

MALLIAVIN CALCULUS APPROACH TO LONG EXIT TIMES FROM AN UNSTABLE EQUILIBRIUM

BY YURI BAKHTIN¹ AND ZSOLT PAJOR-GYULAI

New York University

For a one-dimensional smooth vector field in a neighborhood of an unstable equilibrium, we consider the associated dynamics perturbed by small noise. Using Malliavin calculus tools, we obtain precise vanishing noise asymptotics for the tail of the exit time and for the exit distribution conditioned on atypically long exits.

1. Introduction. Exit problems for small random perturbations of random dynamical systems have been studied for several decades. The most celebrated asymptotic results in this direction are large deviation estimates of the Freidlin–Wentzell theory and their extensions; see the classical book [9].

There are situations when large deviation results are not sufficient for detailed analysis of the system’s behavior. In particular, the analysis of noisy heteroclinic networks, that is, systems with multiple unstable equilibria connected to each other by heteroclinic orbits, requires studying distributional scaling limit theorems for exit points and exit times. This approach allows for an iteration scheme that leads to a detailed description of typical diffusion paths on time scales logarithmic in noise magnitude $\varepsilon > 0$; see [1, 3, 4]. In those papers, the results on the asymptotics of exits from neighborhoods of critical points extend the results of [11] and [7] where the leading deterministic logarithmic term for the exit time τ^ε and the leading random correction to the logarithmic term were computed. Namely, for a class of initial conditions near the critical point (or the associated stable manifold), it was established that if $\lambda > 0$ is the leading eigenvalue of the linearization of the dynamics near the critical point, then

$$\tau^\varepsilon = \frac{1}{\lambda} \log \frac{1}{\varepsilon} + \theta_\varepsilon,$$

where random variables θ_ε converge in distribution as $\varepsilon \rightarrow 0$. Moreover, the limiting distribution is nontrivial and has explicit representations. The associated distributions of exit locations were also studied in [8] and [2].

To extend the results of [1, 3, 4] to longer time scales, one needs to study rare events responsible for the unlikely transitions in the heteroclinic networks. One

Received October 2017; revised February 2018.

¹Supported in part by NSF Grant DMS-1460595.

MSC2010 subject classifications. 60H07, 60H10, 60J60.

Key words and phrases. Vanishing noise limit, unstable equilibrium, exit problem, polynomial decay, Malliavin calculus.

such rare event can be described as withstanding the repulsion near an unstable critical point for an atypically long time. Thus, we need to study the asymptotics of the tail of τ_ε . The best results in this direction, to the best of our knowledge, were established in [12], where a large deviation estimate in the form of logarithmic equivalence was obtained:

$$\lim_{\varepsilon \rightarrow 0} \frac{\log \mathbf{P}(\tau^\varepsilon > \frac{\alpha}{\lambda} \log \frac{1}{\varepsilon})}{\log \varepsilon} = \alpha - 1, \quad \alpha > 1.$$

In this paper, for the one-dimensional case, we provide more delicate estimates proving that for all $\alpha > 1$,

$$(1) \quad \mathbf{P}\left(\tau^\varepsilon > \frac{\alpha}{\lambda} \log \frac{1}{\varepsilon}\right) = \Lambda \varepsilon^{\alpha-1} (1 + o(1)), \quad \varepsilon \rightarrow 0,$$

and computing the precise value of the constant Λ ; see our main result, Theorem 1.1, below. We give an explicit expression for Λ in terms of the starting point ranging through a neighborhood of the critical point that we describe. We also explicitly compute the limiting distribution of the exit location conditioned on the rare event of withstanding the repulsion for an atypically long time, and it turns out that it does not depend on the starting point within that neighborhood.

It is essential for our proof to study the asymptotic behavior of densities of certain auxiliary random variables. Besides the traditional methods of stochastic analysis, we use the Malliavin calculus tools. We believe that our approach can be extended to higher dimensions, although the extension is not straightforward, and it will be addressed in a separate paper.

Let us now introduce the setting more formally. We will consider the family of stochastic differential equations

$$(2) \quad dX_\varepsilon(t) = b(X_\varepsilon(t)) dt + \varepsilon \sigma(X_\varepsilon(t)) dW(t),$$

on a bounded interval $\mathcal{I} = [q_-, q_+] \subseteq \mathbb{R}$, where the drift is given by a vector field $b \in \mathcal{C}^2(\mathbb{R})$ and the random perturbation is given via a standard Brownian motion W with respect to a filtration $(\mathcal{F}_t)_{t \geq 0}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The noise magnitude is given by a small parameter $\varepsilon > 0$ in front of the diffusion coefficient $\sigma \in \mathcal{C}^2(\mathbb{R})$ which is assumed to satisfy $\sigma(0) > 0$. Although we are interested only in the evolution within \mathcal{I} , we can assume that b and σ are globally Lipschitz without changing the setting.

Standard results on stochastic differential equations (see, e.g., [10]) imply that for any starting location $X^\varepsilon(0) \in \mathcal{I}$, equation (2) has a unique strong solution up to

$$\tau_{\mathcal{I}}^\varepsilon = \inf\{t \geq 0 : X_\varepsilon(t) \in \partial\mathcal{I}\},$$

the exit time from \mathcal{I} .

