leqslant \xi^\varepsilon, \\ T, & \xi > \xi^\varepsilon, \end{cases}$$

*is strictly increasing in $\xi$ on $[0, \xi^\varepsilon(\omega)]$ for almost every $\omega \in \Omega$, forms a stochastic time change, and satisfies $u_{\xi^\varepsilon}^\varepsilon = T$. Moreover, for almost every $\omega \in \Omega$, $\xi \mapsto u_\xi^\varepsilon(\omega)$ is differentiable with derivative $\varepsilon^2 / c_{u_\xi^\varepsilon(\omega)}(\omega)$ on $[0, \xi^\varepsilon(\omega)]$.*

We proceed to define a time-changed Brownian motion. Set

$$\widetilde{W}_\xi^\varepsilon = \int_0^{u_\xi^\varepsilon} \varepsilon^{-1} \sqrt{c_s} \, dW_s, \quad \xi \geqslant 0. \tag{3.2}$$

This is a $\mathbb{P}$-Brownian motion, stopped at $\xi^\varepsilon$, relative to the filtration $\mathbb{G}^\varepsilon = (\mathcal{G}_\xi^\varepsilon)_{\xi \in \mathbb{R}_+}$ with $\mathcal{G}_\xi^\varepsilon = \mathcal{F}_{u_\xi^\varepsilon}$. By Lemma 3.1 and Itô's formula, it suffices to show that the process $(\widetilde{X}_\xi^\varepsilon)_{\xi \geqslant 0}$, defined by

$$\widetilde{X}_\xi^\varepsilon := X_{u_\xi^\varepsilon}^\varepsilon / \varepsilon, \quad \xi \geqslant 0, \tag{3.3}$$

is the unique strong solution of the SDE

$$d\widetilde{X}_\xi^\varepsilon = \left( \varepsilon \frac{b_{u_\xi^\varepsilon}}{c_{u_\xi^\varepsilon}} - L_{u_\xi^\varepsilon} g\left( M_{u_\xi^\varepsilon} \widetilde{X}_\xi^\varepsilon \right) \right) \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\xi + \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\widetilde{W}_\xi^\varepsilon, \qquad \widetilde{X}_0^\varepsilon = x_0^\varepsilon / \varepsilon. \tag{3.4}$$

Next, observe that it suffices to show that the SDE (3.4) has a unique strong solution on $[\![0, \xi^\varepsilon]\!]$ (after $\xi^\varepsilon$, $\widetilde{X}^\varepsilon$ is trivially constant) under a measure $\mathbb{Q}$ that is equivalent to $\mathbb{P}$ on $\mathcal{F}_T = \mathcal{G}_{\xi^\varepsilon}$. Set

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \exp\left( -\int_0^T \frac{b_t}{\sqrt{c_t}} dW_t - \frac{1}{2} \int_0^T \frac{(b_t)^2}{c_t} dt \right) := \exp\left( -\int_0^{\xi^\varepsilon} \varepsilon \frac{b_{u_\xi^\varepsilon}}{c_{u_\xi^\varepsilon}} d\widetilde{W}_\xi^\varepsilon - \frac{1}{2} \int_0^{\xi^\varepsilon} \varepsilon^2 \left( \frac{b_{u_\xi^\varepsilon}}{c_{u_\xi^\varepsilon}} \right)^2 d\xi \right), \tag{3.5}$$

so that the first part of the drift of (3.4) is absorbed by the corresponding change of measure. (Note that $\mathbb{Q}$ is well defined by Novikov's condition given that $\int_0^t \frac{b_u^2}{c_u} du \leq \frac{1}{\kappa}$.) By Girsanov's Theorem,

$$\widetilde{W}_\xi^{\varepsilon, \mathbb{Q}} = \widetilde{W}_\xi^\varepsilon + \int_0^{\xi \wedge \xi^\varepsilon} \varepsilon \frac{b_{u_y^\varepsilon}}{c_{u_y^\varepsilon}} dy, \quad \xi \geqslant 0, \tag{3.6}$$

in turn is a $\mathbb{Q}$-Brownian motion, stopped at $\xi^\varepsilon$, relative to the filtration $\mathbb{G}^\varepsilon$. Thus, it suffices to show that there is a unique strong solution of

$$d\widetilde{X}^\varepsilon_\xi = -L_{u^\varepsilon_\xi} g\left(M_{u^\varepsilon_\xi} \widetilde{X}^\varepsilon_\xi\right) \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\xi + \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\widetilde{W}^{\varepsilon,\mathbb{Q}}_\xi, \qquad \widetilde{X}^\varepsilon_0 = x^\varepsilon_0/\varepsilon. \tag{3.7}$$

This is established in the following result:

**Proposition 3.2.** *For each $\varepsilon > 0$, there is a unique strong solution $\widetilde{X}^\varepsilon$ of the SDE* (3.7).

*Proof.* For $n \in \mathbb{N}$, define the bounded function

$$g^{(n)}(x) = \text{sgn}(x) \left(|g(x)| \wedge n\right),$$

and consider the same SDE as above but with truncated drift,

$$d\widetilde{X}^{\varepsilon,n}_\xi = -L_{u^\varepsilon_\xi} g^{(n)}\left(M_{u^\varepsilon_\xi} \widetilde{X}^\varepsilon_\xi\right) \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\xi + \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\widetilde{W}^{\varepsilon,\mathbb{Q}}_\xi. \tag{3.8}$$

The function $g^{(n)}$ is Lipschitz continuous with Lipschitz constant $K_n$ (because $g$ is locally Lipschitz and $\{g \leqslant n\}$ is compact). As a consequence, the (random) function $f : (\xi, \omega, x) \mapsto -L_{u^\varepsilon_\xi} g^{(n)}(M_{u^\varepsilon_\xi} x)$ satisfies

$$|f(\xi, \omega, x) - f(\xi, \omega, y)| \leqslant K_n L_{u^\varepsilon_\xi}(\omega) M_{u^\varepsilon_\xi}(\omega) |x - y|, \quad \text{for all } (\xi, \omega, x, y) \in \mathbb{R}_+ \times \Omega \times \mathbb{R}^2.$$

The random variable $K = K_n \sup_{t \in [0,T]} L_t M_t$ is almost surely finite by continuity of $L$ and $M$ on $[0, T]$. Hence, [30, Theorem V.7] shows that there exists a unique strong solution of the truncated SDE (3.8). Define the stopping time

$$\xi^{\varepsilon,n} := \inf\left\{\xi \in [0, \xi^\varepsilon] : \left|M_{u^\varepsilon_\xi} X^{\varepsilon,n}_\xi\right| \geqslant g^{-1}(n)\right\} \wedge \xi^\varepsilon, \tag{3.9}$$

where $g^{-1}(n) = \inf\{x > 0 : g(x) > n\}$. On $[\![0, \xi^{\varepsilon,n}]\!]$, the processes $\widetilde{X}^\varepsilon$ and $\widetilde{X}^{\varepsilon,n}$ satisfy the same SDE with the same initial condition and are therefore indistinguishable. The squared truncated process $(\widetilde{X}^{\varepsilon,n}_\xi)^2_{\xi \in \mathbb{R}_+}$ has dynamics

$$d\left(\widetilde{X}^{\varepsilon,n}_\xi\right)^2 = \left(1 - 2L_{u^\varepsilon_\xi} \widetilde{X}^{\varepsilon,n}_\xi g^{(n)}\left(M_{u^\varepsilon_\xi} \widetilde{X}^{\varepsilon,n}_\xi\right)\right) \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\xi + 2\sqrt{\left(\widetilde{X}^{\varepsilon,n}_\xi\right)^2} \text{sgn}\left(\widetilde{X}^{\varepsilon,n}_\xi\right) \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\widetilde{W}^{\varepsilon,\mathbb{Q}}_\xi.$$

By Lévy's characterisation of Brownian motion (cf. [22, Theorem 3.3.16]), the process

$$\widetilde{B}^{\varepsilon,\mathbb{Q},n}_\xi = \int_0^\xi \text{sgn}(\widetilde{X}^{\varepsilon,n}_y) d\widetilde{W}^{\varepsilon,\mathbb{Q}}_y$$

is a $\mathbb{Q}$-Brownian motion, stopped at time $\xi^\varepsilon$. Moreover, $x \mapsto xg^{(n)}(x)$ is an even function. Therefore, we can rewrite the dynamics of the squared truncated process as

$$d\left(\widetilde{X}^{\varepsilon,n}_\xi\right)^2 = \left(1 - 2L_{u^\varepsilon_\xi} \sqrt{\left(\widetilde{X}^{\varepsilon,n}_\xi\right)^2} g^{(n)}\left(M_{u^\varepsilon_\xi} \sqrt{\left(\widetilde{X}^{\varepsilon,n}_\xi\right)^2}\right)\right) \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\xi + 2\sqrt{\left(\widetilde{X}^{\varepsilon,n}_\xi\right)^2} \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\widetilde{B}^{\varepsilon,\mathbb{Q},n}_\xi.$$

Let $Y^{\varepsilon,n}$ be the unique strong solution of

$$dY^{\varepsilon,n}_\xi = \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\xi + 2\sqrt{Y^{\varepsilon,n}_\xi} \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\widetilde{B}^{\varepsilon,\mathbb{Q},n}_\xi, \qquad Y^{\varepsilon,n}_0 = (x^\varepsilon_0)^2/\varepsilon^2.$$

This process is – for each $n$ – the square of a 1-dimensional Bessel process started at $(x^\varepsilon_0)^2/\varepsilon^2$ and stopped at time $\xi^\varepsilon$ (cf. [31, Definition XI.1.1]). In particular, it is a submartingale that has finite

8

moments at all bounded stopping times – independent of $n$; see [31, Chapter XI]. The comparison theorem for SDEs in the form of Lemma C.1 yields

$$\mathbb{Q}\left[Y_\xi^{\varepsilon,n} \geqslant \big(\widetilde{X}_\xi^{\varepsilon,n}\big)^2, \text{ for all } \xi \in \mathbb{R}_+\right] = 1.$$

Hence, for $y > 0$ and $n \in \mathbb{N}$, by the definition of $\xi^{\varepsilon,n}$, the above comparison argument, the bound on $M$ and Doob's maximal inequality applied to the non-negative submartingale $Y^{\varepsilon,n}$, we obtain

$$\mathbb{Q}\left[\xi^{\varepsilon,n} < y \wedge \xi^\varepsilon\right] \leqslant \mathbb{Q}\left[\sup\left\{\left|M_{u_\xi^\varepsilon}\widetilde{X}_\xi^{\varepsilon,n}\right| : \xi \in [0, y \wedge \xi^\varepsilon]\right\} > g^{-1}(n)\right]$$

$$\leqslant \mathbb{Q}\left[\sup\left\{\left(\widetilde{X}_\xi^{\varepsilon,n}\right)^2 : \xi \in [0, y \wedge \xi^\varepsilon]\right\} > \kappa^2 g^{-1}(n)^2\right]$$

$$\leqslant \mathbb{Q}\left[\sup\left\{Y_\xi^{\varepsilon,n} : \xi \in [0, y \wedge \xi^\varepsilon]\right\} > \kappa^2 g^{-1}(n)^2\right] \leqslant \frac{\mathbb{E}_\mathbb{Q}\left[Y_{y \wedge \xi^\varepsilon}^{\varepsilon,n}\right]}{\kappa^2 g^{-1}(n)^2} \leqslant \frac{\mathrm{BES1}(y)}{\kappa^2 g^{-1}(n)^2},$$

where $\mathrm{BES1}(y)$ denotes the expectation at time $y$ of the square of a 1-dimensional Bessel process started at $(x_0^\varepsilon)^2/\varepsilon^2$. Letting $n \to \infty$ and using that $\lim_{n\to\infty} g^{-1}(n) = \infty$ shows that for arbitrary $x > 0$,

$$\mathbb{Q}\left[\lim_{n\to\infty} \xi^{\varepsilon,n} \wedge x \wedge \xi^\varepsilon = x \wedge \xi^\varepsilon\right] = 1.$$

Therefore, the solution of (3.7) exists $\mathbb{Q}$-a.s. on $\mathbb{R}_+$ and in particular, on $[\![0, \xi^\varepsilon]\!]$. The solution on $\mathbb{R}_+$ is unique as it coincides $\mathbb{Q}$-a.s. with all the solutions on the smaller interval $[\![0, \xi^{\varepsilon,n}]\!]$. These solutions are unique by global existence and uniqueness for functionally Lipschitz SDEs [30, Theorem V.7]. □

The next result holds for processes as in our main setting ($b$ adapted and locally bounded, $c$, $L$ and $M$ adapted, continuous and positive), provided Assumption 1 is satisfied.

**Proposition 3.3.** *Suppose that Assumption 1 is satisfied. Then for each $k \geq 0$,*

$$\sup_{s\in[0,t]} \mathbb{E}\left[\left|\widetilde{X}_s^\varepsilon\right|^k\right] < \infty, \quad \text{for all } t \geq 0. \tag{3.10}$$

*Proof.* This assertion follows by a similar argument as in the proof of Proposition 3.2. Here, we compare $(\widetilde{X}^\varepsilon)^2$ to the unique strong solution of the SDE

$$dY_\xi^{(\varepsilon)} = \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\xi + 2\sqrt{Y_\xi^{(\varepsilon)}}\, \mathbb{1}_{\{\xi \leqslant \xi^\varepsilon\}} d\widetilde{B}_\xi^{\varepsilon,\mathbb{Q}}, \qquad Y_0^{(\varepsilon)} = (x_0^\varepsilon)^2/\varepsilon^2,$$

where

$$\widetilde{B}_\xi^{\varepsilon,\mathbb{Q}} = \int_0^\xi \mathrm{sgn}\big(\widetilde{X}_y^\varepsilon\big) d\widetilde{W}_y^{\varepsilon,\mathbb{Q}}$$

is a $\mathbb{Q}$-Brownian motion stopped at $\xi^\varepsilon$. We then use Doob's maximal inequality and that the square of the 1-dimensional Bessel process is a submartingale and has finite moments of all orders at all finite times [31, Chapter XI]. Combined with Hölder inequality and Assumption 1, this in turn yields (3.10) and thereby completes the proof. □

# 4 Proofs of Theorems 1.2 and 1.4 for Bounded Coefficients

We now turn to the proof of our main results, Theorems 1.2 and 1.4. By the localization argument from Section 2, we can and will assume throughout without loss of generality that there is $\kappa \in (0, 1)$ such that

$$\int_0^T \frac{b_u^2}{c_u} du \leqslant \frac{1}{\kappa}, H_t \leqslant \frac{1}{\kappa}, K_t \leqslant \frac{1}{\kappa}, c_t \in \left[\kappa, \frac{1}{\kappa}\right], L_t \in \left[\kappa, \frac{1}{\kappa}\right], M_t \in \left[\kappa, \frac{1}{\kappa}\right], \quad t \in [0, T]. \qquad (4.1)$$

Under (4.1), Assumptions 1 and 2 are trivially satisfied. Thus we can appeal to the results from the appendices for processes with such bounded coefficients and, after this localization, Theorem 1.2 is simply a corollary of Theorem 1.4 because $\mathcal{S}^p$ convergence implies uniform convergence in probability. In order to prove Theorem 1.2 for general coefficients, it therefore remains to establish Theorem 1.4 with bounded coefficients as in (4.1). To complete the proof of Theorem 1.4 for general coefficients, it additionally remains to establish Lemma A.3 to ensure the uniform integrability required for the localization argument from Section 2, and Lemma D.6 to ensure the integrability of the limit.

Furthermore, we assume in the following that the function $f$ is nondecreasing on $\mathbb{R}_+$. This is without loss of generality: since $f$ is of finite variation on every compact of $\mathbb{R}$ and even, it can be written as the difference of two non-decreasing and even functions, $f_1$ and $f_2$. Then, applying Theorem 1.2 and 1.4 to $f_1$ and $f_2$ yields the results for $f$. Finally, the result still holds after addition of a constant to $f$; we will therefore assume without loss of generality that $f$ is positive on $\mathbb{R}$.

## 4.1 Local estimation

We start by estimating the integral on the left-hand side of (1.4) "locally", i.e., on the intervals $[t, t + \varepsilon]$ for $t \in [0, T)$ and $\varepsilon > 0$ sufficiently small.[6] More precisely, we study the limit of the normalised integral

$$\varepsilon^{-1} \int_t^{t+\varepsilon} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds. \qquad (4.2)$$

To this end, we proceed in three steps. First, we rescale and time change the process $X^\varepsilon$ as in Section 3 and also use some stopping arguments to bound the integral (4.2) from above and from below by expressions only involving the rescaled and time-changed process $\widetilde{X}^\varepsilon$ from (3.3) and $\mathcal{F}_t$-measurable random variables. In a second step we approximate those $\mathcal{F}_t$-measurable random variables by elementary random variables. In a final step, we use ergodic theorems for one-dimensional diffusions to compute the "local limits" (4.2).

**Step 1: Stopping and time change.** In order to keep the "slowly-varying" processes $c$, $H$, $K$, $L$, and $M$ in a small interval around their values at time $t$, we define for fixed $\varepsilon \in (0, \frac{\kappa^2(T-t)}{4})$ and $\delta \in (0, \frac{\kappa}{2})$ the stopping time

$$\tau_t^{\varepsilon,\delta} = \inf\left\{s \in [t, t+\varepsilon] : \frac{H_s}{c_s} \notin \left[\frac{H_t(1-\delta)}{c_t}, \frac{H_t+\delta}{c_t}\right], c_s \notin [c_t(1-\delta), c_t(1+\delta)],\right.$$

$$\left. L_s \notin [L_t - \delta, L_t + \delta], M_s \notin [M_t - \delta, M_t + \delta], K_s \notin [K_t(1-\delta), K_t + \delta]\right\} \wedge (t+\varepsilon). \qquad (4.3)$$

---

[6]Note that the length of the interval could alternatively be taken equal to $\varepsilon^r$, for any $r \in (0, 2)$, in which case (4.2) is modified to $\varepsilon^{-r} \int_t^{t+\varepsilon^r} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds$. We choose $r = 1$ to obtain the simplest formulas.