Let $(S^t)_{t \in \mathbb{R}}$ be the flow generated by the vector field b , that is, $x(t) = S^t x_0$ is the solution of the autonomous ordinary differential equation

$$(3) \quad \dot{x}(t) = b(x(t)), \quad x(0) = x_0 \in \mathbb{R}.$$

We assume that there is a unique repelling zero of the vector field b . Without loss of generality, we place it at the origin. In other words, we assume that $b(0) = 0$ and, for some $\lambda > 0$ and $\eta \in \mathcal{C}^2(\mathcal{I})$,

$$(4) \quad b(x) = \lambda x + \eta(x)|x|^2, \quad x \in \mathcal{I}.$$

Note that since the origin is the only zero of b in the closed interval \mathcal{I} , this assumption implies that for all $x \neq 0$, there is a uniquely defined finite time $T(x)$ such that $S^{T(x)} \in \partial\mathcal{I}$.

Under the condition (4), the map $f : \mathcal{I} \rightarrow \mathbb{R}$ defined by

$$(5) \quad f(x) = \lim_{t \rightarrow \infty} e^{\lambda t} S^{-t} x = x - \int_0^\infty e^{\lambda s} \eta(S^{-s} x) |S^{-s} x|^2 ds$$

is a \mathcal{C}^2 -diffeomorphism; see [8]. It preserves the order on \mathbb{R} , so

$$f(q_-) < 0 < f(q_+).$$

Our main result is the following.

THEOREM 1.1. *Consider X_ε defined by (2) with initial condition $X_\varepsilon(0) = \varepsilon x$ and let $K(\varepsilon)$ be a function such that*

$$(6) \quad \lim_{\varepsilon \downarrow 0} \varepsilon^\beta K(\varepsilon) = 0 \quad \forall \beta > 0.$$

Then, for all $\alpha > 1$,

$$(7) \quad \lim_{\varepsilon \downarrow 0} \sup_{|x| \leq K(\varepsilon)} \left| \varepsilon^{-(\alpha-1)} \mathbf{P} \left(\tau_{\mathcal{I}}^\varepsilon > \frac{\alpha}{\lambda} \log \varepsilon^{-1} \right) - \sqrt{\frac{\lambda}{\pi}} \frac{e^{-\lambda \left(\frac{x}{\sigma(0)}\right)^2}}{\sigma(0)} (|f(q_+)| + |f(q_-)|) \right| = 0$$

and

$$(8) \quad \lim_{\varepsilon \downarrow 0} \sup_{|x| \leq K(\varepsilon)} \left| \mathbf{P} \left(X_\varepsilon(\tau_{\mathcal{I}}^\varepsilon) = q_\pm \mid \tau_{\mathcal{I}}^\varepsilon > \frac{\alpha}{\lambda} \log \varepsilon^{-1} \right) - \frac{|f(q_\pm)|}{|f(q_+)| + |f(q_-)|} \right| = 0.$$

REMARK 1.2. The second term in (7) plays the role of Λ in (1). Note that in (7) the coefficient in front of $|f(q_+)| + |f(q_-)|$ is the centered Gaussian density with variance $\sigma^2(0)/(2\lambda)$ evaluated at x .

One could approach Theorem 1.1 using that the distribution of X_ε conditioned to stay in the bounded domain for an atypically long time, approaches the quasi-stationary distribution exponentially fast; see [6]. However, our situation is more subtle since both, the time scale and the system, depend on ε . So, instead of attempting to appeal to the general quasi-stationary distribution theory, we adopt the following plan: we change the coordinates by conjugating the drift to a linear one;

then we use Duhamel’s principle to represent the solution and express the event of interest $\{\tau_{\mathcal{I}}^\varepsilon > \frac{\alpha}{\lambda} \log \frac{1}{\varepsilon}\}$ in terms of a random factor in the resulting variation of constants formula; the convergence of that factor in distribution is known and has been employed in the literature cited above, but the desired result concerns unlikely events, and we need much stronger regularity, namely, uniform convergence of the associated probability densities. We use Malliavin calculus to study these densities. In fact, the Malliavin calculus approach works only for small values of α , and to extend it to the longer time scales we need to invoke an additional recursive scheme that can be seen as studying the quasi-stationary distribution since each step is performed under the no-exit conditioning.

In this program, the density estimates are based on a well-known formula for the density in terms of the Malliavin derivative and the divergence operator; see Proposition 4.1. This formula has been used to derive upper bounds on densities (see [14], Section 2.1.1), but in general it is viewed as not very useful (see, e.g., [13]) where a replacement formula is suggested. However, in our context, Proposition 4.1 turns out to be very efficient.

2. Proof of Theorem 1.1. We will study the system in a small neighborhood of the origin and after the process has escaped this small neighborhood.

Let us start with the first part. The diffeomorphism $f : \mathcal{I} \rightarrow \mathbb{R}$ introduced in (5) and its inverse $g = f^{-1}$ provide a conjugation between the flow (S^t) and a linear flow:

$$(9) \quad f(S^t x) = e^{\lambda t} f(x), \quad \text{or} \quad f'(x)b(x) = \lambda f(x).$$

Note that the integrand in (5) is quadratic when x is close to zero, and thus we have $f(0) = 0$ and $f'(0) = 1$. Outside of \mathcal{I} , we define f so that f' and f'' are bounded.

Let $Y_\varepsilon(t) = f(X_\varepsilon(t))$ for times prior to the escape from \mathcal{I} . Itô’s formula and (9) then imply that this process satisfies the stochastic differential equation

$$(10) \quad dY_\varepsilon(t) = \lambda Y_\varepsilon(t) dt + \varepsilon \tilde{\sigma}(Y_\varepsilon(t)) dW(t) + \frac{\varepsilon^2}{2} h(Y_\varepsilon(t)) dt$$

for $t < \tau_{\mathcal{I}}^\varepsilon$, where $\tilde{\sigma}(y) = f'(g(y))\sigma(g(y))$ and $h(y) = f''(g(y))\sigma^2(g(y))$. Due to boundedness of f' and f'' , $\tilde{\sigma}$ and h are also bounded.

Let us choose $R > 0$ sufficiently small to ensure that $\mathcal{V} = g([-R, R]) \subseteq \mathcal{I}$ and $\tilde{\sigma}(x) > 0$ for all $x \in [-R, R]$. The following result describes the behavior of $\tau_{\mathcal{V}}^\varepsilon$, the exit time from \mathcal{V} .

THEOREM 2.1. *Let $Y_\varepsilon(0) = \varepsilon x$, where $|x| \leq K(\varepsilon)$ with $K(\varepsilon)$ satisfying (6). There is a family of random variables (M_ε) such that for all $\varepsilon > 0$, $\mathbf{P}(M_\varepsilon = 0) = 0$,*

$$(11) \quad \tau_{\mathcal{V}}^\varepsilon = \frac{1}{\lambda} \log \frac{R}{\varepsilon} - \frac{1}{\lambda} \log |M_\varepsilon|,$$

$$(12) \quad \mathbf{P}(\{Y_\varepsilon(\tau_{\mathcal{V}}^\varepsilon) = \pm R\} \Delta \{\pm M_\varepsilon > 0\}) = 0,$$

and, if $a(\varepsilon)$ is a monotone function such that $a(\varepsilon) \sim c\varepsilon^\theta, \varepsilon \rightarrow 0$, for some $c, \theta > 0$, then the following estimate holds:

$$(13) \quad \lim_{\varepsilon \downarrow 0} \sup_{|x| \leq K(\varepsilon)} \left| \mathbf{P}(0 < \pm M_\varepsilon \leq a(\varepsilon)) - \sqrt{\frac{\lambda}{\pi}} \frac{e^{-\lambda(\frac{x}{\sigma(0)})^2}}{\sigma(0)} a(\varepsilon) \right| = o(a(\varepsilon)).$$

REMARK 2.2. The coefficient in front of $a(\varepsilon)$ in (13) is the centered Gaussian density with variance $\sigma^2(0)/(2\lambda)$ evaluated at x .

We give the proof of Theorem 2.1 in Section 3.

After exiting from \mathcal{V} , the deterministic dynamics dominates the evolution, which is captured by the following standard large deviation estimates.

PROPOSITION 2.3. Let $X_\varepsilon(0) = g(\pm R)$. There are constants $c_1, c_2, \varepsilon_0 > 0$ such that

$$(14) \quad \mathbf{P}(|\tau_{\mathcal{I}}^\varepsilon - T(g(\pm R))| > \varepsilon N) \leq c_1 e^{-c_2 N^2}, \quad \varepsilon < \varepsilon_0, N \geq 1,$$

and

$$\mathbf{P}(X_\varepsilon(\tau_{\mathcal{I}}) = q_\pm) \geq 1 - c_1 e^{-c_2/\varepsilon^2}, \quad \varepsilon < \varepsilon_0.$$

PROOF OF THEOREM 1.1. Combining $\tau_{\mathcal{I}}^\varepsilon = \tau_{\mathcal{V}}^\varepsilon + (\tau_{\mathcal{I}}^\varepsilon - \tau_{\mathcal{V}}^\varepsilon)$ with (11), (14) and the strong Markov property yields the representation

$$(15) \quad \tau_{\mathcal{I}}^\varepsilon = \frac{1}{\lambda} \log \varepsilon^{-1} + C - \frac{1}{\lambda} \log |M_\varepsilon| + \theta_\varepsilon, \quad C = \frac{1}{\lambda} \log R + T(X_\varepsilon(\tau_{\mathcal{V}}^\varepsilon)),$$

where θ_ε is a random variable such that

$$(16) \quad \mathbf{P}(|\theta_\varepsilon| > \varepsilon N) \leq c_1 e^{-c_2 N^2}, \quad N \geq 1,$$

for some $c_1, c_2 > 0$. Although R appears in the definition of C , one can easily show that C does not, in fact, depend on the choice of R . Identities (15) and (12) imply

$$(17) \quad \mathbf{P}\left(\tau_{\mathcal{I}}^\varepsilon \geq \frac{\alpha}{\lambda} \log \varepsilon^{-1}\right) = \mathbf{P}(|M_\varepsilon| \leq e^{\lambda(C+\theta_\varepsilon)} \varepsilon^{\alpha-1}) = I_- + I_+,$$

where

$$(18) \quad I_\pm = \mathbf{P}(0 < \pm M_\varepsilon \leq e^{\lambda(C^\pm + \theta_\varepsilon)} \varepsilon^{\alpha-1}), \quad C^\pm = \frac{1}{\lambda} \log R + T(g(\pm R)).$$