By uniform continuity of $c, H, K, L, M$ on $[0, T]$, there exists a random variable $\varepsilon_\delta > 0$ (which is independent of $t$) such that, for $\varepsilon \in (0, \frac{\kappa^2(T-t)}{4})$, we have

$$\tau_t^{\varepsilon,\delta} = t + \varepsilon \quad \text{on } \{0 < \varepsilon \leqslant \varepsilon_\delta\}.$$

It follows that for $\varepsilon \in (0, \frac{\kappa^2(T-t)}{4})$,

$$\varepsilon^{-1} \int_t^{t+\varepsilon} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds = \varepsilon^{-1} \int_t^{\tau_t^{\varepsilon,\delta}} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds \quad \text{on } \{0 < \varepsilon \leqslant \varepsilon_\delta\}. \tag{4.4}$$

We proceed to study the integral on the right-hand side of (4.4). To this end, we pass to time-changed quantities as in Section 3, with the difference that we start time at $t$. So set

$$\xi_t^\varepsilon = \int_t^T \varepsilon^{-2} c_s ds, \quad \xi_t^{\varepsilon,\delta} = \int_t^{\tau_t^{\varepsilon,\delta}} \varepsilon^{-2} c_s ds, \quad u_\xi^{\varepsilon,t} = u_{\xi + \int_0^t \varepsilon^{-2} c_s ds}^\varepsilon, \quad \widetilde{X}_\xi^{\varepsilon,t} = X_{u_\xi^{\varepsilon,t}}^\varepsilon / \varepsilon.$$

Here, $\xi_t^\varepsilon$ and $\xi_t^{\varepsilon,\delta}$ denote the lengths of the intervals $[t, T]$ and $[t, \tau_t^{\varepsilon,\delta}]$ after the time change, $u_\xi^{\varepsilon,t}$ is the family of stopping times introduced in Lemma 3.1, shifted to start at the time change of $t$, and $\widetilde{X}^{\varepsilon,t}$ denotes the rescaled and time-changed process $\widetilde{X}^\varepsilon$ restarted at the time change of $t$. Note that for $\varepsilon \in (0, \frac{\kappa^2(T-t)}{4})$,

$$\varepsilon^{-1} c_t(1-\delta) \leqslant \xi_t^{\varepsilon,\delta} \leqslant \varepsilon^{-1} c_t(1+\delta) \quad \text{on } \{0 < \varepsilon \leqslant \varepsilon_\delta\}. \tag{4.5}$$

Moreover, note that since $c \in [\kappa, \kappa^{-1}]$ and $\delta \in (0, 1)$, we have $\xi_t^\varepsilon \geqslant \varepsilon^{-2} \kappa(T-t)$ and $2\varepsilon^{-1} c_t(1+\delta) \leqslant 4\varepsilon^{-1}/\kappa$. Together with $\kappa(T-t) \geqslant 4\kappa^{-1}\varepsilon$, this yields

$$2\varepsilon^{-1} c_t(1+\delta) \leqslant \xi_t^\varepsilon. \tag{4.6}$$

Hence, even though $\varepsilon^{-1} c_t(1+\delta)$ might be larger than $\xi_t^{\varepsilon,\delta}$ with positive probability, it is always smaller than the remaining time to the time horizon $\xi_t^\varepsilon$ after the time change. By considering $\widetilde{X}^{\varepsilon,t}$ on the interval $[0, \xi_t^{\varepsilon,\delta}]$, we can now separate the quickly-oscillating displacement from the other, more slowly-varying processes in the estimation:

**Lemma 4.1.** *Let $t \in [0, T)$, $\varepsilon \in (0, \frac{\kappa^2(T-t)}{4})$ and $\delta \in (0, \frac{\kappa}{2})$. Then:*

$$\mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}} (1-\delta)^2 H_t\left(\frac{1}{\varepsilon^{-1} c_t(1-\delta)} \int_0^{\varepsilon^{-1} c_t(1-\delta)} f\left(K_t(1-\delta)\widetilde{X}_\xi^{\varepsilon,t}\right) d\xi\right)$$

$$\leqslant \mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}} \varepsilon^{-1} \int_t^{t+\varepsilon} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr \tag{4.7}$$

$$\leqslant \mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}} (1+\delta)(H_t + \delta)\left(\frac{1}{\varepsilon^{-1} c_t(1+\delta)} \int_0^{\varepsilon^{-1} c_t(1+\delta)} f\left((K_t + \delta)\widetilde{X}_\xi^{\varepsilon,t}\right) d\xi\right).$$

*Proof.* On $\{0 < \varepsilon \leqslant \varepsilon_\delta\}$, (4.4) and the time change $s = u_\xi^{\varepsilon,t}$ give

$$\varepsilon^{-1} \int_t^{t+\varepsilon} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds = \varepsilon^{-1} \int_t^{\tau_t^{\varepsilon,\delta}} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) dr = \varepsilon \int_0^{\xi_t^{\varepsilon,\delta}} \frac{H_{u_\xi^{\varepsilon,t}}}{c_{u_\xi^{\varepsilon,t}}} f\left(K_{u_\xi^{\varepsilon,t}} \widetilde{X}_\xi^{\varepsilon,t}\right) d\xi.$$

Now, (4.7) follows by using that, by definition of $\tau_t^{\varepsilon,\delta}$,

$$\frac{H_{u_\xi^{\varepsilon,t}}}{c_{u_\xi^{\varepsilon,t}}} \in \left[(1-\delta)\frac{H_t}{c_t}, \frac{H_t + \delta}{c_t}\right], \quad \text{and} \quad K_{u_\xi^{\varepsilon,t}} \in [K_t(1-\delta), K_t + \delta], \quad \xi \in [0, \xi_t^{\varepsilon,\delta}],$$

and also taking into account (4.5), and that the function $f$ is even and non decreasing on $\mathbb{R}_+$. $\square$

**Step 2: Approximation by elementary random variables.** We now turn to the estimation of the terms that appear in the bounds from Lemma 4.1. To this end, we approximate $K_t$ and $c_t$ by elementary $\mathcal{F}_t$-measurable random variables and the SDE for $\widetilde{X}^{\varepsilon,t}$ by an SDE with coefficients that are constant over time and elementary $\mathcal{F}_t$-measurable random variables.

To this end, for $n \in \mathbb{N}$ with $n \geqslant \frac{2}{\kappa}$, $\Upsilon_t \in \{c_t, L_t, M_t, K_t\}$, and $i \in \mathbb{N}$ set

$$\Omega^{i,\Upsilon_t,n} := \left\{ \frac{i}{n} \leqslant \Upsilon_t < \frac{(i+1)}{n} \right\}, \tag{4.8}$$

and define the random variables $c_t^{n,+}, c_t^{n,-}, L_t^{n,+}, L_t^{n,-}, M_t^{n,+}, M_t^{n,-}, K_t^{n,+}, K_t^{n,-}$ for $n \in \mathbb{N}$ with $n \geqslant \frac{2}{\kappa}$ by

$$\Upsilon_t^{n,+} = \sum_{i=0}^{\infty} \frac{i+1}{n} \mathbb{1}_{\{\Omega^{i,\Upsilon_t,n}\}} \quad \text{and} \quad \Upsilon_t^{n,-} = \sum_{i=0}^{\infty} \frac{i}{n} \mathbb{1}_{\{\Omega^{i,\Upsilon_t,n}\}}, \tag{4.9}$$

where $\Upsilon \in \{c, L, M, K\}$. Note that, for fixed $n \geqslant \frac{2}{\kappa}$, and for $\delta \in \left(0, \frac{\kappa}{2}\right)$,

$$\mathbb{P}\left[\Omega^{i,\Upsilon_t,n}\right] = 0 \text{ for } \Upsilon_t \in \{c_t, L_t, M_t\} \text{ and } i \in \mathbb{N} \text{ with } \frac{i}{n} \leqslant \delta \tag{4.10}$$

because $c, L, M \geqslant \kappa$.[7] This implies that $c_t^{n,+}, c_t^{n,-}, L_t^{n,+}, L_t^{n,-}, M_t^{n,+}, M_t^{n,-} > \delta$. Moreover, note that for each $n \in \mathbb{N}$ with $n \geqslant \frac{2}{\kappa}$ by the fact that $c_t \geqslant \kappa$,

$$c_t^{n,+} \leqslant c_t + \frac{1}{n} \leqslant c_t + \frac{\kappa}{2} \leqslant 2c_t.$$

Together with (4.6), this yields the important estimate

$$\varepsilon^{-1} c_t^{n,+}(1+\delta) \leqslant \xi_t^{\varepsilon}. \tag{4.11}$$

By construction,

$$\Upsilon_t^{n,-} \leqslant \Upsilon_t \leqslant \Upsilon_t^{n,+}, \quad n \geqslant \frac{2}{\kappa}, \quad \text{and} \quad \lim_{n\to\infty} \Upsilon_t^{n,-} = \Upsilon_t = \lim_{n\to\infty} \Upsilon_t^{n,+}, \quad \Upsilon \in \{c, L, M, K\}, \tag{4.12}$$

and we have for $\Upsilon_t \in \{c_t, L_t, M_t\}$,

$$\bigcup_{i=1}^{\infty} \Omega^{i,\Upsilon_t,n} = \Omega \quad \text{and} \quad \bigcup_{i=0}^{\infty} \Omega^{i,K_t,n} = \Omega. \tag{4.13}$$

We proceed to approximate $\widetilde{X}^{\varepsilon,t}$ (or more precisely $(\widetilde{X}^{\varepsilon,t})^2$). It follows from (3.7) that the process $\widetilde{X}^{\varepsilon,t}$ satisfies on $[\![0, \xi_t^{\varepsilon}]\!]$ the SDE

$$d\widetilde{X}_{\xi}^{\varepsilon,t} = -L_{u_{\xi}^{\varepsilon,t}} g\left(M_{u_{\xi}^{\varepsilon,t}} \widetilde{X}_{\xi}^{\varepsilon,t}\right) d\xi + d\widetilde{W}_{\xi}^{\varepsilon,\mathbb{Q},t}, \quad \widetilde{X}_0^{\varepsilon,t} = X_t^{\varepsilon}/\varepsilon,$$

where $\widetilde{W}_{\xi}^{\varepsilon,\mathbb{Q},t} := \widetilde{W}_{\xi+\xi_0^{\varepsilon,t}}^{\varepsilon,\mathbb{Q}} - \widetilde{W}_{\xi_0^{\varepsilon,t}}^{\varepsilon,\mathbb{Q}}$ is the $\mathbb{Q}$-Brownian motion from (3.6) restarted at $\xi_0^{\varepsilon,t} = \int_0^t \varepsilon^{-2} c_s ds \leqslant \xi^{\varepsilon}$. Define the process $B^{\varepsilon,\mathbb{Q},t}$ by

$$B_{\xi}^{\varepsilon,\mathbb{Q},t} = \int_0^{\xi} \operatorname{sgn}(\widetilde{X}_y^{\varepsilon,t}) d\widetilde{W}_y^{\varepsilon,\mathbb{Q},t}, \quad \xi \geqslant 0.$$

---

[7]Note that $P[\Omega^{0,K_t,n}]$ may be positive as $K$ is only nonnegative.

By Lévy's characterisation [22, Theorem 3.3.16], this is a Brownian motion, stopped at $\xi_t^\varepsilon$. As the function $g$ is odd, Itô's formula gives

$$d\left(\widetilde{X}_\xi^{\varepsilon,t}\right)^2 = \left(1 - 2L_{u_\xi^{\varepsilon,t}}\sqrt{\left(\widetilde{X}_\xi^{\varepsilon,t}\right)^2}g\left(M_{u_\xi^{\varepsilon,t}}\sqrt{\left(\widetilde{X}_\xi^{\varepsilon,t}\right)^2}\right)\right)\mathbb{1}_{\{\xi\leqslant\xi_t^\varepsilon\}}d\xi + 2\sqrt{\left(\widetilde{X}_\xi^{\varepsilon,t}\right)^2}\mathbb{1}_{\{\xi\leqslant\xi_t^\varepsilon\}}dB_\xi^{\varepsilon,\mathbb{Q},t}.$$

To bound $(\widetilde{X}_\xi^{\varepsilon,t})^2$ from above and from below, we proceed as follows. For $\varepsilon > 0$, $\delta > 0$ and constants $l, m > \delta$ (which are independent of $\varepsilon$) and an $\mathcal{F}_t$-measurable initial value $y \geqslant 0$ (which may depend on $\varepsilon$), let $Y^{y,l,m,\varepsilon,\delta,+}$, and $Y^{y,l,m,\varepsilon,\delta,-}$ be the unique strong solutions of the following two SDEs:

$$dY_\xi^{y,l,m,\varepsilon,\delta,\pm} = \left(1 - 2\left(l \mp \delta\right)\sqrt{Y_\xi^{y,l,m,\varepsilon,\delta,\pm}}g\left((m \mp \delta)\sqrt{Y_\xi^{y,l,m,\varepsilon,\delta,\pm}}\right)\right)\mathbb{1}_{\{\xi\leqslant\xi_t^\varepsilon\}}d\xi$$
$$+ 2\sqrt{Y_\xi^{l,m,\varepsilon,\delta,\pm}}\mathbb{1}_{\{\xi\leqslant\xi_t^\varepsilon\}}dB_\xi^{\varepsilon,\mathbb{Q},t}, \quad Y_0^{y,l,m,\varepsilon,\delta,\pm} = y. \tag{4.14}$$

Existence and uniqueness of strong solutions for (4.14) follows from Lemma D.1.[8] Note that the SDEs (4.14) depend on $\varepsilon$ only via the Brownian motion $B^{\varepsilon,\mathbb{Q},t}$ and their starting value $y$.

Moreover, for an $\mathcal{F}_t$-measurable random variable $y \geqslant 0$ and each $n \in \mathbb{N} \setminus \{0\}$ define the continuous semimartingales $(Y_\xi^{y,L_t^{n,+},M_t^{n,+},\varepsilon,\delta,-})_{\xi\geq 0}$ and $(Y_\xi^{y,L_t^{n,-},M_t^{n,-},\varepsilon,\delta,+})_{\xi\geq 0}$ by

$$Y_\xi^{y,L_t^{n,+},M_t^{n,+},\varepsilon,\delta,-} = \sum_{i\geqslant 0}\sum_{j\geqslant 0}Y_\xi^{y,(i+1)/n,(j+1)/n,\varepsilon,\delta,-}\mathbb{1}_{\Omega^{i,L_t,n}}\mathbb{1}_{\Omega^{j,M_t,n}}, \tag{4.15}$$

$$Y_\xi^{y,L_t^{n,-},M_t^{n,-},\varepsilon,\delta,+} = \sum_{i\geqslant 0}\sum_{j\geqslant 0}Y_\xi^{y,i/n,j/n,\varepsilon,\delta,+}\mathbb{1}_{\Omega^{i,L_t,n}}\mathbb{1}_{\Omega^{j,M_t,n}}. \tag{4.16}$$

Now (4.12), the definition of $\tau_t^{\varepsilon,\delta}$ in (4.3), the assumption that $g$ is non-decreasing and the comparison theorem for SDEs in the form of Lemma C.1 give, for each $n \in \mathbb{N}$ with $n \geqslant \frac{2}{\kappa}$ and all $i, j \in \mathbb{N}$,

$$\mathbb{P}\Bigg[Y_\xi^{Y_0^\varepsilon,(i+1)/n,(j+1)/n,\varepsilon,\delta,-}\mathbb{1}_{\Omega^{i,L_t,n}}\mathbb{1}_{\Omega^{j,M_t,n}} \leqslant \left(\widetilde{X}_\xi^{\varepsilon,t}\right)^2\mathbb{1}_{\Omega^{i,L_t,n}}\mathbb{1}_{\Omega^{j,M_t,n}}$$
$$\leqslant Y_\xi^{Y_0^\varepsilon,i/n,j/n,\varepsilon,\delta,+}\mathbb{1}_{\Omega^{i,L_t,n}}\mathbb{1}_{\Omega^{j,M_t,n}} \text{ for all } 0 \leqslant \xi \leqslant \xi_t^{\varepsilon,\delta}\Bigg] = 1,$$

where

$$Y_0^\varepsilon := \left(\widetilde{X}_0^{\varepsilon,t}\right)^2 = (X_t^\varepsilon)^2/\varepsilon^2.$$

Together with (4.15) and (4.16), (4.8), (4.10) and (4.13), this yields

$$\mathbb{P}\left[Y_\xi^{Y_0^\varepsilon,L_t^{n,+},M_t^{n,+},\varepsilon,\delta,-} \leqslant \left(\widetilde{X}_\xi^{\varepsilon,t}\right)^2 \leqslant Y_\xi^{Y_0^\varepsilon,L_t^{n,-},M_t^{n,-},\varepsilon,\delta,+} \text{ for all } 0 \leqslant \xi \leqslant \xi_t^{\varepsilon,\delta}\right] = 1. \tag{4.17}$$

To simplify the notation in the subsequent results, define for constants $c, k, l, m$ (independent of $\varepsilon$) with $c, l, m > \delta$ and an $\mathcal{F}_t$-valued random variable $y$ (which may depend on $\varepsilon$) the following two random variables:

$$v_y^{\varepsilon,\delta,+}(c,k,l,m) := \frac{1}{\varepsilon^{-1}c(1+\delta)}\int_0^{\varepsilon^{-1}c(1+\delta)}f\left((k+\delta)\sqrt{Y_\xi^{y,l,m,\varepsilon,\delta,+}}\right)d\xi,$$

$$v_y^{\varepsilon,\delta,-}(c,k,l,m) := \frac{1}{\varepsilon^{-1}c(1-\delta)}\int_0^{\varepsilon^{-1}c(1-\delta)}f\left(k(1-\delta)\sqrt{Y_\xi^{y,l,m,\varepsilon,\delta,-}}\right)d\xi.$$

---

[8]More precisely, the solution of (4.14) corresponds to a solution of (D.1), stopped at time $\xi_t^\varepsilon$.

Note that by (4.6), $v_y^{\varepsilon,\delta,+}(c,k,l,m)$ and $v_y^{\varepsilon,\delta,-}(c,k,l,m)$ are $\mathcal{F}_T$-measurable for $c \leqslant 1/\kappa$. Moreover, by comparison of SDEs in their initial values, they are non decreasing in $y$ (see Lemma C.1 and the proof of Lemma D.1).