Let us simplify the definition of C^\pm . Since $T(g(\pm R))$ equals the time it takes for the linear flow to travel between R and $f(q_\pm)$, that is, $\frac{1}{\lambda} \log \frac{|f(q_\pm)|}{R}$, we see that

$$(19) \quad C^\pm = \frac{1}{\lambda} \log |f(q_\pm)|.$$

We can write

$$I_{\pm} \leq \mathbf{P}(0 < \pm M_{\varepsilon} \leq e^{\lambda(C^{\pm} + \varepsilon^{\beta})} \varepsilon^{\alpha-1}) + \mathbf{P}(|\theta_{\varepsilon}| > \varepsilon^{\beta}),$$

where we choose $\beta \in (0, 1)$. The second term decays exponentially fast as $\varepsilon \rightarrow 0$ by (16) and we may apply (13) to the first term with $a(\varepsilon) = e^{\lambda(C^{\pm} + \varepsilon^{\beta})} \varepsilon^{\alpha-1}$ to conclude

$$(20) \quad I_{\pm} \leq \sqrt{\frac{\lambda}{\pi}} \frac{e^{\lambda C^{\pm}} \varepsilon^{\alpha-1}}{\sigma(0)} e^{-\lambda(\frac{x}{\sigma(0)})^2} (1 + \mathcal{O}(\varepsilon^{\beta})) + o(\varepsilon^{\alpha-1}),$$

with the error term being uniform over $|x| \leq K(\varepsilon)$. Note that invoking (13) was justified by

$$Y_{\varepsilon}(0) = f(\varepsilon x) = \varepsilon x + o(\varepsilon x) = \varepsilon x(1 + o(1)).$$

The analogous lower bound follows similarly:

$$(21) \quad \begin{aligned} I_{\pm} &\geq \mathbf{P}(0 < \pm M_{\varepsilon} \leq e^{\lambda(C^{\pm} - \varepsilon^{\beta})} \varepsilon^{\alpha-1}; |\theta_{\varepsilon}| \leq \varepsilon^{\beta}) \\ &\geq \mathbf{P}(0 < \pm M_{\varepsilon} \leq e^{\lambda(C^{\pm} - \varepsilon^{\beta})} \varepsilon^{\alpha-1}) - \mathbf{P}(|\theta_{\varepsilon}| \geq \varepsilon^{\beta}) \\ &= \sqrt{\frac{\lambda}{\pi}} \frac{e^{\lambda C^{\pm}} \varepsilon^{\alpha-1}}{\sigma(0)} e^{-\lambda(\frac{x}{\sigma(0)})^2} (1 + \mathcal{O}(\varepsilon^{\beta})) + o(\varepsilon^{\alpha-1}). \end{aligned}$$

Combining (20), (21) and (17), we obtain

$$\varepsilon^{-(\alpha-1)} \mathbf{P}\left(\tau_{\mathcal{I}}^{\varepsilon} \geq \frac{\alpha}{\lambda} \log \varepsilon^{-1}\right) = \sqrt{\frac{\lambda}{\pi}} \frac{e^{\lambda C^{-}} + e^{\lambda C^{+}}}{\sigma(0)} e^{-\lambda(\frac{x}{\sigma(0)})^2} + o(1),$$

with the error term being uniform over $|x| \leq K(\varepsilon)$. Using (19), we complete the proof of (7).

To prove (8), we write

$$\begin{aligned} &\mathbf{P}(X_{\varepsilon}(\tau_{\mathcal{I}}^{\varepsilon}) = q_{\pm}; \lambda \tau_{\mathcal{I}}^{\varepsilon} > (1 - \alpha) \log \varepsilon^{-1}) \\ &= \mathbf{P}(X_{\varepsilon}(\tau_{\mathcal{I}}^{\varepsilon}) = q_{\pm}; |M_{\varepsilon}| \leq e^{\lambda(C + \theta_{\varepsilon})} \varepsilon^{\alpha-1}) \\ &= \mathbf{P}(X_{\varepsilon}(\tau_{\mathcal{I}}^{\varepsilon}) = g(\pm R); |M_{\varepsilon}| \leq e^{\lambda(C + \theta_{\varepsilon})} \varepsilon^{\alpha-1}) + \mathcal{O}(e^{-c_2/\varepsilon^2}) \\ &= \mathbf{P}(0 < \pm M_{\varepsilon} \leq e^{\lambda(C^{\pm} + \theta_{\varepsilon})} \varepsilon^{\alpha-1}) + \mathcal{O}(e^{-c_2/\varepsilon^2}) \\ &= I_{\pm} + \mathcal{O}(e^{-c_2/\varepsilon^2}). \end{aligned}$$

Here, the first equality is due to (15), the second one is due to Proposition 2.3, while third one holds by (12). Therefore,

$$\mathbf{P}\left(X_{\varepsilon}(\tau_{\mathcal{I}}^{\varepsilon}) = q_{\pm} \mid \tau_{\mathcal{I}} > \frac{\alpha}{\lambda} \log \varepsilon^{-1}\right) = \frac{I_{\pm} + \mathcal{O}(e^{-c_2/\varepsilon^2})}{I_{-} + I_{+}} \longrightarrow \frac{e^{\lambda C^{\pm}}}{e^{\lambda C^{-}} + e^{\lambda C^{+}}},$$

as $\varepsilon \rightarrow 0$ and the proof is completed by (19)–(21). \square

3. The linear equation. The goal of this section is to prove Theorem 2.1. To this end, we apply Duhamel’s formula to the stochastic differential equation (10) to obtain

$$(22) \quad Y_\varepsilon(t) = \varepsilon e^{\lambda t} M_\varepsilon(t), \quad M_\varepsilon(t) = x + U_\varepsilon(t) + V_\varepsilon(t),$$

where we used $Y_\varepsilon(0) = \varepsilon x$ and introduced

$$U_\varepsilon(t) = \int_0^t e^{-\lambda s} \tilde{\sigma}(Y_\varepsilon(s)) dW(s), \quad V_\varepsilon(t) = \frac{\varepsilon}{2} \int_0^t e^{-\lambda s} h(Y_\varepsilon(s)) ds.$$

Note that there is $C > 0$ such that for all $t \geq 0$ and $x \in \mathbb{R}$,

$$(23) \quad \langle U_\varepsilon \rangle_t < C, \quad |V_\varepsilon(t)| \leq C\varepsilon.$$

By the definition of τ_V^ε , we have

$$(24) \quad R = |Y_\varepsilon(\tau_V^\varepsilon)| = \varepsilon e^{\lambda \tau_V^\varepsilon} |M_\varepsilon(\tau_V^\varepsilon)|,$$

which we can rearrange to obtain

$$\tau_V^\varepsilon = \frac{1}{\lambda} \log \frac{R}{\varepsilon} - \frac{1}{\lambda} \log |M_\varepsilon(\tau_V^\varepsilon)|.$$

This establishes (11) once we set $M_\varepsilon = M_\varepsilon(\tau_V^\varepsilon)$. Note immediately that

$$\mathbf{P}(M_\varepsilon = 0) = \mathbf{P}(\tau_T^\varepsilon = \infty) = 0$$

by the uniform ellipticity of $\tilde{\sigma}$. Moreover, the sign of $Y_\varepsilon(t)$ coincides with the sign of $M_\varepsilon(t)$ for all $t > 0$, and thus (12) is verified as well. Therefore, it remains to prove (13), and the rest of the section is dedicated to this goal.

The strategy of proof is the following. We first replace the stopping time τ_V^ε with the deterministic time

$$(25) \quad T_\varepsilon = \frac{1}{\lambda} \log \frac{R}{\varepsilon} - \frac{1}{\lambda} \log a(\varepsilon),$$

where we recall that $a(\varepsilon) \sim c\varepsilon^\theta$, which turns out to be a good substitute on the rare events that we seek to study. Next, we use tools from Malliavin calculus to establish that for an appropriately chosen shorter fixed time $T'_\varepsilon \leq T_\varepsilon$, the random variable $M_\varepsilon(T'_\varepsilon)$ has a density around zero that converges to the density of a centered Gaussian random variable with variance $(2\lambda)^{-1} \sigma^2(0)$, which establishes the desired estimate for $M_\varepsilon(T'_\varepsilon)$. Finally, we use the Markov property to iteratively extend this conclusion to $M_\varepsilon(T_\varepsilon)$. The following well-known exponential martingale inequality (see, e.g., Problem 12.10 in [5]) will be useful many times.

LEMMA 3.1. *Let $M(t)$ be a martingale with quadratic variation process $\langle M \rangle_t$. Then*

$$\mathbf{P}\left(\sup_{t \geq 0} |M(t)| \geq a; \langle M \rangle_\infty \leq b\right) \leq 2e^{-\frac{a^2}{2b}}$$

for any $a, b > 0$. In particular, due to (23),

$$\mathbf{P}\left(\sup_{t \geq 0} |U_\varepsilon(t)| \geq a\right) \leq 2e^{-Ca^2}, \quad \mathbf{E}\left[\sup_{t \geq 0} |U_\varepsilon(t)|^p\right] \leq C_p, \quad p \geq 1,$$

for some constants $C, C_p < \infty$.

We start by showing that $M_\varepsilon(T_\varepsilon)$ is a good substitute for $M_\varepsilon = M_\varepsilon(\tau_\gamma^\varepsilon)$ on the set $|M_\varepsilon| \leq a(\varepsilon)$. When working with deterministic times, we do not restrict the dynamics to times before the exit from \mathcal{I} , so we use our global Lipschitzness assumption to ensure the existence of the strong solution. From now on, we use the same letter C to denote various positive constants.

LEMMA 3.2. *For every $\gamma \in (0, 1)$ and $\gamma' > 0$, there is $\varepsilon_0 > 0$ such that*

$$\sup_{|x| \leq K(\varepsilon)} \mathbf{P}(|M_\varepsilon - M_\varepsilon(T_\varepsilon)| \geq a(\varepsilon)\varepsilon^\gamma; |M_\varepsilon| \leq a(\varepsilon)) < \varepsilon^{\gamma'}, \quad \varepsilon < \varepsilon_0.$$

PROOF. Observe that $\{|M_\varepsilon| \leq a(\varepsilon)\} = \{\tau_\gamma^\varepsilon \geq T_\varepsilon\}$ by (24) and (25). This allows us to write

$$\begin{aligned} & \mathbf{P}\left(|U_\varepsilon(\tau_\gamma^\varepsilon) - U_\varepsilon(T_\varepsilon)| \geq \frac{a(\varepsilon)\varepsilon^\gamma}{2}; |M_\varepsilon| \leq a(\varepsilon)\right) \\ &= \mathbf{P}\left(\left|\int_{T_\varepsilon}^{\tau_\gamma^\varepsilon} e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(s)) dW(s)\right| \geq \frac{a(\varepsilon)\varepsilon^\gamma}{2}; \tau_\gamma^\varepsilon \geq T_\varepsilon\right) \\ &\leq \mathbf{P}\left(\sup_{t \geq T_\varepsilon} \left|\int_{T_\varepsilon}^t e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(s)) dW(s)\right| \geq \frac{a(\varepsilon)\varepsilon^\gamma}{2}\right) \\ &\leq 2 \exp\left\{-C \frac{a^2(\varepsilon)\varepsilon^{2\gamma}}{e^{-2\lambda T_\varepsilon}}\right\} = 2 \exp\left\{-C \frac{a^2(\varepsilon)\varepsilon^{2\gamma}}{e^{-2(\log \frac{R}{\varepsilon} - \log a(\varepsilon))}}\right\} \\ &= 2 \exp\left\{-C \frac{\varepsilon^{2\gamma}}{\varepsilon^2}\right\} \leq \frac{\varepsilon^{\gamma'}}{2}, \end{aligned}$$