Now combining (4.17) and (4.12) with Lemma 4.1 and the fact that $f$ is even and nondecreasing on $\mathbb{R}_+$ yields the following result.[9]

**Lemma 4.2.** *Let* $t \in [0,T)$, $\varepsilon \in (0, \frac{\kappa^2(T-t)}{4})$, $\delta \in (0, \frac{\kappa}{2})$, *and* $n \in \mathbb{N} \setminus \{0\}$. *Then:*

$$
\mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}} (1-\delta)^2 \, H_t \frac{c_t^{n,-}}{c_t^{n,+}} v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}(c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+})
$$

$$
\leqslant \mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}} \varepsilon^{-1} \int_t^{t+\varepsilon} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr \tag{4.18}
$$

$$
\leqslant \mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}} (1+\delta)(H_t + \delta) \frac{c_t^{n,+}}{c_t^{n,-}} v_{Y_0^\varepsilon}^{\varepsilon,\delta,+}(c_t^{n,+}, K_t^{n,+}, L_t^{n,-}, M_t^{n,-}),
$$

*where* $c^{n,+}$, $c^{n,-}$, $K^{n,+}$, $K^{n,-}$, $L^{n,+}$, $L^{n,-}$, $M^{n,+}$, $M^{n,-}$ *are defined as in* (4.9).

**Step 3: Limit theorems.** We now combine the sandwiching inequalities from Lemma 4.2 with an ergodic theorem for one-dimensional diffusions (cf. Lemma D.4) to calculate the following "local" scaling limit:

$$
\lim_{\varepsilon \to 0} \varepsilon^{-1} \int_t^{t+\varepsilon} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr, \quad t \in [0,T).
$$

To this end, we first establish an ergodic result, which is non-standard in that both the time horizon and the underlying process change with the small parameter at hand. As a consequence, the ergodic limit only holds in probability here rather than almost surely. To formulate this result, define for constants $k \geqslant 0$ and $l, m > \delta > 0$, the two continuous functions

$$
w^{\delta,+}(k,l,m) := \frac{\int_{\mathbb{R}_+} f\left(\frac{(k+\delta)}{m-\delta} x\right) \exp\left(-2\frac{l-\delta}{m-\delta} G(x)\right) dx}{\int_{\mathbb{R}_+} \exp\left(-2\frac{l-\delta}{m-\delta} G(x)\right) dx},
$$

$$
w^{\delta,-}(k,l,m) := \frac{\int_{\mathbb{R}_+} f\left(\frac{k(1-\delta)}{m+\delta} x\right) \exp\left(-2\frac{l+\delta}{m+\delta} G(x)\right) dx}{\int_{\mathbb{R}_+} \exp\left(-2\frac{l+\delta}{m+\delta} G(x)\right) dx}.
$$

With this notation, our ergodic result reads as follows:

**Lemma 4.3.** *Let* $t \in [0,T)$, $\delta \in (0, \frac{\kappa}{2})$, *and* $n \in \mathbb{N}$ *with* $n \geqslant \frac{2}{\kappa}$ *be fixed. Then the following two limits hold in probability:*

$$
\lim_{\varepsilon \to 0} v_{Y_0^\varepsilon}^{\varepsilon,\delta,\pm}(c_t^{n,\pm}, K_t^{n,\pm}, L_t^{n,\mp}, M_t^{n,\mp}) = w^{\delta,\pm}(K_t^{n,\pm}, L_t^{n,\mp}, M_t^{n,\mp}). \tag{4.19}
$$

*Proof.* We only spell out the argument for the "$-$"-limit in (4.19); the "$+$"-limit is established analogously. By one of the equivalent characterizations of convergence in probability, we have to show that

$$
\lim_{\varepsilon \to 0} \mathbb{E}\left[\left| v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}(c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) - w^{\delta,-}(K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) \right| \wedge 1 \right] = 0.
$$

---

[9]Note that compared to Lemma 4.1, the processes in the upper and lower bounds are replaced by the simpler approximating diffusions introduced in (4.14) here and the frozen coefficients are approximated by finitely many values, as $c, K, L$ and $M$ are bounded from above.

Since both $v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}(c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+})$ and $w^{\delta,-}(K_t^{n,-}, L_t^{n,+}, M_t^{n,+})$ are $\mathcal{F}_T$-measurable and the measure $\mathbb{Q}$ defined in (3.5) is equivalent to $\mathbb{P}$ on $\mathcal{F}_T$, it suffices to show that

$$\lim_{\varepsilon \to 0} \mathbb{E}_{\mathbb{Q}} \left[ \left| v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}(c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) - w^{\delta,-}(K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) \right| \wedge 1 \right] = 0.$$

For $i_c, i_K, i_L, i_M \in \mathbb{N}$ set

$$\Omega_{i_c, i_K, i_L, i_M} := \Omega^{i_c, c_t, n} \cap \Omega^{i_K, K_t, n} \cap \Omega^{i_L, L_t, n} \cap \Omega^{i_M, M_t, n},$$

where, for $\Upsilon_t \in \{c_t, K_t, L_t, M_t\}$ and $i \in \{i_c, i_K, i_L, i_M\}$, the set $\Omega^{i, \Upsilon_t, n}$ is defined as in (4.8). By dominated convergence and (4.8) it suffices to show that

$$\lim_{\varepsilon \to 0} \mathbb{E}_{\mathbb{Q}} \left[ \left( \left| v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}(c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) - w^{\delta,-}(K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) \right| \wedge 1 \right) \mathbb{1}_{\Omega_{i_c, i_K, i_L, i_M}} \right] = 0,$$

for all $i_c, i_K, i_L, i_M \in \mathbb{N}$. So fix $i_c, i_K, i_L, i_M \in \mathbb{N}$ with $\mathbb{Q}[\Omega_{i_c, i_K, i_L, i_M}] > 0$. Note that by (4.10) and $\mathbb{P} \approx \mathbb{Q}$, this implies in particular that $i_c/n, i_L/n, i_M/n > \delta$. Using that $c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+}$ take the constant values $i_c/n, i_K/n, (i_L + 1)/n, (i_M + 1)/n$ on $\Omega_{i_c, i_K, i_L, i_M}$ we have, for each fixed $\varepsilon > 0$,

$$\mathbb{E}_{\mathbb{Q}} \left[ \left( \left| v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}(c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) - w^{\delta,-}(K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) \right| \wedge 1 \right) \mathbb{1}_{\Omega_{i_c, i_K, i_L, i_M}} \right]$$
$$= \mathbb{E}_{\mathbb{Q}} \left[ \left( \left| v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}\left( \frac{i_c}{n}, \frac{i_K}{n}, \frac{i_L + 1}{n}, \frac{i_M + 1}{n} \right) - w^{\delta,-}\left( \frac{i_c}{n}, \frac{i_K}{n}, \frac{i_L + 1}{n}, \frac{i_M + 1}{n} \right) \right| \wedge 1 \right) \mathbb{1}_{\Omega_{i_c, i_K, i_L, i_M}} \right].$$
(4.20)

We proceed to estimate the right-hand side of (4.20). To simplify notation, set

$$\bar{w} := w^{\delta,-}\left( \frac{i_c}{n}, \frac{i_K}{n}, \frac{i_L + 1}{n}, \frac{i_M + 1}{n} \right) \tag{4.21}$$

and note that this is a constant. For $\lambda > 0$, we split up the expectation to the disjoint events $\{0 \leqslant Y_0^\varepsilon \leqslant \frac{1}{\lambda}\}$ and $\{Y_0^\varepsilon > \frac{1}{\lambda}\}$. On the event $\{0 \leqslant Y_0^\varepsilon \leqslant \frac{1}{\lambda}\}$, we use that $v_y^{\varepsilon,\delta,-}(c, k, l, m)$ is non decreasing in $y$ together with the elementary inequality $|z - \bar{w}| \leqslant |z_{\min} - \bar{w}| + |z_{\max} - \bar{w}|$ for $\bar{w} \in \mathbb{R}$ and $z_{\min} \leqslant z \leqslant z_{\max} \in \mathbb{R}$. On the event $\{Y_0^\varepsilon > \frac{1}{\lambda}\}$, we use that the random variable inside the expectation on the right-hand side of (4.20) is bounded from above by 1. Together, this yields

$$\mathbb{E}_{\mathbb{Q}} \left[ \left| v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}\left( \frac{i_c}{n}, \frac{i_K}{n}, \frac{i_L + 1}{n}, \frac{i_M + 1}{n} \right) - \bar{w} \right| \wedge 1 \right]$$
$$\leqslant \mathbb{E}_{\mathbb{Q}} \left[ \left| v_0^{\varepsilon,\delta,-}\left( \frac{i_c}{n}, \frac{i_K}{n}, \frac{i_L + 1}{n}, \frac{i_M + 1}{n} \right) - \bar{w} \right| \wedge 1 \right]$$
$$+ \mathbb{E}_{\mathbb{Q}} \left[ \left| v_{\frac{1}{\lambda}}^{\varepsilon,\delta,-}\left( \frac{i_c}{n}, \frac{i_K}{n}, \frac{i_L + 1}{n}, \frac{i_M + 1}{n} \right) - \bar{w} \right| \wedge 1 \right] + \mathbb{Q}\left[ Y_0^\varepsilon > \frac{1}{\lambda} \right]. \tag{4.22}$$

Next, note that for $y \in \{0, \frac{1}{\lambda}\}$ independent of $\varepsilon$, the random variables $v_y^{\varepsilon,\delta,-}(i_c/n, i_K/n, (i_L + 1)/n, (i_M + 1)/n)$ depend on $\varepsilon$ only via the Brownian motion $B^{\mathbb{Q},\varepsilon,t}$ which also is the only source of stochasticity. In particular, the law of $v_y^{\varepsilon,\delta,-}(i_c/n, i_K/n, (i_L + 1)/n, (i_M + 1)/n)$ does not depend on $\varepsilon$. Thus, if we replace the $\mathbb{Q}$-Brownian motion $B^{\mathbb{Q},\varepsilon,t}$ by any other fixed Brownian motion $B$ (on some different probability space), the result does not change. Hence, we can apply the ergodic theorem for one-dimensional diffusions in the form of Lemma D.4 to conclude that the first two

15

terms on the right-hand side of (4.22) converge to zero as $\varepsilon \to 0$.[10] Finally, by Markov's inequality and Lemma A.1,[11] there is a constant $\overline{C}_2$ independent of $\varepsilon$ such that

$$\mathbb{Q}\left[Y_0^\varepsilon > \frac{1}{\lambda}\right] = \mathbb{Q}\left[\left(\widetilde{X}_0^{\varepsilon,t}\right)^2 > \frac{1}{\lambda}\right] = \mathbb{Q}\left[\left(\frac{X_t^\varepsilon}{\varepsilon}\right)^2 > \frac{1}{\lambda}\right] \leqslant \lambda \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{X_t^\varepsilon}{\varepsilon}\right)^2\right] \leqslant \lambda \overline{C}_2. \qquad (4.23)$$

The claim in turn follows by letting $\lambda$ go to zero. $\qquad\square$

Sending the localization parameter $\delta$ from Lemma 4.1 to zero and the discretization parameter $n$ from (4.8) to infinity, we now obtain the following scaling limit:

**Proposition 4.4.** *For $t \in [0,T)$, the following limit holds in probability:*

$$\lim_{\varepsilon \to 0} \varepsilon^{-1} \int_t^{t+\varepsilon} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr = H_t \frac{\int_\mathbb{R} f\left(\frac{K_t}{M_t}x\right) \exp\left(-2\frac{L_t}{M_t}G\left(x\right)\right) dx}{\int_\mathbb{R} \exp\left(-2\frac{L_t}{M_t}G\left(x\right)\right) dx}.$$

*Proof.* Fix $t \in [0,T)$. For $\varepsilon \in (0, \frac{\kappa^2(T-t)}{2})$, $\delta \in (0, \frac{\kappa}{2})$, and $n \in \mathbb{N}$ with $n \geqslant \frac{2}{\kappa}$, set

$$a_t^{\varepsilon,\delta,n,-} := \mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}}(1-\delta)^2 H_t \frac{c_t^{n,-}}{c_t^{n,+}} v_{Y_0^\varepsilon}^{\varepsilon,\delta,-}(c_t^{n,-}, K_t^{n,-}, L_t^{n,+}, M_t^{n,+}),$$

$$a_t^{\varepsilon,\delta,n,+} := \mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}}(1+\delta)(H_t+\delta) \frac{c_t^{n,+}}{c_t^{n,-}} v_{Y_0^\varepsilon}^{\varepsilon,\delta,+}(c_t^{n,+}, K_t^{n,+}, L_t^{n,-}, M_t^{n,-}),$$

$$a_t^{1,\varepsilon,\delta} := \mathbb{1}_{\{\varepsilon \leqslant \varepsilon_\delta\}} \varepsilon^{-1} \int_t^{t+\varepsilon} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds,$$

$$a_t^{1,\varepsilon} := \varepsilon^{-1} \int_t^{t+\varepsilon} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds,$$

$$a_t^2 := H_t w(K_t, L_t, M_t),$$

where, for constants $k \geqslant 0$ and $l, m > 0$,

$$w(k,l,m) := \frac{\int_\mathbb{R} f\left(\frac{k}{m}x\right) \exp\left(-2\frac{l}{m}G\left(x\right)\right) dx}{\int_\mathbb{R} \exp\left(-2\frac{l}{m}G\left(x\right)\right) dx}.$$

Note that the choices for $\varepsilon$ and $\kappa \in (0,1)$ ensure that $t + \varepsilon \leqslant T$. Furthermore, by Lemma 4.2, we have

$$a_t^{\varepsilon,\delta,n,-} \leqslant a_t^{1,\varepsilon,\delta} \leqslant a_t^{\varepsilon,\delta,n,+}. \qquad (4.24)$$

We want to show that, in probability, $\lim_{\varepsilon \to 0} a_t^{1,\varepsilon} = a_t^2$. To this end, we use the subsequence criterion for convergence in probability, cf. [21, Lemma 4.2]. Let $(\varepsilon_m)_{m \in \mathbb{N}}$ be a sequence of positive real numbers converging to zero. By Lemma 4.3 and the subsequence criterion, there exists a subsequence $(m_k)_{k \in \mathbb{N}}$ of $\mathbb{N}$ such that

$$\lim_{k \to \infty} v_{Y_0^{\varepsilon_{m_k}}}^{\varepsilon_{m_k},\delta,\pm}(c_t^{n,\pm}, K_t^{n,\pm}, L_t^{n,\mp}, M_t^{n,\mp}) = w^{\delta,\pm}(K_t^{n,\pm}, L_t^{n,\mp}, M_t^{n,\mp}) \text{ a.s.}$$

---

[10]This also uses crucially the estimate (4.11) and the fact that $i_c/n \geqslant \kappa/2$.

[11]Note that all assumptions in Lemma A.1 are satisfied by (4.1).

Hence,

$$\lim_{k\to\infty} a_t^{\varepsilon_{m_k},\delta,n,-} = (1-\delta)^2 \, H_t \frac{c_t^{n,-}}{c_t^{n,+}} w^{\delta,-}(K_t^{n,-}, L_t^{n,+}, M_t^{n,+}) \text{ a.s.,}$$

$$\lim_{k\to\infty} a_t^{\varepsilon_{m_k},\delta,n,+} = (1+\delta)\,(H_t+\delta)\frac{c_t^{n,+}}{c_t^{n,-}} w^{\delta,+}(K_t^{n,+}, L_t^{n,-}, M_t^{n,-}) \text{ a.s.}$$

Moreover, (4.24) and $\varepsilon_\delta > 0$ give

$$\lim_{k\to\infty} a_t^{\varepsilon_{m_k},\delta,n,-} \leqslant \liminf_{k\to\infty} a_t^{1,\varepsilon_{m_k},\delta} = \liminf_{k\to\infty} a_t^{1,\varepsilon_{m_k}} \leqslant \limsup_{k\to\infty} a_t^{1,\varepsilon_{m_k}} = \limsup_{k\to\infty} a_t^{1,\varepsilon_{m_k},\delta}$$

$$\leqslant \lim_{k\to\infty} a_t^{\varepsilon_{m_k},\delta,n,+} \quad \text{a.s.} \tag{4.25}$$

Finally, using that $\Upsilon_t^{n,\pm}$ (defined in (4.9)) converges almost surely to $\Upsilon_t$ as $n \to \infty$ for $\Upsilon \in \{K, L, M\}$, we obtain by dominated convergence (and continuity of $f$) that

$$\lim_{\delta\to 0} \lim_{n\to\infty} \lim_{k\to\infty} a_t^{\varepsilon_{m_k},\delta,n,-} = H_t w(K_t, L_t, M_t) = \lim_{\delta\to 0} \lim_{n\to\infty} \lim_{k\to\infty} a_t^{\varepsilon_{m_k},\delta,n,+}.$$

Together with (4.25), this shows that

$$\liminf_{k\to\infty} a_t^{1,\varepsilon_{m_k}} = \limsup_{k\to\infty} a_t^{1,\varepsilon_{m_k}} = H_t w(K_t, L_t, M_t) = a_t^2 \text{ a.s.}$$

The assertion in turn follows from the subsequence criterion. $\qquad\square$

## 4.2 Concatenation of the local estimates

We now piece together the local estimates from Proposition 4.4 to establish Theorem 1.4. For each $\varepsilon \in (0, T)$, define as above the product-measurable processes $(a_t^{1,\varepsilon})_{t\in[0,T]}$ and $(a_t^2)_{t\in[0,T]}$ by[12]

$$a_t^{1,\varepsilon} = \mathbb{1}_{\{0\leqslant t\leqslant T-\varepsilon\}}\varepsilon^{-1} \int_t^{t+\varepsilon} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds, \quad a_t^2 = H_t w(K_t, L_t, M_t).$$

It follows from Fubini's theorem that, for each $t \in [0, T]$:

$$\int_0^t a_s^{1,\varepsilon} ds = \int_0^T \int_0^T \mathbb{1}_{\{0\leqslant s\leqslant t\wedge(T-\varepsilon)\}}\mathbb{1}_{\{s\leqslant r\leqslant s+\varepsilon\}}\varepsilon^{-1} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr ds$$

$$= \int_0^T \int_0^T \mathbb{1}_{\{r-\varepsilon\vee 0\leqslant s\leqslant r\wedge t\wedge(T-\varepsilon)\}} ds \, \mathbb{1}_{\{0\leqslant r\leqslant (t+\varepsilon)\wedge T\}}\varepsilon^{-1} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr$$

$$= \int_0^T \int_0^T \left(\mathbb{1}_{\{r-\varepsilon\vee 0\leqslant s\leqslant r\}} - \mathbb{1}_{\{t\wedge(T-\varepsilon)<s\leqslant r\}}\right) ds \, \mathbb{1}_{\{0\leqslant r\leqslant (t+\varepsilon)\wedge T\}}\varepsilon^{-1} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr$$

$$= \int_0^\varepsilon \varepsilon^{-1} r H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr + \int_\varepsilon^{(t+\varepsilon)\wedge T} H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr$$

$$- \int_{t\wedge(T-\varepsilon)}^{(t+\varepsilon)\wedge T} \varepsilon^{-1}\left(r - \left(t\wedge(T-\varepsilon)\right)\right) H_r f\left(\frac{K_r X_r^\varepsilon}{\varepsilon}\right) dr \tag{4.26}$$

$$\leqslant \int_0^T H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds. \tag{4.27}$$

---

[12]The indicator of the set $\{0 \leqslant t \leqslant T - \varepsilon\}$ is needed now as we define $a_t^{1,\varepsilon}$ on $[0, T]$ for fixed $\varepsilon > 0$.