for any $\gamma' > 0$, and sufficiently small ε , where we use the boundedness of $\tilde{\sigma}$ and Lemma 3.1 in the last inequality. At the same time, on $\{\tau_\gamma^\varepsilon \geq T_\varepsilon\}$, we have

$$|V_\varepsilon(\tau_\gamma^\varepsilon) - V_\varepsilon(T_\varepsilon)| \leq \frac{\varepsilon}{2\lambda} \sup_{t \leq \tau_\gamma^\varepsilon} |a(Y_\varepsilon(t))| e^{-\lambda T_\varepsilon} \leq \frac{\varepsilon^2 a(\varepsilon) \|h\|_\infty}{2\lambda R},$$

where we used that h is bounded. Therefore, the triangle inequality and a simple union bound gives

$$\begin{aligned} & \mathbf{P}(|M_\varepsilon - M_\varepsilon(T_\varepsilon)| \geq a(\varepsilon)\varepsilon^\gamma; |M_\varepsilon| \leq a(\varepsilon)) \\ &\leq \mathbf{P}\left(|U_\varepsilon(\tau_\gamma^\varepsilon) - U_\varepsilon(T_\varepsilon)| \geq \frac{a(\varepsilon)\varepsilon^\gamma}{2}; |M_\varepsilon| \leq a(\varepsilon)\right) \end{aligned}$$

$$\begin{aligned}
 &+ \mathbf{P}\left(\left|V_\varepsilon(\tau_V^\varepsilon) - V_\varepsilon(T_\varepsilon)\right| \geq \frac{a(\varepsilon)\varepsilon^\gamma}{2}; |M_\varepsilon| \leq a(\varepsilon)\right) \\
 &\leq \varepsilon^{\gamma'}
 \end{aligned}$$

provided $\gamma \in (0, 1)$ and ε is sufficiently small. \square

Our next goal is to show that $M_\varepsilon(T_\varepsilon)$ satisfies the desired small ball asymptotics. The next proposition is the key technical result of this paper and will be proved in Section 4 with Malliavin calculus tools.

PROPOSITION 3.3. *Assume $Y_\varepsilon(0) = \varepsilon x$ and let $T'_\varepsilon > 0$ be a function of ε such that*

$$(26) \quad \frac{\lambda T'_\varepsilon}{\log \varepsilon^{-1}} \in \left[1 - \frac{c}{\log \varepsilon^{-1}}, 2 - \kappa\right]$$

for some $c, \kappa > 0$ and for all sufficiently small ε . The random variable $M'_\varepsilon = M_\varepsilon(T'_\varepsilon)$ has a continuous, bounded density $p^x_\varepsilon(z)$. Moreover, for every $K(\varepsilon)$ satisfying (6),

$$\lim_{\varepsilon \downarrow 0} \sup_{|x| \leq K(\varepsilon)} \sup_{z \in \mathbb{R}} |p^x_\varepsilon(z) - p^x(z)| = 0,$$

where p^x is the density of a Gaussian random variable with mean x and variance $(2\lambda)^{-1}\sigma^2(0)$:

$$p^x(z) = \sqrt{\frac{\lambda}{\pi \sigma(0)}} e^{-\lambda\left(\frac{z-x}{\sigma(0)}\right)^2}.$$

In the proof of the theorem on the outcome of the linear evolution, we will use the Markov property many times, so it is convenient to denote by \mathbf{P}_y the joint distribution of the driving Wiener process W and the process Y_ε with initial point $Y_\varepsilon(0) = y$.

PROOF OF THEOREM 2.1. As observed in the beginning of this section, it only remains to show that M_ε satisfies (13). Let us derive the theorem from the following statement that we will prove later: if $b(\varepsilon) = o(a(\varepsilon))$ and $K(\varepsilon)$ satisfies (6), then $P(\varepsilon, x) = \mathbf{P}_{\varepsilon x}(b(\varepsilon) \leq M_\varepsilon(T_\varepsilon) \leq a(\varepsilon))$ satisfies

$$(27) \quad \lim_{\varepsilon \rightarrow 0} \sup_{|x| \leq K(\varepsilon)} \left| \frac{P(\varepsilon, x)}{a(\varepsilon)} - p^x(0) \right| = 0.$$

Let us choose γ' large enough such that $\varepsilon^{\gamma'} = o(a(\varepsilon))$, $\varepsilon \rightarrow 0$. Lemma 3.2 and (27) imply that if $\gamma \in (0, 1)$, then

$$(28) \quad \begin{aligned}
 \mathbf{P}_{\varepsilon x}(0 < M_\varepsilon \leq a(\varepsilon)) &\leq \varepsilon^{\gamma'} + \mathbf{P}_{\varepsilon x}(0 < M_\varepsilon \leq a(\varepsilon); |M_\varepsilon - M_\varepsilon(T_\varepsilon)| < a(\varepsilon)\varepsilon^\gamma) \\
 &\leq \varepsilon^{\gamma'} + \mathbf{P}_{\varepsilon x}(-a(\varepsilon)\varepsilon^\gamma \leq M_\varepsilon(T_\varepsilon) \leq a(\varepsilon)(1 + \varepsilon^\gamma))
 \end{aligned}$$