Rearranging and recalling the growth condition (1.3) for $f$ and the fact that $K$ and $L$ are uniformly bounded from above by $1/\kappa$ yields

$$\sup_{t\in[0,T]} \left| \int_0^t a_s^{1,\varepsilon} ds - \int_0^t H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds \right| \leqslant \int_0^\varepsilon H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds$$

$$+ \sup_{t\in[0,T]} \left( \int_t^{(t+\varepsilon)\wedge T} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds \right)$$

$$+ \sup_{t\in[0,T-\varepsilon]} \left( \int_t^{t+\varepsilon} H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds \right)$$

$$\leqslant 3 \sup_{t\in[0,T]} \left( \int_t^{(t+\varepsilon)\wedge T} \frac{1}{\kappa} C_f \left( \left( \left| \frac{X_s^\varepsilon}{\varepsilon} \right| \frac{1}{\kappa} \right)^{q'} + 1 \right) ds \right)$$

$$\leqslant 3 C_f \left( \frac{1}{\kappa} \right)^{1+q'} \varepsilon \left( \sup_{t\in[0,T]} \left| \frac{X_t^\varepsilon}{\varepsilon} \right|^{q'} + \kappa^{q'} \right).$$

Thus, it follows from Lemma B.4 that

$$\lim_{\varepsilon\to 0} \mathbb{E}\left[ \sup_{t\in[0,T]} \left| \int_0^t a_s^{1,\varepsilon} ds - \int_0^t H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right) ds \right|^p \right] = 0. \tag{4.28}$$

Now, by Proposition 4.4, for each $t \in [0,T)$, we have

$$\mathbb{P}\text{-}\lim_{\varepsilon\to 0} a_t^{1,\varepsilon} = a_t^2.$$

Next, recall that Assumptions 1 and 2 are satisfied due to the uniform boundedness assumption (4.1). Therefore, Lemma D.6 ensures that $a^2$ is in $L^p(\mathbb{P} \otimes \mathrm{Leb}_{|[0,T]})$. Moreover, Jensen's inequality and computations similar to the ones leading to (4.27) give

$$\int_0^T |a_t^{1,\varepsilon}|^{p(1+\eta)} dt \leqslant \int_0^T H_s^{p(1+\eta)} f\left(\frac{K_t X_t^\varepsilon}{\varepsilon}\right)^{p(1+\eta)} dt,$$

which, with Lemma A.3 yields

$$\sup_{\varepsilon>0} \mathbb{E}\left[ \int_0^T |a_t^{1,\varepsilon}|^{p(1+\eta)} dt \right] < \infty. \tag{4.29}$$

By de la Vallée-Poussin's criterion for uniform integrability (cf., e.g., the remark before [21, Lemma 4.10]), (4.29) implies that $(a^{1,\varepsilon})^p$ is uniformly integrable with respect to $\mathbb{P} \otimes \mathrm{Leb}_{|[0,T]}$.

Hence, it follows from Lemma D.5 that

$$\lim_{\varepsilon\to 0} \mathbb{E}\left[ \sup_{t\in[0,T]} \left| \int_0^t a_s^{1,\varepsilon} ds - \int_0^t a_s^2 ds \right|^p \right] = 0. \tag{4.30}$$

Theorem 1.4 now follows by putting together (4.30) and (4.28). □

**Remark 4.5.** To prove the variant of the results mentioned in Remark 1.5, the argument in Section 4.2 needs to be changed at two places. The first is the use of moments of $X^\varepsilon/\varepsilon$ to obtain (4.23). This requires a version of Lemma A.1 valid under the alternative assumptions of Remark 1.5. To

this end, we argue as in Veretennikov [36, Lemmas 1-5]: Fix $\varepsilon > 0$ and compare the SDE satisfied by $\overline{X}^\varepsilon_\cdot = X^\varepsilon_{\varepsilon^2\cdot}$ under $\mathbb{Q}$,

$$d\overline{X}^\varepsilon_s = -\sigma^2 L_{\varepsilon^2 s} g\left(M_{\varepsilon^2 s}\overline{X}^\varepsilon_s\right)ds + \sigma d\overline{W}^{\varepsilon,\mathbb{Q}}_s, \qquad \overline{X}^\varepsilon_0 = x^\varepsilon_0/\varepsilon,$$

where $\overline{W}^{\varepsilon,\mathbb{Q}}$ is some $\mathbb{Q}$-Brownian motion, to the SDE satisfied by the process $\overline{X}^{V,\varepsilon}$ under $\mathbb{Q}$ given by[13]

$$d\overline{X}^{V,\varepsilon}_s = -\frac{\sigma^2 C}{2}\frac{1}{\overline{X}^{V,\varepsilon}_s}ds + \sigma d\overline{W}^{\varepsilon,\mathbb{Q}}_s + dA_s, \qquad \overline{X}^{V,\varepsilon}_0 = \frac{x^\varepsilon_0}{\varepsilon}\vee 2, \tag{4.31}$$

where $A$ is a non-decreasing process increasing on $\{\overline{X}^{V,\varepsilon} = 1\}$, starting at 0 and such that

$$A_t = \int_0^t \mathbb{1}_{\{\overline{X}^{V,\varepsilon}_s = 1\}}dA_s, \qquad \text{and} \qquad \mathbb{E}_\mathbb{Q}\left[\int_0^{\varepsilon^{-2}T}\mathbb{1}_{\{\overline{X}^{V,\varepsilon}_s = 1\}}ds\right] = 0.$$

The process $\overline{X}^{V,\varepsilon}$ has a non-sticky reflecting boundary at 1. In view of (1.5), [36, Lemma 2] then shows that

$$\mathbb{Q}\left[|\overline{X}^\varepsilon_t| \leqslant |\overline{X}^{V,\varepsilon}_t|, \ \forall\, t \in [0, \varepsilon^{-2}T]\right] = 1. \tag{4.32}$$

The existence of all moments for $X^\varepsilon/\varepsilon$ then follows from the existence of all moments for $\overline{X}^{V,\varepsilon}$, cf. [36, Lemmas 3-5] and the assumption $\limsup_{\varepsilon \to 0} |x^\varepsilon_0|/\varepsilon < \infty$ made on the initial condition.

The second adaptation is necessary in the arguments after (4.27). Fix $0 < \varepsilon \leqslant 1$. Starting from (4.26), using that $f$ and $H$ are nonnegative and $f$ is non-decreasing, the comparison result (4.32) (which remains true under $\mathbb{P}$), the elementary inequality $(a+b)^p \leqslant 2^p(a^p + b^p)$ for $a, b \geqslant 0$, Jensen's inequality (using also that $p \geq 1$ and $\varepsilon \leqslant 1$), the growth condition (1.3) for $f$, Fubini's theorem and the moments assumptions on $H$ and $K$, we obtain for fixed $t \in [0, T]$

$$\mathbb{E}\left[\left|\int_0^t a^{1,\varepsilon}_s ds - \int_0^t H_s f\left(\frac{K_s X^\varepsilon_s}{\varepsilon}\right)ds\right|^p\right]$$

$$\leqslant \mathbb{E}\left[\left|\int_0^\varepsilon H_s f\left(K_s \overline{X}^\varepsilon_{\varepsilon^{-2}s}\right)ds + 2\int_{t\wedge(T-\varepsilon)}^{(t+\varepsilon)\wedge T} H_s f\left(K_s \overline{X}^\varepsilon_{\varepsilon^{-2}s}\right)ds\right|^p\right]$$

$$\leqslant \mathbb{E}\left[\left|\int_0^\varepsilon H_s f\left(K_s \overline{X}^{V,\varepsilon}_{\varepsilon^{-2}s}\right)ds + 2\int_{t\wedge(T-\varepsilon)}^{(t+\varepsilon)\wedge T} H_s f\left(K_s \overline{X}^{V,\varepsilon}_{\varepsilon^{-2}s}\right)ds\right|^p\right]$$

$$\leqslant 2^p \mathbb{E}\left[\int_0^\varepsilon H^p_s f\left(K_s \overline{X}^{V,\varepsilon}_{\varepsilon^{-2}s}\right)^p ds\right] + 4^p \mathbb{E}\left[\int_{t\wedge(T-\varepsilon)}^{(t+\varepsilon)\wedge T} H^p_s f\left(K_s \overline{X}^{V,\varepsilon}_{\varepsilon^{-2}s}\right)^p ds\right]$$

$$\leqslant 4^p C^p_f \int_0^\varepsilon \mathbb{E}\left[H^p_s\left(K^{pq'}_s\left|\overline{X}^{V,\varepsilon}_{\varepsilon^{-2}s}\right|^{pq'} + 1\right)\right]ds$$

$$+ 8^p C^p_f \int_{t\wedge(T-\varepsilon)}^{(t+\varepsilon)\wedge T} \mathbb{E}\left[H^p_s\left(K^{pq'}_s\left|\overline{X}^{V,\varepsilon}_{\varepsilon^{-2}s}\right|^{pq'} + 1\right)\right]ds. \tag{4.33}$$

Next, apply the Cauchy–Schwarz inequality twice and use the integrability Assumption 1 and [36, Lemma 5]. This together with the elementary inequality $x^{1/2} \leq x + 1$ for $x \geq 0$ gives for some

---

[13]This process is denoted by $v$ in [36, Lemma 2]. Note, moreover, that $\overline{X}^{V,\varepsilon}$ depends on $\varepsilon$ only through the Brownian motion $\overline{W}^{\varepsilon,\mathbb{Q}}$.

constant $C$ independent of $s$,

$$\mathbb{E}\left[H_s^p\left(K_s^{pq'}\left|\overline{X}_{\varepsilon^{-2}s}^{V,\varepsilon}\right|^{pq'}+1\right)\right] \leqslant \mathbb{E}\left[H_s^{2p}K_s^{2pq'}\right]^{\frac{1}{2}}\mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\left|X_{\varepsilon^{-2}s}^{V,\varepsilon}\right|^{2pq'}\right]^{\frac{1}{2}}+\mathbb{E}\left[H_s^p\right]$$

$$\leqslant \mathbb{E}\left[H_s^{2p}K_s^{2pq'}\right]^{\frac{1}{2}}\mathbb{E}\left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)\right]^{\frac{1}{4}}\mathbb{E}_{\mathbb{Q}}\left[\left|X_{\varepsilon^{-2}s}^{V,\varepsilon}\right|^{4pq'}\right]^{\frac{1}{4}}+\mathbb{E}\left[H_s^p\right]$$

$$\leqslant C\mathbb{E}\left[H_s^{2p}K_s^{2pq'}\right]^{\frac{1}{2}}+\mathbb{E}\left[H_s^p\right]\leqslant C\left(\mathbb{E}\left[H_s^{2p}K_s^{2pq'}\right]+1\right)+\mathbb{E}\left[H_s^p\right].$$

Plugging this into (4.33), using Fubini's theorem and Assumption 2 finally gives

$$\mathbb{E}\left[\left|\int_0^t a_s^{1,\varepsilon}ds - \int_0^t H_s f\left(\frac{K_s X_s^\varepsilon}{\varepsilon}\right)ds\right|^p\right] \to 0 \;\; \text{as} \;\; \varepsilon \to 0.$$

This establishes convergence in $L^p$ (and a fortiori in probability) for all $t \in [0,T]$.[14]

## A    Integrability Results

In this appendix, we establish moment estimates that are needed at various stages of the proof of Theorem 1.4. Most importantly, we show in Lemma A.3 that the $p$-th moment of the expression on the left-hand side of (1.4) is uniformly integrable. This is crucial both for the reduction to bounded coefficients in Section 2 and the concatenation argument in Section 4.2. The main ingredient to this result is to establish that $X^\varepsilon/\varepsilon$ has uniformly bounded moments. This is done in Lemma A.1.

### A.1    A First Moment Estimate

The following result is the key ingredient for Lemma A.3. It is also used in the proof of Lemma 4.3.

**Lemma A.1.** *Let $N \geqslant 2$ be a real number. Suppose that Assumption 1 is satisfied and*

$$\mathbb{E}\left[\sup_{u \in [0,T]} c_u^{2N}\right] + \mathbb{E}\left[\sup_{u \in [0,T]}(L_u c_u \wedge M_u)^{-\frac{2(q+1)N}{q-1}}\right] < \infty \qquad \text{if } q > 1,$$

*or*

$$\mathbb{E}\left[\sup_{u \in [0,T]} c_u^{2N}\right] < \infty \quad and \quad \text{essinf}_{u \in [0,T]}(L_u c_u \wedge M_u) > 0 \qquad \text{if } q = 1.$$

*Then for each $0 \leqslant n \leqslant N$, there exists a constant $\overline{C}_n > 0$ such that for all $t \in [0,T]$ and $\varepsilon > 0$,*

$$\mathbb{E}_{\mathbb{Q}}\left[\left|\frac{X_t^\varepsilon}{\varepsilon}\right|^n\right] \leqslant \overline{C}_n, \tag{A.1}$$

*where $\mathbb{Q}$ is the probability measure defined in (3.5).*

---

[14]Uniform convergence results would require extending our maximal inequality to sublinear mean-reversion speeds.

*Proof.* Note that $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is square integrable under $\mathbb{P}$ by Assumption 1 and Novikov's criterion. The Cauchy-Schwarz inequality implies that

$$\mathbb{E}_{\mathbb{Q}}\left[|A|\right] \leqslant \mathbb{E}\left[A^2\right]^{\frac{1}{2}} \mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right]^{\frac{1}{2}} < \infty$$

for any random variable $A$ that is square-integrable under $\mathbb{P}$. This estimate will be used throughout the proof without further mention.

Define the process $(\overline{X}^\varepsilon_s)_{s\in[0,\varepsilon^{-2}T]}$ by

$$\overline{X}^\varepsilon_s := X^\varepsilon_{\varepsilon^2 s}/\varepsilon.$$

Then, $\overline{X}^\varepsilon$ satisfies under $\mathbb{Q}$ the SDE

$$d\overline{X}^\varepsilon_s = -L_{\varepsilon^2 s}c_{\varepsilon^2 s}g\left(M_{\varepsilon^2 s}\overline{X}^\varepsilon_s\right)ds + \sqrt{c_{\varepsilon^2 s}}d\overline{W}^{\varepsilon,\mathbb{Q}}_s, \quad \overline{X}^\varepsilon_0 = x^\varepsilon_0/\varepsilon. \tag{A.2}$$

Moreover, for $2 \leqslant n \leqslant N$, the $n$-th power of $\left|\overline{X}^\varepsilon\right|$ satisfies the SDE

$$d\left|\overline{X}^\varepsilon_s\right|^n = \left(-n\,\overline{X}^\varepsilon_s\left|\overline{X}^\varepsilon_s\right|^{n-2} L_{\varepsilon^2 s}c_{\varepsilon^2 s}g\left(M_{\varepsilon^2 s}\overline{X}^\varepsilon_s\right) + \frac{n(n-1)}{2}c_{\varepsilon^2 s}\left|\overline{X}^\varepsilon_s\right|^{n-2}\right)ds$$

$$+ n\,\operatorname{sgn}\left(\overline{X}^\varepsilon_s\right)\left|\overline{X}^\varepsilon_s\right|^{n-1}\sqrt{c_{\varepsilon^2 s}}d\overline{W}^{\varepsilon,\mathbb{Q}}_s, \quad \left|\overline{X}^\varepsilon_0\right|^n = (x^\varepsilon_0/\varepsilon)^n, \tag{A.3}$$

where $\overline{W}^{\varepsilon,\mathbb{Q}}$ is a $\mathbb{Q}$-Brownian motion.

Comparing (A.3) for $n = 2$ to the SDE

$$d\overline{Y}^\varepsilon_s = c_{\varepsilon^2 s}ds + 2\sqrt{\overline{Y}^\varepsilon_s}\sqrt{c_{\varepsilon^2 s}}d\overline{B}^{\varepsilon,\mathbb{Q}}_s, \quad \overline{Y}^\varepsilon_0 = (x^\varepsilon_0)^2/\varepsilon^2,$$

where $\overline{B}^{\varepsilon,\mathbb{Q}}_s = \int_0^s \operatorname{sgn}\left(\overline{X}^\varepsilon_u\right)d\overline{W}^{\varepsilon,\mathbb{Q}}_u$ for $s \in [0, \varepsilon^{-2}T]$, we obtain by Lemma C.1 that

$$\mathbb{Q}\left[\left|\overline{X}^\varepsilon_s\right|^2 \leqslant \overline{Y}^\varepsilon_s, \text{ for all } s \in [0, \varepsilon^{-2}T]\right] = 1.$$

By Lemma D.2 and the integrability assumption on $c$ this implies that $\sup_{t\in[0,\varepsilon^2 T]}\left|\overline{X}^\varepsilon_t\right|$ has moments of orders up to $2N$ under $\mathbb{Q}$. Together with Hölder's inequality, it follows that, for each $2 \leqslant n \leqslant N$,

$$\mathbb{E}_{\mathbb{Q}}\left[\int_0^{\varepsilon^{-2}T} n^2\left|\overline{X}^\varepsilon_s\right|^{2(n-1)}c_{\varepsilon^2 s}ds\right] \leqslant n^2\varepsilon^{-2}T\mathbb{E}_{\mathbb{Q}}\left[\sup_{s\in[0,\varepsilon^{-2}T]}\left(\overline{Y}^\varepsilon_s\right)^n\right]^{\frac{n-1}{n}}\mathbb{E}_{\mathbb{Q}}\left[\sup_{u\in[0,T]}c^n_u\right]^{\frac{1}{n}} < \infty.$$

Hence, the local martingale term in (A.3) is a $\mathbb{Q}$-martingale.