$$(29) \quad = p^x(0)a(\varepsilon) + o(a(\varepsilon)).$$

Similarly,

$$\begin{aligned}
 & \mathbf{P}_{\varepsilon x}(0 < M_\varepsilon \leq a(\varepsilon)) \\
 & \geq \mathbf{P}_{\varepsilon x}(0 < M_\varepsilon \leq a(\varepsilon); |M_\varepsilon - M_\varepsilon(T_\varepsilon)| < a(\varepsilon)\varepsilon^\gamma) \\
 (30) \quad & \geq \mathbf{P}_{\varepsilon x}(a(\varepsilon)\varepsilon^\gamma \leq M_\varepsilon(T_\varepsilon) \leq a(\varepsilon)(1 - \varepsilon^\gamma); |M_\varepsilon - M_\varepsilon(T_\varepsilon)| < a(\varepsilon)\varepsilon^\gamma) \\
 & = \mathbf{P}_{\varepsilon x}(a(\varepsilon)\varepsilon^\gamma \leq M_\varepsilon(T_\varepsilon) \leq a(\varepsilon)(1 - \varepsilon^\gamma)) - \Delta(\varepsilon, x), \\
 & = p^x(0)a(\varepsilon) + o(a(\varepsilon)) - \Delta(\varepsilon, x),
 \end{aligned}$$

where

$$\begin{aligned}
 \Delta(\varepsilon, x) &= \mathbf{P}_{\varepsilon x}(a(\varepsilon)\varepsilon^\gamma \leq M_\varepsilon(T_\varepsilon) \leq a(\varepsilon)(1 - \varepsilon^\gamma); |M_\varepsilon - M_\varepsilon(T_\varepsilon)| \geq a(\varepsilon)\varepsilon^\gamma) \\
 &\leq \Delta_1(\varepsilon, x) + \Delta_2(\varepsilon, x),
 \end{aligned}$$

with

$$\begin{aligned}
 \Delta_1(\varepsilon, x) &= \mathbf{P}_{\varepsilon x}(\tau_{\mathcal{V}}^\varepsilon \geq T_\varepsilon; |M_\varepsilon - M_\varepsilon(T_\varepsilon)| \geq a(\varepsilon)\varepsilon^\gamma), \\
 \Delta_2(\varepsilon, x) &= \mathbf{P}_{\varepsilon x}(\tau_{\mathcal{V}}^\varepsilon \leq T_\varepsilon; |M_\varepsilon(T_\varepsilon)| \leq a(\varepsilon)(1 - \varepsilon^\gamma)).
 \end{aligned}$$

Due to (28) and (30), the desired relation (13) for the positive sign follows from

$$(31) \quad \Delta_i(\varepsilon, x) = o(a(\varepsilon)), \quad i = 1, 2,$$

uniformly in x . Since $\{\tau_{\mathcal{V}}^\varepsilon \geq T_\varepsilon\} = \{|M_\varepsilon| \leq a(\varepsilon)\}$, Lemma 3.2 immediately implies that $\Delta_1(\varepsilon, x) < \varepsilon^{\gamma'}$ for any $\gamma' > 0$ and sufficiently small ε . To estimate $\Delta_2(\varepsilon, x)$, we note that $|M_\varepsilon(T_\varepsilon)| \leq a(\varepsilon)(1 - \varepsilon^\gamma)$ implies

$$|Y_\varepsilon(T_\varepsilon)| = \varepsilon e^{\lambda T_\varepsilon} |M_\varepsilon(T_\varepsilon)| \leq R(1 - \varepsilon^\gamma),$$

and applying the strong Markov property to the process Y_ε after $\tau_{\mathcal{V}}^\varepsilon$, we obtain

$$\Delta_2(\varepsilon, x) \leq \max_{z=\pm R} \mathbf{P}_z\left(\inf_{t \in [0, T_\varepsilon]} |Y_\varepsilon(t)| \leq R(1 - \varepsilon^\gamma)\right).$$

Duhamel’s formula (22) and the reverse triangle inequality imply

$$|Y_\varepsilon(t)| \geq e^{\lambda t} \left| |Y_\varepsilon(0)| - \varepsilon |U_\varepsilon(t)| - \varepsilon |V_\varepsilon(t)| \right| \geq |Y_\varepsilon(0)| - \varepsilon |U_\varepsilon(t)| - \varepsilon |V_\varepsilon(t)|,$$

and, combining the last two displays, (23), and Lemma 3.1, we obtain

$$\Delta_2(\varepsilon, x) \leq \mathbf{P}\left(\sup_{t \in [0, T_\varepsilon]} |U_\varepsilon(t)| \geq R\varepsilon^{-(1-\gamma)} - C\right) \leq 3e^{-\frac{C}{\varepsilon^{2(1-\gamma)}}} < \varepsilon^{\gamma'},$$

for any $\gamma < 1$ and sufficiently small ε . Thus, (31) is verified, completing the proof of (13) for the positive sign. The result for $-M_\varepsilon$ is proved the same way.