We now show that the supremum of the positive part of the integrand in the $ds$-term in (A.3) is $\mathbb{Q}$-integrable. To this end, we compute

$$\sup_{s\in[0,\varepsilon^{-2}T]}\left\{\left(-n\,\overline{X}^\varepsilon_s\left|\overline{X}^\varepsilon_s\right|^{n-2} L_{\varepsilon^2 s}c_{\varepsilon^2 s}g\left(M_{\varepsilon^2 s}\overline{X}^\varepsilon_s\right) + \frac{n(n-1)}{2}c_{\varepsilon^2 s}\left|\overline{X}^\varepsilon_s\right|^{n-2}\right)^+\right\}$$

$$\leqslant \sup_{s\in[0,\varepsilon^{-2}T]}\frac{n(n-1)}{2}c_{\varepsilon^2 s}\left|\overline{X}^\varepsilon_s\right|^{n-2}.$$

Now the claim follows from the Cauchy-Schwarz inequality, the stated integrability assumptions and the fact that $\sup_{t\in[0,\varepsilon^{-2}T]}\left|\overline{X}^\varepsilon_t\right|$ has moments of orders up to $2N$ under $\mathbb{Q}$.

21

Thus, the function $j_{n,\varepsilon} : [0, \varepsilon^{-2}T] \to \mathbb{R} \cup \{-\infty\}$ given by

$$j_{n,\varepsilon}(t) = -n\mathbb{E}_{\mathbb{Q}}\left[\overline{X}_t^\varepsilon \left|\overline{X}_t^\varepsilon\right|^{n-2} L_{\varepsilon^2 t} c_{\varepsilon^2 t} g\left(M_{\varepsilon^2 t} \overline{X}_t^\varepsilon\right)\right] + \frac{n(n-1)}{2}\mathbb{E}_{\mathbb{Q}}\left[c_{\varepsilon^2 t}\left|\overline{X}_t^\varepsilon\right|^{n-2}\right] \tag{A.4}$$

is well defined and bounded from above. Define the function $J_{n,\varepsilon} : [0, \varepsilon^{-2}T] \to \mathbb{R}$ by $J_{n,\varepsilon}(t) = \mathbb{E}_{\mathbb{Q}}\left[\left|\overline{X}_t^\varepsilon\right|^n\right]$. After taking $\mathbb{Q}$-expectations in (A.3), Fubini's theorem and the fundamental theorem of calculus for Lebesgue-measurable functions show that $J_{n,\varepsilon}$ is absolutely continuous and almost everywhere differentiable with derivative $j_{n,\varepsilon}$. In particular $j_{n,\varepsilon}$ is almost everywhere finite and Lebesgue-integrable.

We proceed to derive a differential inequality for $J_{n,\varepsilon}$. To this end, we first use the fact that $g$ is odd and that by condition (1.2), $xg(x) \geqslant 0$ and $xg(x) \geqslant a|x|^{q+1} - \widetilde{a}$ for some constants $a, \widetilde{a} > 0$ and all $x \in \mathbb{R}$. This yields

$$j_{n,\varepsilon}(t) \leqslant -n\mathbb{E}_{\mathbb{Q}}\left[\frac{L_{\varepsilon^2 t} c_{\varepsilon^2 t} \wedge M_{\varepsilon^2 t}}{M_{\varepsilon^2 t}}\left|\overline{X}_t^\varepsilon\right|^{n-2}\left(a\left|\overline{X}_t^\varepsilon\right|^{q+1} M_{\varepsilon^2 t}^{q+1} - \widetilde{a}\right)\right] + \frac{n(n-1)}{2}\mathbb{E}_{\mathbb{Q}}\left[c_{\varepsilon^2 t}\left|\overline{X}_t^\varepsilon\right|^{n-2}\right]$$

$$\leqslant -an\mathbb{E}_{\mathbb{Q}}\left[\left(L_{\varepsilon^2 t} c_{\varepsilon^2 t} \wedge M_{\varepsilon^2 t}\right)^{q+1}\left|\overline{X}_t^\varepsilon\right|^{n-1+q}\right] + \mathbb{E}_{\mathbb{Q}}\left[\left(n\widetilde{a} + \frac{n(n-1)}{2}c_{\varepsilon^2 t}\right)\left|\overline{X}_t^\varepsilon\right|^{n-2}\right]. \tag{A.5}$$

We now estimate the first term on the right-hand side of (A.5). We consider two cases separately for the growth condition (1.2): strictly superlinear growth ($\liminf_{x\to\infty}\frac{g(x)}{x^q} > 0$ for $q > 1$) and linear growth ($\liminf_{x\to\infty}\frac{g(x)}{x} > 0$). In the first case, we use the reverse Hölder inequality (with power $\frac{n}{n+q-1}$) and the inequality $|x|^{\frac{n+q-1}{n}} \geqslant |x| - 1$ for $x \in \mathbb{R}$. This gives

$$\mathbb{E}_{\mathbb{Q}}\left[\left(L_{\varepsilon^2 t} c_{\varepsilon^2 t} \wedge M_{\varepsilon^2 t}\right)^{q+1}\left|\overline{X}_t^\varepsilon\right|^{n-1+q}\right] \geqslant \mathbb{E}_{\mathbb{Q}}\left[\left|\overline{X}_t^\varepsilon\right|^n\right]^{\frac{n+q-1}{n}}\mathbb{E}_{\mathbb{Q}}\left[\left(L_{\varepsilon^2 t} c_{\varepsilon^2 t} \wedge M_{\varepsilon^2 t}\right)^{-\frac{(q+1)n}{q-1}}\right]^{-\frac{q-1}{n}}$$

$$\geqslant \left(\mathbb{E}_{\mathbb{Q}}\left[\left|\overline{X}_t^\varepsilon\right|^n\right] - 1\right)\mathbb{E}_{\mathbb{Q}}\left[\sup_{u\in[0,T]}\left(L_u c_u \wedge M_u\right)^{-\frac{(q+1)n}{q-1}}\right]^{-\frac{q-1}{n}}.$$

In the second case, we obtain similarly from Assumption 2 (see Remark 1.3) that

$$\mathbb{E}_{\mathbb{Q}}\left[\left(L_{\varepsilon^2 t} c_{\varepsilon^2 t} \wedge M_{\varepsilon^2 t}\right)^2\left|\overline{X}_t^\varepsilon\right|^n\right] \geqslant \text{ess inf}_{u\in[0,T]}\left\{\left(L_u c_u \wedge M_u\right)^2\right\}\mathbb{E}_{\mathbb{Q}}\left[\left|\overline{X}_t^\varepsilon\right|^n\right]$$

Plugging this into (A.5) and setting

$$C_n := \begin{cases} an \, \mathbb{E}_{\mathbb{Q}}\left[\sup_{u\in[0,T]}\left(L_u c_u \wedge M_u\right)^{-\frac{(q+1)n}{q-1}}\right]^{-\frac{q-1}{n}}, & \text{if } q > 1, \\ an \, \text{ess inf}_{u\in[0,T]}\left\{\left(L_u c_u \wedge M_u\right)^2\right\}, & \text{if } q = 1, \end{cases}$$

this yields

$$j_{n,\varepsilon}(t) \leqslant -C_n J_{n,\varepsilon}(t) + C_n \mathbb{1}_{\{q>1\}} + \mathbb{E}_{\mathbb{Q}}\left[\left(n\widetilde{a} + \frac{n(n-1)}{2}c_{\varepsilon^2 t}\right)\left|\overline{X}_t^\varepsilon\right|^{n-2}\right]. \tag{A.6}$$

We now establish (A.1) for $n = 2$ and use it to prove the result for general integer $n \geqslant 3$ by induction. For $n = 2$, the $\mathbb{Q}$-expectation on the right hand side of (A.6) can be bounded by $\mathbb{E}_{\mathbb{Q}}\left[2\widetilde{a} + \sup_{u\in[0,T]} c_u\right]$ and then (A.1) follows from Lemma A.2. For $n \geqslant 3$, assume that $J_{n-1,\varepsilon}(t) \leqslant$

$\overline{C}_{n-1}$ for some $\overline{C}_{n-1} > 0$. Then Hölder's inequality and the inequality $(a+b)^{n-1} \leq 2^{n-1}(a^{n-1}+b^{n-1})$ for $a, b \geq 0$ give

$$j_{n,\varepsilon}(t) \leqslant -C_n J_{n,\varepsilon}(t) + C_n \mathbb{1}_{\{q>1\}} + 2\left((n\tilde{a})^{n-1} + \left(\frac{n(n-1)}{2}\right)^{n-1}\mathbb{E}_\mathbb{Q}\left[\sup_{u \in [0,T]} c_u^{n-1}\right]\right)^{\frac{1}{n-1}} \overline{C}_{n-1}^{\frac{n-2}{n-1}}.$$

Now (A.1) follows from Lemma A.2. $\qquad\square$

The following differential inequality is used in the proof of Lemma A.1.

**Lemma A.2.** *Let $C, C' > 0 \in \mathbb{R}$ be constants and $J : [0, T] \to \mathbb{R}$ be an absolutely continuous function with almost everywhere derivative $j$ satisfying the differential inequality*

$$j(t) \leqslant C - C'J(t) \text{ a.e.}, \quad J(0) = J_0.$$

*Then there exists a positive constant $C''$ such that $J(t) \leqslant C''$ on $[0, T]$.*

*Proof.* Define the function $t \mapsto C_1(t) := j(t) + C'J(t)$. Then $C_1 \leqslant C$ a.e. and $J$ satisfies the ODE

$$j(t) = C_1(t) - C'J(t) \text{ a.e}, \quad J(0) = J_0.$$

This ODE has the explicit solution $J(t) = e^{-C't}\left(J_0 + \int_0^t C_1(s)e^{C's}ds\right)$, which yields the desired bound:

$$J(t) \leqslant J_0 + \frac{C}{C'}, \quad t \in [0, T]. \qquad\square$$

Note that we could not use Gronwall inequality here, as the factor in front of the function $J$ is negative.

## A.2   A Uniform Integrability Result

With the help of Lemma A.1, we can now establish our result on uniform integrability:

**Lemma A.3.** *Suppose that Assumptions 1 and 2 are satisfied. Then:*

$$\sup_{\varepsilon>0} \mathbb{E}\left[\int_0^T H_t^{p(1+\eta)} f\left(\frac{K_t X_t^\varepsilon}{\varepsilon}\right)^{p(1+\eta)} dt\right] < \infty. \tag{A.7}$$

*Proof.* Using the polynomial bound (1.3) for $f$ and Hölder's inequality, we obtain

$$\mathbb{E}\left[\int_0^T H_t^{p(1+\eta)} f\left(\frac{K_t X_t^\varepsilon}{\varepsilon}\right)^{p(1+\eta)} dt\right]$$

$$\leqslant C_f \mathbb{E}\left[\int_0^T H_t^{p(1+\eta)} \left|\frac{K_t X_t^\varepsilon}{\varepsilon}\right|^{pq'(1+\eta)} dt\right] + C_f \mathbb{E}\left[\int_0^T H_t^{p(1+\eta)} dt\right]$$

$$\leqslant C_f \mathbb{E}\left[\int_0^T \left(H_t K_t^{q'}\right)^{2p(1+\eta)} dt\right]^{\frac{1}{2}} \mathbb{E}\left[\int_0^T \left|\frac{X_t^\varepsilon}{\varepsilon}\right|^{2pq'(1+\eta)} dt\right]^{\frac{1}{2}} + C_f \mathbb{E}\left[\int_0^T H_t^{p(1+\eta)} dt\right]. \tag{A.8}$$

By Assumption 2, the first expectation in the first term and the last term on the right-hand side of (A.8) are finite and independent of $\varepsilon$. By Fubini's theorem, Lemma A.1, Hölder's inequality and Assumption 1 the second expectation in the first term is also finite,

$$
\mathbb{E}\left[\int_0^T \left|\frac{X_t^\varepsilon}{\varepsilon}\right|^{2pq'(1+\eta)} dt\right] = \int_0^T \mathbb{E}\left[\left|\frac{X_t^\varepsilon}{\varepsilon}\right|^{2pq'(1+\eta)}\right] dt = \int_0^T \mathbb{E}_\mathbb{Q}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\left|\frac{X_t^\varepsilon}{\varepsilon}\right|^{2pq'(1+\eta)}\right] dt
$$

$$
\leqslant \int_0^T \mathbb{E}_\mathbb{Q}\left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)^2\right]^{\frac{1}{2}} \mathbb{E}_\mathbb{Q}\left[\left|\frac{X_t^\varepsilon}{\varepsilon}\right|^{4pq'(1+\eta)}\right]^{\frac{1}{2}} dt
$$

$$
= T\mathbb{E}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right]^{\frac{1}{2}} \sqrt{\overline{C}_{4pq'(1+\eta)}} < \infty.
$$

This concludes the proof. $\qquad\square$

# B   A Maximal Inequality for Square-Root Processes

In this section, we establish a maximal inequality for square-root processes which is inspired by a result of Peskir [29].[15]  This estimate is crucial for establishing Lemma B.4 which in turn is necessary to concatenate the infinitesimal estimates from Proposition 4.4 to establish Theorem 1.4 in Section 4.2, but also of independent interest.

**Proposition B.1.** *Let $(Y_t)_{t\geqslant 0}$ be the unique strong solution of the SDE*

$$
dY_t = \nu(\theta - Y_t)\,dt + \sigma\sqrt{Y_t}dB_t, \quad Y_0 = y_0, \tag{B.1}
$$

*where $y_0 > 0$, the constants $\nu, \theta, \sigma > 0$ satisfy $\frac{2\nu\theta}{\sigma^2} < 1$, and $(B_t)_{t\geqslant 0}$ is a Brownian motion on some filtered probability space. Set $\gamma := \frac{2\nu}{\sigma^2}$ and, for $n \in \mathbb{N}$, define*

$$
C^1(y_0, \gamma, \sigma, n) := \left(1 + 8^n\left(y_0^n + \gamma^{-2n} + 4^n\sigma^n\right)\right)\left(2 + \left(2^n y_0^{n-1} + \frac{2 + 8^n\gamma^{-2n}}{y_0}\right)\left(1 + \frac{12}{y_0}\left((n+1)!\right)\gamma^{-1}\right)\right),
$$

$$
C^2(y_0, \gamma, n) := 4^n\gamma^{-n}\left(2 + \left(2^n y_0^{n-1} + \frac{2 + 8^n\gamma^{-2n}}{y_0}\right)\left(1 + \frac{12}{y_0}\left((n+1)!\right)\gamma^{-1}\right)\right).
$$

*Then for any $n \geqslant 0$ and any finite stopping time $\tau$,*

$$
\mathbb{E}\left[\max_{0\leqslant t\leqslant \tau} Y_t^n\right] \leqslant C^1(y_0, \gamma, \sigma, n) + C^2(y_0, \gamma, n)\mathbb{E}\left[\log(\tau \vee 1)^n\right]. \tag{B.2}
$$

*Proof.* For $n = 0$, the claim is trivial. So fix $n > 0$. The SDE satisfied by $Y^n$ for $n \in \mathbb{N}\backslash\{0\}$ is

$$
d\left(Y_t^n\right) = \left(n\left(\nu\theta + \frac{n-1}{2}\sigma^2\right)\left(Y_t^n\right)^{\frac{n-1}{n}} - n\nu\left(Y_t^n\right)\right) dt + n\sigma\left(Y_t^n\right)^{\frac{n-\frac{1}{2}}{n}} dB_t, \quad Y_0^n = y_0^n.
$$

We sharpen the arguments of [29] in the present context. We first compute for $z \geqslant y_0^n$ the scale function $S$ and the speed measure $m$ of the diffusion $Y^n$. The derivative of $S$ is given by

$$
S'(z) = \exp\left(-2\int_{y_0^n}^z \frac{n\left(\nu\theta + \frac{n-1}{2}\sigma^2\right)y^{\frac{n-1}{n}} - n\nu y}{n^2\sigma^2 y^{\frac{2n-1}{n}}}dy\right) = \left(\frac{z}{y_0^n}\right)^{-\frac{2\left(\nu\theta + \frac{n-1}{2}\sigma^2\right)}{n\sigma^2}}\exp\left(\frac{2\nu}{\sigma^2}\left(z^{\frac{1}{n}} - y_0\right)\right).
$$

---

[15]For square-root processes, the arguments from [29] can be sharpened to obtain constants explicit in the model parameters. This is needed for the application of the estimate in the companion paper [11].

The speed measure $m$ in turn is given by

$$m(y_0^n, z] = \int_{y_0^n}^z \frac{2}{n^2 \sigma^2 S'(y)} y^{-\frac{2n-1}{n}} dy$$

$$= \frac{2}{n^2 \sigma^2} \exp\left(\frac{2\nu y_0}{\sigma^2}\right) y_0^{-\frac{2\nu\theta}{\sigma^2}+1-n} \int_{y_0^n}^z y^{\frac{-(2n-1)\sigma^2+2\left(\nu\theta+\frac{n-1}{2}\sigma^2\right)}{n\sigma^2}} \exp\left(-\frac{2\nu}{\sigma^2} y^{\frac{1}{n}}\right) dy$$

$$= \frac{2}{n^2 \sigma^2} \exp\left(\frac{2\nu y_0}{\sigma^2}\right) y_0^{-\frac{2\nu\theta}{\sigma^2}+1-n} \int_{y_0^n}^z y^{\frac{2\nu\theta}{n\sigma^2}-1} \exp\left(-\frac{2\nu}{\sigma^2} y^{\frac{1}{n}}\right) dy$$

$$= \frac{2}{n^2 \sigma^2} \exp\left(\frac{2\nu y_0}{\sigma^2}\right) y_0^{-\frac{2\nu\theta}{\sigma^2}+1-n} \int_{y_0}^{z^{\frac{1}{n}}} x^{\frac{2\nu\theta}{\sigma^2}-1} \exp\left(-\frac{2\nu}{\sigma^2} x\right) dx.$$

Define the function $F : [y_0^n, \infty) \to [0, \infty)$ (as in [29, Theorem 2.5]) for $x \geqslant y_0^n$ by

$$F(x) = \int_{y_0^n}^x m(y_0^n, z] S'(z) dz = \frac{2}{n\sigma^2} \int_{y_0^n}^x z^{-\frac{2\left(\nu\theta+\frac{n-1}{2}\sigma^2\right)}{n\sigma^2}} \exp\left(\frac{2\nu}{\sigma^2} z^{\frac{1}{n}}\right) \int_{y_0}^{z^{\frac{1}{n}}} y^{\frac{2\nu\theta}{\sigma^2}-1} \exp\left(-\frac{2\nu}{\sigma^2} y\right) dy \, dz$$

$$= \frac{2}{\sigma^2} \int_{y_0}^{x^{\frac{1}{n}}} v^{-\frac{2\nu\theta}{\sigma^2}} e^{\frac{2\nu}{\sigma^2} v} \int_{y_0}^v y^{\frac{2\nu\theta}{\sigma^2}-1} e^{-\frac{2\nu}{\sigma^2} y} dy \, dv$$

$$= \frac{2}{\sigma^2} \int_{y_0}^{x^{\frac{1}{n}}} v^{-\gamma\theta} e^{\gamma v} \int_{y_0}^v y^{\gamma\theta-1} e^{-\gamma y} dy \, dv, \tag{B.3}$$

where $\gamma := \frac{2\nu}{\sigma^2}$. It is not difficult to check that $F$ is strictly increasing with $\lim_{x\to\infty} F(x) = +\infty$, continuously differentiable on $[y_0^n, \infty)$ and twice continuously differentiable on $(y_0^n, \infty)$. Next, define the function $G : [0, \infty) \to [0, \infty)$ by $G(x) = \mathbb{1}_{\{x \geqslant y_0^n\}} F(x)$. Note that $G$ is twice continuously differentiable everywhere, except at $y_0^n$. Applying Itô's formula as in [32, Exercise 4.20] to $G(Y_t^n)$ (note that the infinitesimal generator of the diffusion $Y^n$ applied to $G$ gives 0 on $[0, y_0^n)$ and 1 on $(y_0^n, +\infty)$, and that $G(Y_0^n) = 0$), we obtain

$$G(Y_t^n) \leqslant t + \int_0^t G'(Y_t^n) n\sigma (Y_t^n)^{\frac{n-\frac{1}{2}}{n}} dB_s.$$

By localization of $Y^n$ and the monotone convergence theorem ($G$ is nondecreasing), it follows that, for any finite stopping time $\tau$,

$$\mathbb{E}[G(Y_\tau^n)] \leqslant \mathbb{E}[\tau].$$

Denote by $H : [0, \infty) \to [y_0^n, \infty)$ the inverse of $F$, which like $F$ is increasing and continuously differentiable on $[0, \infty)$. Then by Lenglart's domination principle in the form of [29, Lemma 2.1] with $Z_t = G(Y_t^n)$ and $A_t = t$, we obtain for any finite stopping time $\tau$,[16]

$$\mathbb{E}\left[\sup_{0 \leqslant t \leqslant \tau} H(G(Y_t^n))\right] \leqslant \mathbb{E}\left[\widetilde{H}(\tau)\right],$$

where the function $\widetilde{H} : [0, \infty) \to [0, \infty]$ is given by

$$\widetilde{H}(y) = y \int_y^\infty \frac{1}{z} H'(z) dz + 2H(y).$$

---

[16]Note that the assumption that $H(0) = 0$ can be replaced by $H(0) \geqslant 0$ in [29, Lemma 2.1].

Now using that $H(G(x)) = x$ for $x \in [y_0^n, \infty)$ and $H(G(x)) = H(0) = y_0^n \geqslant x$ for $x \in [0, y_0^n)$, we obtain for any finite stopping time $\tau$,

$$\mathbb{E}\left[\sup_{0 \leqslant t \leqslant \tau} Y_t^n\right] \leqslant \mathbb{E}\left[\sup_{0 \leqslant t \leqslant \tau} H(G(Y_t^n))\right] \leqslant \mathbb{E}\left[\widetilde{H}(\tau)\right]. \tag{B.4}$$

We proceed to estimate $\widetilde{H}(y)$. A change of variable yields

$$\widetilde{H}(y) \leqslant \left(\sup_{x \geqslant y_0^n} \frac{F(x)}{x} \int_x^\infty \frac{dz}{F(z)} + 2\right) H(y), \quad y \in [0, \infty). \tag{B.5}$$

In order to estimate both factors on the right-hand side of (B.5), we need to establish lower and upper bounds for the function $F$.

First, we establish an upper bound for $F$. It follows from (B.3) and the assumption $\gamma\theta < 1$ that for $x \in [y_0^n, \infty)$,

$$F(x) = \frac{2}{\sigma^2} \int_{y_0}^{x^{\frac{1}{n}}} v^{-\gamma\theta} \mathrm{e}^{\gamma v} \int_{y_0}^v y^{\gamma\theta-1} \mathrm{e}^{-\gamma y} dy\ dv \leqslant \frac{2}{\sigma^2} y_0^{-\gamma\theta} y_0^{\gamma\theta-1} \int_{y_0}^{x^{\frac{1}{n}}} \mathrm{e}^{\gamma v} \int_{y_0}^v \mathrm{e}^{-\gamma y} dy\ dv$$

$$\leqslant \frac{2}{\gamma\sigma^2 y_0} \int_{y_0}^{x^{\frac{1}{n}}} \mathrm{e}^{\gamma v} \left(\mathrm{e}^{-\gamma y_0} - \mathrm{e}^{-\gamma v}\right) dv \leqslant \frac{2\mathrm{e}^{-\gamma y_0}}{\gamma\sigma^2 y_0} \int_{y_0}^{x^{\frac{1}{n}}} \mathrm{e}^{\gamma v} dv \leqslant \frac{2\mathrm{e}^{-\gamma y_0}}{\gamma^2\sigma^2 y_0} \left(\mathrm{e}^{\gamma x^{\frac{1}{n}}} - \mathrm{e}^{\gamma y_0}\right)$$

$$\leqslant \frac{2\mathrm{e}^{-\gamma y_0}}{\gamma^2\sigma^2 y_0} \mathrm{e}^{\gamma x^{\frac{1}{n}}}. \tag{B.6}$$

Next, we establish a lower bound for $F$. Set

$$\bar{x} := 2^n y_0^n + 1 + 8^n \gamma^{-2n} \geqslant 2^n y_0^n + \left(\frac{4\log(2)}{\gamma} \vee \frac{8}{\gamma^2}\right)^n \geqslant \left(y_0 + \frac{2\log(2)}{\gamma}\right)^n \vee \frac{8^n}{\gamma^{2n}},$$

where the first inequality follows from the fact that $\frac{8}{\gamma^2} > \frac{4\log(2)}{\gamma}$ for $\gamma < 2/\log(2)$ and $\frac{4\log(2)}{\gamma} \leqslant 2\log(2)^2 < 1$ for $\gamma \geqslant 2/\log(2)$, and the second inequality follows from the elementary inequality $(a+b)^n \leqslant 2^n(a^n + b^n)$ for $a, b, n \geqslant 0$.

Then using $\gamma\theta < 1$ together with the elementary inequalities of Lemma B.2, we obtain for $x \geqslant \bar{x}$:

$$F(x) = \frac{2}{\sigma^2} \int_{y_0}^{x^{\frac{1}{n}}} v^{-\gamma\theta} \mathrm{e}^{\gamma v} \int_{y_0}^v y^{\gamma\theta-1} \mathrm{e}^{-\gamma y} dy\ dv \geqslant \frac{2}{\sigma^2} \int_{y_0}^{x^{\frac{1}{n}}} v^{-\gamma\theta} v^{\gamma\theta-1} \mathrm{e}^{\gamma v} \int_{y_0}^v \mathrm{e}^{-\gamma y} dy\ dv$$

$$\geqslant \frac{2}{\gamma\sigma^2} \int_{y_0}^{x^{\frac{1}{n}}} v^{-1} \mathrm{e}^{\gamma v} \left(\mathrm{e}^{-\gamma y_0} - \mathrm{e}^{-\gamma v}\right) dv \geqslant \frac{\mathrm{e}^{-\gamma y_0}}{\gamma\sigma^2} \int_{y_0+\frac{\log(2)}{\gamma}}^{x^{\frac{1}{n}}} v^{-1} \mathrm{e}^{\gamma v} dv$$

$$\geqslant \frac{\mathrm{e}^{-\gamma y_0}}{\gamma\sigma^2} x^{-\frac{1}{n}} \int_{y_0+\frac{\log(2)}{\gamma}}^{x^{\frac{1}{n}}} \mathrm{e}^{\gamma v} dv = \frac{\mathrm{e}^{-\gamma y_0}}{\gamma^2\sigma^2} x^{-\frac{1}{n}} \left(\mathrm{e}^{\gamma x^{\frac{1}{n}}} - \mathrm{e}^{\gamma(y_0+\frac{\log(2)}{\gamma})}\right)$$

$$\geqslant \frac{\mathrm{e}^{-\gamma y_0}}{2\gamma^2\sigma^2} x^{-\frac{1}{n}} \mathrm{e}^{\gamma x^{\frac{1}{n}}} \tag{B.7}$$

$$\geqslant \frac{\mathrm{e}^{-\gamma y_0}}{2\gamma^2\sigma^2} \mathrm{e}^{\frac{\gamma}{2} x^{\frac{1}{n}}}. \tag{B.8}$$

In view of (B.8) and since $F$ is increasing, it follows that

$$H(y) \leqslant \left(\frac{2\log(y)}{\gamma} + 2y_0 + \frac{4}{\gamma}\log(2\gamma\sigma)\right)^n \leqslant \left(\frac{2\log(y)}{\gamma} + 2y_0 + 8\sigma\right)^n$$

$$\leqslant 4^n \gamma^{-n} \log(y)^n + 8^n y_0^n + 32^n \sigma^n, \quad \text{for } y \geqslant F(\bar{x}). \tag{B.9}$$

26

Here, we have used in the last two steps the elementary inequalities $\log(x) \leqslant x$ for $x \in (0, \infty)$ and $(a+b)^n \leqslant 2^n(a^n + b^n)$ for $a, b, n \geqslant 0$. Moreover, for $y \in [0, F(\bar{x})]$, using again that $F$ is increasing, we have

$$H(y) \leqslant \bar{x}. \tag{B.10}$$

Combining (B.9) and (B.10), we obtain for all $y \in [0, \infty)$,

$$H(y) \leqslant 4^n \gamma^{-n} \log(y \vee 1)^n + 1 + 8^n \left(y_0^n + \gamma^{-2n} + 4^n \sigma^n\right). \tag{B.11}$$

Finally, we derive an upper bound for $\sup_{x \geqslant 1} \frac{F(x)}{x} \int_x^\infty \frac{dz}{F(z)} + 2$. First, by (B.7), a change of variables, and Lemma B.3 (noting that $(\bar{x})^{\frac{1}{n}} \gamma \geqslant 1$), we obtain for $x \geqslant \bar{x}$ that

$$\int_x^\infty \frac{dz}{F(z)} \leqslant 2\gamma^2 \sigma^2 e^{\gamma y_0} \int_x^\infty z^{\frac{1}{n}} e^{-\gamma z^{\frac{1}{n}}} dz = 2n\gamma^2 \sigma^2 e^{\gamma y_0} \int_{x^{\frac{1}{n}}}^\infty w^n e^{-\gamma w} dw$$

$$\leqslant 6n(n!)\gamma \sigma^2 e^{\gamma y_0} x e^{-\gamma x^{\frac{1}{n}}} \leqslant 6(n+1)!\gamma \sigma^2 e^{\gamma y_0} x e^{-\gamma x^{\frac{1}{n}}}. \tag{B.12}$$

Putting together Equations (B.6) and (B.12), we obtain for $x \geqslant \bar{x}$,

$$\frac{F(x)}{x} \int_x^\infty \frac{dz}{F(z)} \leqslant \frac{12}{y_0} \left((n+1)!\right) \gamma^{-1}. \tag{B.13}$$

Using that $F$ is increasing, and the estimate (B.13), we obtain for $x \in [y_0, \bar{x}]$,

$$\frac{F(x)}{x} \int_x^\infty \frac{dz}{F(z)} \leqslant \frac{1}{y_0} \left( \int_x^{\bar{x}} \frac{F(x)}{F(z)} dz + F(\bar{x}) \int_{\bar{x}}^\infty \frac{dz}{F(z)} \right)$$

$$\leqslant \frac{1}{y_0} \left( \int_x^{\bar{x}} dz + \bar{x} \frac{F(\bar{x})}{\bar{x}} \int_{\bar{x}}^\infty \frac{dz}{F(z)} \right)$$

$$\leqslant \frac{1}{y_0} \left( \bar{x} + \bar{x} \frac{12}{y_0} \left((n+1)!\right) \gamma^{-1} \right). \tag{B.14}$$

Combining (B.14) and (B.13) gives

$$\sup_{x \geqslant 1} \frac{F(x)}{x} \int_x^\infty \frac{dz}{F(z)} + 2 \leqslant 2 + \frac{\bar{x}}{y_0} \left( 1 + \frac{12}{y_0} \left((n+1)!\right) \gamma^{-1} \right). \tag{B.15}$$

Now, the result follows from (B.4), (B.5), (B.11), and (B.15). $\qquad \square$

The following elementary estimates are used in the proof of Proposition B.1:

**Lemma B.2.** *Let $\gamma > 0$ and $y \geqslant 0$. Then:*

$$\frac{1}{x} \exp(\gamma x) \geqslant \exp\left(\frac{\gamma}{2} x\right), \quad \text{for all } x \geqslant \frac{8}{\gamma^2}, \tag{B.16}$$

$$\exp(\gamma x) - \exp(\gamma y) \geqslant \frac{1}{2} \exp(\gamma x), \quad \text{for all } x \geqslant y + \frac{\log(2)}{\gamma}, \tag{B.17}$$

$$\exp(-\gamma y) - \exp(-\gamma x) \geqslant \frac{1}{2} \exp(-\gamma y), \quad \text{for all } x \geqslant y + \frac{\log(2)}{\gamma}. \tag{B.18}$$

**Lemma B.3.** *Let $\gamma > 0$, $n \in \mathbb{N}$ and $y \geqslant \frac{1}{\gamma}$. Then:*

$$\int_y^\infty x^n \exp(-\gamma x) dx \leqslant 3(n!)\gamma^{-1} y^n \exp(-\gamma y).$$

*Proof.* Set $w := \gamma y \geqslant 1$. Then by a change of variables,

$$\int_y^\infty x^n \exp(-\gamma x) dx = \gamma^{-1-n} \int_w^\infty z^n \exp(-z) dz.$$

Moreover, integration by parts (and induction) together with $w \geqslant 1$ give

$$\int_w^\infty z^n \exp(-z) dz = \sum_{k=0}^n w^{n-k} \frac{n!}{(n-k)!} \exp(-w) \leqslant w^n \exp(-w) \sum_{k=0}^n \frac{n!}{(n-k)!}$$

$$\leqslant 3(n!)w^n \exp(-w). \qquad \square$$

Using Proposition B.1, we now establish a moment estimate for the supremum of $X^\varepsilon/\varepsilon$ that is used in Section 4.2.

**Lemma B.4.** *Suppose that Assumption 1 is satisfied and there is $\kappa \in (0,1)$ such that*

$$L_t, M_t \in \left[\kappa, \frac{1}{\kappa}\right], \quad t \in [0, T]. \tag{B.19}$$

*Set*

$$z_0 := 2\limsup_{\varepsilon \to 0} (x_0^\varepsilon)^2/\varepsilon^2 \vee 1. \tag{B.20}$$

*Then, for $n \in \mathbb{N}$ and $\varepsilon > 0$:*

$$\mathbb{E}\left[\max_{0 \leqslant t \leqslant T} \left(\frac{X_t^\varepsilon}{\varepsilon}\right)^n\right] \leqslant \mathbb{E}_\mathbb{P}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right]^{\frac{1}{2}} \left(\sqrt{C^1(z_0, a\kappa^2, 2, n)} + \sqrt{C^2(z_0, a\kappa^2, n)} \times \right.$$

$$\left. \times \mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right]^{\frac{1}{2}} \mathbb{E}\left[\log\left(\int_0^T \varepsilon^{-2} c_t dt \vee 1\right)^{2n}\right]^{\frac{1}{4}}\right)$$

$$= O\left(\log\left(\frac{1}{\varepsilon}\right)^{\frac{n}{2}}\right). \tag{B.21}$$

*Here, the functions $C^1$ and $C^2$ are defined as in Proposition B.1, and $a$ is a positive constant such that $xg(x) \geqslant a|x|^2 - \frac{1}{4}$ for all $x \in \mathbb{R}$.*

*Proof.* Using Proposition B.1, we show below that

$$\mathbb{E}_\mathbb{Q}\left[\max_{0 \leqslant t \leqslant T} \left(\frac{X_t^\varepsilon}{\varepsilon}\right)^{2n}\right] \leqslant C^1(z_0, a\kappa^2, 2, n) + C^2(z_0, a\kappa^2, n)\mathbb{E}_\mathbb{Q}\left[\log(\xi^\varepsilon \vee 1)^n\right], \tag{B.22}$$

where $\xi^\varepsilon := \int_0^T \varepsilon^{-2} c_t dt$ is as in (3.1). Then by the Cauchy–Schwarz inequality, the elementary inequality $\sqrt{b+c} \leqslant \sqrt{b} + \sqrt{c}$ for $b, c \geqslant 0$, and again the Cauchy–Schwarz inequality, we obtain

$$\mathbb{E}\left[\max_{0 \leqslant t \leqslant T} \left(\frac{X_t^\varepsilon}{\varepsilon}\right)^n\right] \leqslant \mathbb{E}_\mathbb{Q}\left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)^2\right]^{\frac{1}{2}} \mathbb{E}_\mathbb{Q}\left[\max_{0 \leqslant t \leqslant T} \left(\frac{X_t^\varepsilon}{\varepsilon}\right)^{2n}\right]^{\frac{1}{2}}$$

$$\leqslant \mathbb{E}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right]^{\frac{1}{2}} \left(\sqrt{C^1(z_0, a\kappa^2, 2, n)} + \sqrt{C^2(z_0, a\kappa^2, n)}\mathbb{E}_\mathbb{Q}\left[\log(\xi^\varepsilon \vee 1)^n\right]^{\frac{1}{2}}\right)$$

$$\leqslant \mathbb{E}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right]^{\frac{1}{2}} \left(\sqrt{C^1(z_0, a\kappa^2, 2, n)} + \sqrt{C^2(z_0, a\kappa^2, n)}\mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right]^{\frac{1}{2}} \mathbb{E}\left[\log(\xi^\varepsilon \vee 1)^{2n}\right]^{\frac{1}{4}}\right)$$

$$= O\left(\log\left(\frac{1}{\varepsilon}\right)^{\frac{n}{2}}\right).$$

It remains to prove (B.22). The definition of the rescaled and time-changed process (3.3) gives

$$\mathbb{E}_{\mathbb{Q}}\left[\max_{0\leqslant t\leqslant T}\left(\frac{X_t^\varepsilon}{\varepsilon}\right)^{2n}\right]=\mathbb{E}_{\mathbb{Q}}\left[\max_{0\leqslant\xi\leqslant\xi^\varepsilon}\left(\widetilde{X}_\xi^\varepsilon\right)^{2n}\right].\qquad(\text{B.23})$$

Under $\mathbb{Q}$, $(\widetilde{X}_\xi^\varepsilon)^2$ satisfies the SDE

$$d\big(\widetilde{X}_\xi^\varepsilon\big)^2=\left(1-2L_{u_\xi^\varepsilon}\widetilde{X}_\xi^\varepsilon g\left(M_{u_\xi^\varepsilon}\widetilde{X}_\xi^\varepsilon\right)\right)\mathbb{1}_{\{\xi\leqslant\xi^\varepsilon\}}d\xi+2\sqrt{(\widetilde{X}_\xi^\varepsilon)^2}\mathbb{1}_{\{\xi\leqslant\xi^\varepsilon\}}d\widetilde{B}_\xi^{\varepsilon,\mathbb{Q}},$$

where

$$\widetilde{B}_\xi^{\varepsilon,\mathbb{Q}}=\int_0^\xi\mathrm{sgn}\big(\widetilde{X}_y^\varepsilon\big)d\widetilde{W}_y^{\varepsilon,\mathbb{Q}}$$

is a $\mathbb{Q}$-Brownian motion stopped at $\xi^\varepsilon$. Furthermore, the growth condition of $g$ (1.2) implies that there exist a constant $a$ such that $xg(x)\geqslant a\,|x|^2-\frac{1}{4}$. In view of (B.19),

$$1-2L_{u_\xi^\varepsilon}xg\left(M_{u_\xi^\varepsilon}x\right)\leqslant 1-2\frac{L_{u_\xi^\varepsilon}\wedge M_{u_\xi^\varepsilon}}{M_{u_\xi^\varepsilon}}M_{u_\xi^\varepsilon}xg\left(M_{u_\xi^\varepsilon}x\right)\leqslant 1-2\frac{L_{u_\xi^\varepsilon}\wedge M_{u_\xi^\varepsilon}}{M_{u_\xi^\varepsilon}}\left(aM_{u_\xi^\varepsilon}^2\,|x|^2-\frac{1}{4}\right)$$

$$\leqslant\frac{3}{2}-2a\kappa^2\,|x|^2,\quad\text{a.s., for }x\in\mathbb{R}.$$

Using the comparison result established in Lemma C.1, we have $\mathbb{Q}\left[\left(\widetilde{X}_\xi^\varepsilon\right)^2\leqslant Z_\xi,\text{for all }0\leqslant\xi\leqslant\xi^\varepsilon\right]$, where $Z$ is the solution of the SDE

$$dZ_t=\left(\frac{3}{2}-2a\kappa^2Z_t\right)dt+2\sqrt{Z_t}d\widetilde{B}_t^{\varepsilon,\mathbb{Q}},\quad Z_0=z_0,$$

where $z_0$ is defined in (B.20). This is a special case of the equation (B.1) studied in Proposition B.1 with $\nu=2a\kappa^2$, $\theta=\frac{3}{4a\kappa^2}$, $\sigma=2$ and $\gamma=a\kappa^2$. Combining (B.23) and (B.2), we finally obtain the asserted estimate:

$$\mathbb{E}_{\mathbb{Q}}\left[\max_{0\leqslant t\leqslant T}\left(\frac{X_t^\varepsilon}{\varepsilon}\right)^{2n}\right]\leqslant\mathbb{E}_{\mathbb{Q}}\left[\max_{0\leqslant\xi\leqslant\xi^\varepsilon}Z_\xi^n\right]\leqslant C^1(z_0,a\kappa^2,2,n)+C^2(z_0,a\kappa^2,n)\mathbb{E}_{\mathbb{Q}}\left[\log\left(\xi^\varepsilon\vee 1\right)^n\right].$$

This completes the proof. □

## C A Comparison Result for SDEs

In this appendix, we establish a comparison result for one-dimensional SDEs that is used at various points in the proofs of our main results. It extends the standard argument from [22, Proposition 5.2.18] to the case of random initial conditions as well as drift and diffusion coefficients that are not globally Lipschitz:

**Lemma C.1.** *Let $(\Omega,\mathcal{F},\mathbb{F}=(\mathcal{F}_t)_{t\geqslant 0},\mathbb{P})$ be a filtered probability space satisfying the usual conditions. Let $\tau$ be an $\mathbb{F}$-stopping time taking values in $[0,\infty]$, $(c_t)_{t\geqslant 0}$ an $\mathbb{R}^d$-valued, locally bounded and $\mathbb{F}$-adapted process, $(W_t)_{t\geqslant 0}$ a continuous $\mathbb{F}$-adapted process that is a standard Brownian motion on $[\![0,\tau]\!]$, and $\Omega_0\in\mathcal{F}_0$. Suppose $(Y_t^{(1)})_{t\geqslant 0}$ and $(Y_t^{(2)})_{t\geqslant 0}$ are continuous $\mathbb{F}$-adapted processes that satisfy the SDEs*

$$dY_t^{(i)}=b^{(i)}(\omega,t,Y_t^{(i)})\mathbb{1}_{\{t\leqslant\tau\}}dt+h\big(c_t,Y_t^{(i)}\big)\mathbb{1}_{\{t\leqslant\tau\}}dW_t,\quad Y_0^{(i)}=y_0^{(i)},\qquad(\text{C.1})$$

29

*where $y_0^{(i)} \geqslant 0$ is $\mathcal{F}_0$-measurable and $b^{(i)}$ is $\mathcal{F}$-predictable for $i = 1, 2$, and $h$ is $1/2$-Hölder continuous in its second variable:*

$$\left| h(c, x) - h(c, y) \right| \leqslant K(c) \sqrt{|x - y|},$$

*where $K$ is a locally bounded function from $\mathbb{R}^d$ to $\mathbb{R}_+$. Set*

$$\widetilde{\Omega}_0^\tau := \{(\omega, t) \in \Omega \times [0, \infty) : \omega \in \Omega_0, t \in [0, \tau(\omega)]\}$$

*and assume that:*

*(i) $b^{(1)}(\omega, t, y) \leqslant b^{(2)}(\omega, t, y)$ for all $(\omega, t) \in \widetilde{\Omega}_0^\tau$ and $y \in \mathbb{R}_+$;*

*(ii) $y_0^{(1)}(\omega) \leqslant y_0^{(2)}(\omega)$ for all $\omega \in \Omega_0$;*

*(iii) Either $b^{(1)}$ or $b^{(2)}$ is locally one-sided Lipschitz in $x$, uniformly in $(\omega, t) \in \widetilde{\Omega}_0^\tau$, i.e.*

$$\forall x_1 \in \mathbb{R}, \exists r_{x_1} : \forall x_2 \in [x_1 - r_{x_1}, x_1], \quad b^{(i)}(\omega, t, x_1) - b^{(i)}(\omega, t, x_2) \leqslant K_{x_1}(x_1 - x_2),$$

*for some $K_{x_1} > 0$ and all $(t, \omega) \in \Omega_0^\tau$.*

*Then*

$$\mathbb{P}\left[ \mathbb{1}_{\Omega_0} Y_t^{(1)} \leqslant \mathbb{1}_{\Omega_0} Y_t^{(2)}, \text{ for all } t \geqslant 0 \right] = 1. \tag{C.2}$$

*Proof.* We may assume without loss of generality that $b^{(1)}$ satisfies (iii). By a standard localization argument, we may assume that $b^{(1)}$ is globally one-sided Lipschitz in $x$, uniformly in $(\omega, t) \in \widetilde{\Omega}_0^\tau$, with Lipschitz constant $K > 0$. By a further localization argument, we may assume that $c$, the function $K$, $Y^{(1)}$ and $Y^{(2)}$ are bounded by a constant $L > 0$.

By the construction in the proof of [22, Proposition 5.2.13], there exists a nondecreasing sequence $(\varphi_n)_{n \in \mathbb{N} \setminus \{0\}}$ of nonnegative $C^2$ functions such that:

(a) for each $n$, $\varphi_n$ is supported on $[a_n, \infty)$ for some constant $a_n > 0$, and satisfies $0 \leqslant \varphi_n'(x) \leqslant 1$ and $0 \leqslant \varphi_n''(x) \leqslant \frac{2}{nx}$ for $x > 0$;

(b) $\lim_{n \to \infty} \varphi_n(x) = x^+$ for $x \in \mathbb{R}$.

Fix $t > 0$ and $n \in \mathbb{N} \setminus \{0\}$. Set $\Delta_t := Y_t^{(1)} - Y_t^{(2)}$. By Itô's formula, the $1/2$-Hölder continuity of $h$, the fact that $0 \leqslant \varphi_n''(x) \leqslant \frac{2}{nx}$ for $x > 0$ by Property (a), and the assumption that $K \leqslant L$, we obtain

$$\begin{aligned}
\varphi_n(\Delta_t) = {} & \varphi_n(\Delta_0) + \int_0^t \varphi_n'(\Delta_s) \left( b^{(1)}(\omega, s, Y_s^{(1)}) - b^{(2)}(\omega, s, Y_s^{(2)}) \right) \mathbb{1}_{\{s \leqslant \tau\}} ds \\
& + \frac{1}{2} \int_0^t \varphi_n''(\Delta_s) \left( h(c_s, Y_s^{(1)}) - h(c_s, Y_s^{(2)}) \right)^2 \mathbb{1}_{\{s \leqslant \tau\}} ds \\
& + \int_0^t \varphi_n'(\Delta_s) \left( h(c_s, Y_s^{(1)}) - h(c_s, Y_s^{(2)}) \right) \mathbb{1}_{\{s \leqslant \tau\}} dW_s \\
\leqslant {} & \varphi_n(\Delta_0) + \int_0^t \varphi_n'(\Delta_s) \left( b^{(1)}(\omega, s, Y_s^{(1)}) - b^{(2)}(\omega, s, Y_s^{(2)}) \right) \mathbb{1}_{\{s \leqslant \tau\}} ds + \frac{tL}{n} \\
& + \int_0^{t \wedge \tau} \varphi_n'(\Delta_s) \left( h(c_s, Y_s^{(1)}) - h(c_s, Y_s^{(2)}) \right) \mathbb{1}_{\{s \leqslant \tau\}} dW_s. \tag{C.3}
\end{aligned}$$

Now multiply the inequality (C.3) with $\mathbb{1}_{\Omega_0}$, use that $\Delta_0 \leqslant 0$ on $\Omega_0$ by Assumption (ii), and note that $\int_0^\cdot \varphi_n'(\Delta_s) h(c_s, Y_s^{(i)}) dW_s$ is a martingale for $i \in \{1, 2\}$ since $K(c)$, $Y^{(1)}$ and $Y^{(2)}$ are bounded

and $0 \leqslant \varphi'_n(x) \leqslant 1$ for $x > 0$ by Assumption (a). Also taking into account Assumption (i), that $\varphi'(x) = 0$ on $\mathbb{R}_-$ and that $b^{(1)}$ is one-sided Lipschitz in $x$, uniformly in $(\omega, t) \in \widetilde{\Omega}_0^\tau$, with constant $K > 0$, it follows that

$$
\begin{aligned}
\mathbb{E}\left[\mathbb{1}_{\Omega_0}\varphi_n(\Delta_t)\right] &\leqslant 0 + \mathbb{E}\left[\int_0^t \mathbb{1}_{\{\Delta_s > 0\}}\varphi'_n(\Delta_s)\mathbb{1}_{\Omega_0}\mathbb{1}_{\{s \leqslant \tau\}}\Big(b^{(1)}(\omega, s, Y_s^{(1)}) - b^{(1)}(\omega, s, Y_s^{(2)})\Big)ds\right] \\
&\quad + \mathbb{E}\left[\int_0^t \varphi'_n(\Delta_s)\mathbb{1}_{\Omega_0}\mathbb{1}_{\{s \leqslant \tau\}}\Big(b^{(1)}(\omega, s, Y_s^{(2)}) - b^{(2)}(\omega, s, Y_s^{(2)})\Big)ds\right] + \frac{tL}{n} + 0 \\
&\leqslant K\mathbb{E}\left[\int_0^t \mathbb{1}_{\Omega_0}|\Delta_s|\mathbb{1}_{\{\Delta_s > 0\}}ds\right] + 0 + \frac{tL}{n}.
\end{aligned}
$$

Letting $n \to \infty$, monotone convergence, Property (b) and Fubini's theorem give

$$
\mathbb{E}\left[\mathbb{1}_{\Omega_0}\left(\Delta_t\right)^+\right] \leqslant K \int_0^t \mathbb{E}\left[\mathbb{1}_{\Omega_0}\left(\Delta_s\right)^+\right]ds.
$$

Now apply Gronwall's inequality to the function

$$
h(s) := \mathbb{E}\left[\mathbb{1}_{\Omega_0}\left(\Delta_s\right)^+\right] \geqslant 0.
$$

This yields $h(s) = 0$ for $s \in [0, t]$, and in turn $\mathbb{1}_{\Omega_0}Y_s^{(1)} \leqslant \mathbb{1}_{\Omega_0}Y_s^{(2)}$ $\mathbb{P}$-a.s. for $s \in [0, t]$. The result now follows from the continuity of the paths of $Y^1$ and $Y^2$. $\qquad\square$

**Remark C.2.** Note that this rather general comparison result accommodates, in particular, "bang-bang"-controlled SDEs of the form

$$
dX_t^\varepsilon = -\frac{1}{\varepsilon}\mathrm{sgn}(X_t^\varepsilon)dt + dW_t.
$$

This is because our arguments only require the drift functional to be one-sided Lipschitz, as was kindly pointed out to as by one of the anonymous reviewers.

# D   Auxiliary Results

## D.1   Existence results for SDEs

The following two strong existence results are somewhat nonstandard because the volatility functions are not locally Lipschitz at 0 and, in the second result, the drift and volatility coefficient are not necessarily Markov.

**Lemma D.1.** Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, $y_0 \geqslant 0$ be an $\mathcal{F}_0$-measurable random variable, and $(W_t)_{t\geqslant 0}$ an $\mathbb{F}$-Brownian motion. Moreover, let $b : \mathbb{R} \to \mathbb{R}$ be locally Lipschitz, nonnegative on $\mathbb{R}^+$, odd and null at zero. Then the SDE

$$
dY_t = \left(1 - 2\sqrt{Y_t}b\left(\sqrt{Y_t}\right)\right)dt + 2\sqrt{Y_t}dW_t, \quad Y_0 = y_0, \tag{D.1}
$$

has a unique strong solution.

*Proof.* By [22, Corollary 5.3.23], it suffices to show that weak existence and pathwise uniqueness hold for the SDE (D.1). To establish weak existence, consider the SDE

$$dZ_t = -b(Z_t)dt + dW_t, \quad Z_0 = \sqrt{y_0}. \tag{D.2}$$

It follows by the same argument as in Proposition 1.1 that $Z$ has a unique strong solution. Now set $Y := Z^2$ and define the Brownian motion $B$ by

$$B_t = \int_0^t \operatorname{sgn}(Z_t)dW_t.$$

Then $Y$ satisfies the SDE

$$dY_t = \left(1 - 2\sqrt{Y_t}b(\sqrt{Y_t})\right)dt + 2\sqrt{Y_t}dB_t, \quad Y_0 = y_0,$$

and so (D.1) has a weak solution.

Pathwise uniqueness follows from Lemma C.1 with $\tau = +\infty$, $\Omega_0 = \Omega$ and $b^1(\omega, t, y) = b^2(\omega, t, y) = 1 - 2\sqrt{y}b(\sqrt{y})$ and $y_0^1 = y_0^2$. Note that $x \mapsto \sqrt{x}b(\sqrt{x})$ is locally Lipschitz under the assumptions on $b$. $\qquad\square$

**Lemma D.2.** *Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions, $y_0 \geqslant 0$ be an $\mathcal{F}_0$-measurable random variable, and $(W_t)_{t \in [0,T]}$ an $\mathbb{F}$-Brownian motion. Moreover, let $n \in \mathbb{N} \setminus \{0\}$ and $(c_t)_{t \in [0,T]}$ be a positive, continuous and $\mathbb{F}$-adapted process satisfying*

$$\mathbb{E}\left[\int_0^T c_s^{2n}ds\right] < \infty. \tag{D.3}$$

*Then the SDE*

$$dY_t = c_tdt + 2\sqrt{Y_t}\sqrt{c_t}dW_t, \quad Y_0 = y_0, \tag{D.4}$$

*has a unique strong solution. Moreover,*

$$\mathbb{E}\left[\sup_{t \in [0,T]}(Y_t)^{2n}\right] < \infty. \tag{D.5}$$

*Proof.* Existence of a unique strong solution follows by a time change argument. Indeed, use the time change from Lemma 3.1 with $\varepsilon = 1$ and just write $u_\xi$ instead of $u_\xi^1$. Then writing $\widetilde{W}_\xi$ instead of $\widetilde{W}_\xi^1$ and setting $\widetilde{Y}_\xi := Y_{u_\xi}$, it suffices to show that the SDE

$$dY_\xi = d\xi + 2\sqrt{Y_\xi}d\widetilde{W}_\xi, \quad \widetilde{Y}_0 = y_0$$

has a unique strong solution. This is clear as this is the SDE satisfied by the square of a one-dimensional Bessel process started at $y_0$; cf. [31, Definition XI.1.1] and note that this result extends to non-trivial initial condition by virtue of [22, Corollary 5.3.23].

We proceed to derive (D.5). Define the process $Z$ by

$$dZ_t = \sqrt{c_t}dW_t, \quad Z_0 = \sqrt{y_0}. \tag{D.6}$$

Then $Z$ is a martingale with finite $4n$-th moments by (D.3). In particular by the Burkholder-Davis-Gundy inequality, $\mathbb{E}\left[\sup_{t \in [0,T]} Z_t^{4n}\right] < \infty$. Define the process $Y$ by $Y := Z^2$. Then $Y$ is a weak solution of the SDE (D.4). Now (D.5) follows from the fact that $\sup_{t \in [0,T]} Y_t^{2n} = \sup_{t \in [0,T]} Z_t^{4n}$ and uniqueness in law of any strong or weak solution to (D.4). $\qquad\square$

## D.2   An Ergodic Result

This section contains an ergodic theorem for the one-dimensional diffusions defined in (4.14), which is used in the proof of Lemma 4.3. For constants $l, m > \delta > 0$ and $y \geqslant 0$, consider the following two SDEs on some filtered probability space:

$$dY_t^\pm = \left(1 - 2(l \mp \delta)\sqrt{Y_t^\pm}g\left((m \mp \delta)\sqrt{Y_t^\pm}\right)\right)dt + 2\sqrt{Y_t^\pm}dB_t, \quad Y_0 = y, \tag{D.7}$$

for a standard Brownian motion $(B_t)_{t \geqslant 0}$. (Existence and uniqueness follow from Lemma D.1.)

**Lemma D.3.** *The diffusions $Y^+$, $Y^-$ are recurrent; their speed measures are finite and have the following densities:*

$$\nu^\pm(y) = \frac{1}{2}y^{-\frac{1}{2}}\exp\left(-2\frac{l \mp \delta}{m \mp \delta}G\left((m \mp \delta)\sqrt{y}\right) + 2\frac{l \mp \delta}{m \mp \delta}G(m \mp \delta)\right)\mathbb{1}_{\{y \geqslant 0\}}.$$

*Proof.* To prove that $Y^+$, $Y^-$ are recurrent, first note that this property only depends on the respective laws. Whence, it is enough to verify it for any weak solutions of the SDEs (D.7). Such solutions are given by the squares of the solutions of the SDE (3.7) with constant coefficients $L = l \pm \delta$, $M = m \pm \delta$, and without stopping. To prove recurrence of $Y^\pm$ on $\mathbb{R}_+$ it is in turn sufficient to verify recurrence of these solutions on $\mathbb{R}$, which follows from [22, Proposition 5.5.22(a)].

To compute the speed measures of $Y^+$, $Y^-$, first note that the respective scale functions (cf. [22, Equation (5.42)] are

$$p^\pm(y) = \int_1^y \exp\left(-2\int_1^x \frac{1 - 2(l \mp \delta)\sqrt{z}g\left((m \mp \delta)\sqrt{z}\right)}{4z}dz\right)dx$$

$$= \int_1^y x^{-\frac{1}{2}}\exp\left(2\frac{(l \mp \delta)}{(m \mp \delta)}G\left((m \mp \delta)\sqrt{x}\right) - 2\frac{(l \mp \delta)}{(m \mp \delta)}G(m \mp \delta)\right)dx, \quad \text{for } y \geqslant 0,$$

where $G(x) = \int_0^x g(y)dy$. The asserted formulas for the densities of the corresponding speed measures in turn follows directly from the definition [22, Equation (5.51)]. Finiteness follows from an elementary integration near zero and the growth condition (1.2) for the function $g$ near infinity. □

Lemma D.3, the ergodic theorem as in [7, Section II.35], the growth condition for the functions $f$ and $g$, and a change of variable in turn yield the following ergodic limits:

**Lemma D.4.** *Suppose $l, m > \delta > 0$. Then, for any $k \geqslant 0$:*

$$\lim_{x \to \infty} \frac{1}{x}\int_0^x f\left((k + \delta)\sqrt{Y_t^+}\right)dt = \frac{\int_{\mathbb{R}_+} f\left(\frac{(k+\delta)}{m-\delta}y\right)\exp\left(-2\frac{l-\delta}{m-\delta}G(y)\right)dy}{\int_{\mathbb{R}_+}\exp\left(-2\frac{l-\delta}{m-\delta}G(y)\right)dy} \quad a.s.$$

$$\lim_{x \to \infty} \frac{1}{x}\int_0^x f\left(k(1 - \delta)\sqrt{Y_t^-}\right)dt = \frac{\int_{\mathbb{R}_+} f\left(\frac{k(1-\delta)}{m+\delta}y\right)\exp\left(-2\frac{l+\delta}{m+\delta}G(y)\right)dy}{\int_{\mathbb{R}_+}\exp\left(-2\frac{l+\delta}{m+\delta}G(y)\right)dy} \quad a.s.$$

## D.3   A Result from Measure Theory

The following result from measure theory is used in the proof of Theorem 1.4:

**Lemma D.5.** *Let $(a_t^{\varepsilon,1})_{t\in[0,T]}$ be a family of product-measurable processes indexed by $\varepsilon \in (0, \frac{1}{2})$ and $(a_t^2)_{t\in[0,T]}$ a product measurable-process. Suppose that, for $\mathrm{Leb}_{|[0,T]}$-a.e. $t \in [0,T]$, $a_t^{\varepsilon,1}$ converges in probability to $a_t^2$ as $\varepsilon \to 0$. Moreover, assume that $|a^{\varepsilon,1}|^p$ is uniformly integrable (as a family indexed by $\varepsilon$) with respect to $\mathbb{P} \otimes \mathrm{Leb}_{|[0,T]}$ and that $a^2 \in L^p(\mathbb{P} \otimes \mathrm{Leb}_{|[0,T]})$ for some $p \geqslant 1$. Then $\int_0^{\cdot} a_s^{\varepsilon,1}ds \to \int_0^{\cdot} a_s^2 ds$ in $\mathcal{S}^p([0,T])$.*

*Proof.* First, we show that $a^{\varepsilon,1}$ converges to $a^2$ in measure under $\mathbb{P} \otimes \mathrm{Leb}_{|[0,T]}$. Indeed, for fixed $t \in [0,T]$, convergence in probability of $a_t^{\varepsilon,1}$ to $a_t^2$ is equivalent to

$$\mathbb{E}\left[\left|a_t^2 - a_t^{\varepsilon,1}\right| \wedge 1\right] \to 0 \ \text{ as } \ \varepsilon \to 0.$$

Thus, by Fubini's theorem and dominated convergence, we obtain

$$\mathbb{E}\left[\int_0^T \left(\left|a_t^2 - a_t^{\varepsilon,1}\right| \wedge 1\right)dt\right] \to 0 \ \text{ as } \ \varepsilon \to 0,$$

which is equivalent to convergence in measure under $\mathbb{P} \otimes \mathrm{Leb}_{|[0,T]}$ of $a^{\varepsilon,1}$ to $a^2$.

Next [21, Proposition 4.12] implies that $a^{\varepsilon,1} \to a^2$ in $L^p(\mathbb{P} \otimes \mathrm{Leb}_{|[0,T]})$. The assertion in turn follows from Jensen's inequality via

$$\limsup_{\varepsilon\to0} \mathbb{E}\left[\sup_{t\in[0,T]} \left|\int_0^t a_s^{\varepsilon,1}ds - \int_0^t a_s^2 ds\right|^p\right] \leqslant T^{p-1} \limsup_{\varepsilon\to0} \mathbb{E}\left[\int_0^T \left|a_s^{\varepsilon,1} - a_s^2\right|^p ds\right] = 0.$$

$\square$

## D.4 An Integrability Result

The following result is used in the reduction to bounded coefficients in Section 2.

**Lemma D.6.** *Suppose that Assumption 2 is satisfied. Then:*

$$\mathbb{E}\left[\int_0^T \left(H_s \frac{\int_{\mathbb{R}} f\left(\frac{K_s}{M_s}y\right)\exp\left(-2\frac{L_s}{M_s}G(y)\right)dy}{\int_{\mathbb{R}} \exp\left(-2\frac{L_s}{M_s}G(y)\right)dy}\right)^{p(1+\eta)} ds\right] < \infty. \tag{D.8}$$

*Proof.* We start by estimating the fraction appearing in (D.8). By the growth condition (1.3) for $f$ and the fact that $G$ is even (because $g$ is odd), it follows that

$$\frac{\int_{\mathbb{R}} f\left(\frac{K_s}{M_s}y\right)\exp\left(-2\frac{L_s}{M_s}G(y)\right)dy}{\int_{\mathbb{R}} \exp\left(-2\frac{L_s}{M_s}G(y)\right)dy} \leqslant C_f + C_f\left(\frac{K_s}{M_s}\right)^{q'}\frac{\int_0^{\infty} y^{q'}\exp\left(-2\frac{L_s}{M_s}G(y)\right)dy}{\int_0^{\infty} \exp\left(-2\frac{L_s}{M_s}G(y)\right)dy}. \tag{D.9}$$

We proceed to estimate the numerator and denominator in the fraction appearing in (D.9). For the numerator, we use that by the growth condition (1.2) of $g$, there are $\tilde{x}$ and $C > 0$ such that

$$G(x) \geqslant Cx, \quad \forall x \geqslant \tilde{x}.$$

Using this, we obtain

$$\int_0^\infty y^{q'} \exp\left(-2\frac{L_s}{M_s}G(y)\right) dy \leqslant \frac{\tilde{x}^{q'+1}}{1+q'} + \int_{\tilde{x}}^\infty y^{q'} \exp\left(-2\frac{L_s}{M_s}Cy\right) dy$$

$$\leqslant \frac{\tilde{x}^{q'+1}}{1+q'} + \int_0^\infty y^{q'} \exp\left(-2\frac{L_s}{M_s}Cy\right) dy$$

$$\leqslant \frac{\tilde{x}^{q'+1}}{1+q'} + \left(2C\frac{L_s}{M_s}\right)^{-q'-1} \int_0^\infty z^{q'} \exp\left(-z\right) dy$$

$$\leqslant \frac{\tilde{x}^{q'+1}}{1+q'} + \Gamma(q'+1)(2C)^{-q'-1}\left(\frac{M_s}{L_s}\right)^{q'+1}$$

$$\leqslant \tilde{C}\left(1 + \left(\frac{M_s}{L_s}\right)^{q'+1}\right), \tag{D.10}$$

for some constant $\tilde{C} > 0$. For the denominator in the fraction appearing in (D.9), we use that $G$ is continuous and nondecreasing on $\mathbb{R}_+$ with $G(0) = 0$ and $\lim_{x\to\infty} G(x) = \infty$. Thus, there is $c > 0$ such that $G(c) \leqslant \frac{\eta}{4p(1+\eta)}$. This gives

$$\int_0^\infty \exp\left(-2\frac{L_s}{M_s}G(y)\right) dy \geqslant \int_0^c \exp\left(-2\frac{L_s}{M_s}G(y)\right) dy \geqslant \int_0^c \exp\left(-2\frac{L_s}{M_s}G(c)\right) dy$$

$$= c\exp\left(-\frac{L_s}{M_s}G(c)\right) \geqslant c\exp\left(-\frac{L_s}{M_s}\frac{\eta}{4p(1+\eta)}\right). \tag{D.11}$$

It is an elementary exercise in analysis to show that there is a constant $A > 0$ such that

$$\left(1 + x^{-q'-1}\right) \exp\left(\frac{\eta}{4p(1+\eta)}x\right) \leqslant A\left(x^{-q'-1} + \exp\left(\frac{\eta}{4p(1+\eta)}x\right)\right), \quad x > 0.$$

Together with (D.10) and (D.11), it follows that there is a constant $\bar{C} \geqslant 1$ such that

$$\frac{\int_0^\infty y^{q'} \exp\left(-2\frac{L_s}{M_s}G(y)\right) dy}{\int_0^\infty \exp\left(-2\frac{L_s}{M_s}G(y)\right) dy} \leqslant \bar{C}\left(\left(\frac{M_s}{L_s}\right)^{q'+1} + \exp\left(\frac{L_s}{M_s}\frac{\eta}{4p(1+\eta)}\right)\right). \tag{D.12}$$

Now putting together (D.9) and (D.12), and using the elementary inequality $abc \leqslant a^2 + b^4 + c^4$ for $a, b, c > 0$, we obtain

$$H_s \frac{\int_{\mathbb{R}} f\left(\frac{K_s}{M_s}y\right) \exp\left(-2\frac{L_s}{M_s}G(y)\right) dy}{\int_{\mathbb{R}} \exp\left(2\frac{L_s}{M_s}G(y)\right) dy}$$

$$\leqslant C_f \bar{C}\left(H_s + H_s K_s^{q'}\left(\frac{1}{L_s}\right)^{q'}\frac{M_s}{L_s} + H_s K_s^{q'}\left(\frac{1}{M_s}\right)^{q'}\exp\left(\frac{L_s}{M_s}\frac{\eta}{4p(1+\eta)}\right)\right)$$

$$\leqslant C_f \bar{C}\left(H_s + H_s K_s^{q'}\left(\frac{1}{L_s \wedge M_s}\right)^{q'}\frac{M_s}{L_s} + H_s K_s^{q'}\left(\frac{1}{L_s \wedge M_s}\right)^{q'}\exp\left(\frac{L_s}{M_s}\frac{\eta}{4p(1+\eta)}\right)\right)$$

$$\leqslant C_f \bar{C}\left(H_s + 2\left(H_s K_s^{q'}\right)^2 + 2\left(\frac{1}{L_s \wedge M_s}\right)^{4q'} + \left(\frac{M_s}{L_s}\right)^4 + \exp\left(\frac{L_s}{M_s}\frac{\eta}{p(1+\eta)}\right)\right) \tag{D.13}$$

The claimed estimate (D.8) in turn follows from (D.13), Assumption 2 and the elementary inequality $(a+b+c+d)^{p(1+\eta)} \leqslant 4^{p(1+\eta)}\left(a^{p(1+\eta)} + b^{p(1+\eta)} + c^{p(1+\eta)} + d^{p(1+\eta)}\right)$ for $a, b, c, d \geqslant 0$. $\qquad\square$

35

# References

[1] L. Ahrens. *On using shadow prices for the asymptotic analysis of portfolio optimization under proportional transaction costs.* PhD thesis, Christan-Albrechts-Universität zu Kiel, 2015.

[2] R. F. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Appl. Math. Finance*, 10(1):1–18, 2003.

[3] R. F. Almgren and N. Chriss. Optimal execution of portfolio transactions. *J. Risk*, 3:5–40, 2001.

[4] R. F. Almgren and T. M. Li. Option hedging with smooth market impact. *Market Microstucture Liq.*, 2(1), 2016.

[5] R. F. Almgren, C. Thum, E. Hauptmann, and H. Li. Direct estimation of equity market impact. *RISK*, July, 2005.

[6] P. Bank, H. Soner, and M. Voss. Hedging with temporary price impact. *Math. Fin. Econ.*, 11(2):215–239, 2017.

[7] A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae.* Birkhäuser Verlag, Basel, second edition, 2002.

[8] J. Cai and M. Fukasawa. Asymptotic replication with modified volatility under small transaction costs. *Finance Stoch.*, 20(2):381–431, 2016.

[9] J. Cai, M. Rosenbaum, and P. Tankov. Asymptotic lower bounds for optimal tracking: a linear programming approach. *Ann. Appl. Probab.*, 27(4):2455–2514, 2017.

[10] J. Cai, M. Rosenbaum, and P. Tankov. Asymptotic optimal tracking: feedback strategies. *Stochastics*, 89(6–7):943–966, 2017.

[11] T. Cayé, M. Herdegen, and J. Muhle-Karbe. Trading with small nonlinear price impact. Preprint, 2017.

[12] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4):676–713, 1990.

[13] S. N. Ethier and T. G. Kurtz. *Markov processes.* John Wiley & Sons, Inc., New York, 1986.

[14] M. R. Feodoria. *Optimal investment and utility indifference pricing in the presence of small fixed transaction costs.* PhD thesis, Christian-Albrechts-Universität zu Kiel, 2016.

[15] N. Garleanu and L. H. Pedersen. Dynamic trading with predictable returns and transaction costs. *J. Finance*, 68(6):2309–2340, 2013.

[16] N. Garleanu and L. H. Pedersen. Dynamic portfolio choice with frictions. *J. Econ. Theory*, 165:487–516, 2016.

[17] P. Guasoni and M. Weber. Dynamic trading volume. *Math. Finance*, 27(2):313–349, 2017.

[18] P. Guasoni and M. Weber. Nonlinear price impact and portfolio choice. Preprint, 2018.

[19] M. Herdegen and J. Muhle-Karbe. Stability of Radner equilibria with respect to small frictions. *Finance Stoch.*, 22(2):443–502, 2018.

[20] K. Janeček and S. E. Shreve. Futures trading with transaction costs. *Illinois J. Math.*, 54(4):1239–1284, 2010.

[21] O. Kallenberg. *Foundations of modern probability.* Springer, Berlin, second edition, 2002.

[22] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus.* Springer, New York, second edition, 1991.

[23] R. Z. Khas'minskii. A limit theorem for the solutions of differential equations with random right-hand sides. *Theory Probab. Appl.*, 11(3):390–406, 1966.

[24] F. Lillo, J. D. Farmer, and R. N. Mantegna. Master curve for price-impact function. *Nature*, 421:129–130, 2003.

[25] L. Moreau, J. Muhle-Karbe, and H. M. Soner. Trading with small price impact. *Math. Finance*, 27(2):350–400, 2017.

[26] G. C. Papanicolaou, D. W. Stroock, and S. R. S. Varadhan. Martingale approach to some limit theorems. In D. Ruelle, editor, *Conference on Statistical Mechanics, Dynamical Systems and Turbulence.* Duke Univ. Press, 1977.

[27] E. Pardoux and A. Y. Veretennikov. On poisson equation and diffusion approximation 1. *Ann. Prob.*, 29(3):1061–1085, 2001.

[28] E. Pardoux and A. Y. Veretennikov. On poisson equation and diffusion approximation 2. *Ann. Prob.*, 31(3):1166–1192, 2003.

[29] G. Peskir. Bounding the maximal height of a diffusion by the time elapsed. *J. Theoret. Probab.*, 14(3), 2015.

[30] P. E. Protter. *Stochastic integration and differential equations.* Springer, Berlin, second edition, 2005.

[31] D. Revuz and M. Yor. *Continuous martingale and Brownian motion.* Springer, Berlin, third edition, 1999.

[32] S. E. Shreve. *Stochastic calculus for finance II: continuous-time models.* Springer, New York, 2004.

[33] A. V. Skorokhod. *Asymptotic methods in the theory of stochastic differential equations.* American Mathematical Society, Providence, RI, 1989.

[34] R. L. Stratonovich. *Topics in the theory of random noise. Vol. I.* Gordon and Breach, New York, 1963.

[35] R. L. Stratonovich. *Topics in the theory of random noise. Vol. II.* Gordon and Breach, New York, 1967.

[36] A. Y. Veretennikov. On polynomial mixing bounds for stochastic differential equations. *Stoch. Proc. and Appl.*, 70(1):115–128, 1997.