It remains to prove (27). If $\theta \in (0, 1)$, then $T_\varepsilon = T'_\varepsilon$ satisfies (26), so (27) immediately follows from Lemma 3.2 and Proposition 3.3. If $\theta \geq 1$, then we cannot apply Proposition 3.3 directly, so our strategy will be to extend it to longer times using an iterative procedure based on the Markov property. Namely, we let

$N = \lfloor \theta \rfloor + 1$ and consider the shorter time $T'_\varepsilon = T_\varepsilon/N$. There is $\kappa > 0$ such that for sufficiently small ε ,

$$\alpha_\varepsilon := \frac{\lambda T'_\varepsilon}{\log \varepsilon^{-1}} = \frac{1 + \theta - \zeta_\varepsilon}{\lfloor \theta \rfloor + 1} \in \left(1 - \frac{C}{\log \varepsilon^{-1}}, 2 - \kappa \right),$$

where $\zeta_\varepsilon = \mathcal{O}(1/\log \varepsilon^{-1})$. We recall that $C > 0$ stands for a constant that may change on every appearance. In fact, observe also that the last display and $N \in \mathbb{N}$ imply

$$(32) \quad 2 \leq N \leq \frac{\theta}{\alpha_\varepsilon} + 1 + \frac{C}{\log \varepsilon^{-1}}$$

for sufficiently small ε . For simpler notation, let us use the abbreviations

$$T_{\varepsilon,k} = kT'_\varepsilon, \quad M_{\varepsilon,k} = M_\varepsilon(T_{\varepsilon,k}), \quad Y_{\varepsilon,k} = Y_\varepsilon(T_{\varepsilon,k}), \quad \mathcal{F}_{\varepsilon,k} = \mathcal{F}_{T_{\varepsilon,k}}$$

for $k = 1, \dots, N$. We will show by induction that for each $\varepsilon > 0$ there is a sequence $(H_{\varepsilon,k})_{k=0}^\infty$ of centered Gaussian random variables independent of $\mathcal{F}_{\varepsilon,k}$, with variance

$$(33) \quad \mathbf{E}H_{\varepsilon,k}^2 = \frac{\sigma^2(0)}{2\lambda} \frac{1 - \varepsilon^{2(N-k)\alpha_\varepsilon}}{1 - \varepsilon^{2\alpha_\varepsilon}},$$

and such that

$$(34) \quad P(\varepsilon, x) = \mathbf{P}_{\varepsilon x}(b(\varepsilon) \leq M_{\varepsilon,k} + \varepsilon^{k\alpha_\varepsilon} H_{\varepsilon,k} \leq a(\varepsilon)) + o(a(\varepsilon)), \quad \varepsilon \downarrow 0,$$

for all $k = 0, \dots, N$. Using (34) with $k = 0$ completes the proof of (27) as $M_{\varepsilon,0} = x$ and the random variable $H_{\varepsilon,0}$ is centered Gaussian with variance converging to $(2\lambda)^{-1}\sigma^2(0)$ as $\varepsilon \downarrow 0$ by (33).

Now we proceed with the proof of (34) by first noting that the case $k = N$ is trivial as $H_{\varepsilon,N} = 0$. Let us assume that (34) holds for some $k = 1, \dots, N$ and show that it therefore holds for $k - 1$ as well. Due to the relation $Y_\varepsilon(t) = \varepsilon e^{\lambda t} M_\varepsilon(t)$, the Markov property allows us to write

$$\begin{aligned} \mathbf{P}_{\varepsilon x}(b(\varepsilon) \leq M_{\varepsilon,k} + \varepsilon^{k\alpha_\varepsilon} H_{\varepsilon,k} \leq a(\varepsilon)) &= \mathbf{E}_{\varepsilon x}[\mathbf{P}_{\varepsilon x}(b(\varepsilon) \leq M_{\varepsilon,k} + \varepsilon^{k\alpha_\varepsilon} H_{\varepsilon,k} \leq a(\varepsilon) \mid \mathcal{F}_{T_{\varepsilon,k-1}})] \\ &= \int_{-\infty}^\infty \mathbf{P}_{\varepsilon x}(b(\varepsilon) \leq M_{\varepsilon,k} + \varepsilon^{k\alpha_\varepsilon} H_{\varepsilon,k} \leq a(\varepsilon) \mid M_{\varepsilon,k-1} = z) r_x^{(k-1)}(dz), \end{aligned}$$

where we introduced the measure

$$r_x^{(k-1)}(dz) = \mathbf{P}_{\varepsilon x}(M_{\varepsilon,k-1} \in dz).$$

Since $Y_{\varepsilon,k} = \varepsilon e^{\lambda T_{\varepsilon,k}} M_{\varepsilon,k} = \varepsilon^{1-k\alpha_\varepsilon} M_{\varepsilon,k}$, for all $k = 0, \dots, N$, the integrand equals

$$\begin{aligned} \mathbf{P}_{\varepsilon x}(\varepsilon^{1-k\alpha_\varepsilon} b(\varepsilon) \leq Y_{\varepsilon,k} + \varepsilon H_{\varepsilon,k} \leq \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon) \mid Y_{\varepsilon,k-1} = \varepsilon^{1-(k-1)\alpha_\varepsilon} z) \\ = \mathbf{P}_{\varepsilon^{1-(k-1)\alpha_\varepsilon} z}(\varepsilon^{1-k\alpha_\varepsilon} b(\varepsilon) \leq Y_{\varepsilon,1} + \varepsilon H_{\varepsilon,k} \leq \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon)), \end{aligned}$$

where we used the Markov property and the independence of $H_{\varepsilon,k}$ and $\mathcal{F}_{\varepsilon,k}$ in the last step. Note that $Y_{\varepsilon,1}$ is independent of $H_{\varepsilon,k}$. Combining the last three displays and the change of variables $z \rightarrow \varepsilon^{(k-1)\alpha_\varepsilon} z$ gives

$$\begin{aligned}
 \mathbf{P}_{\varepsilon,x}(b(\varepsilon) \leq M_{\varepsilon,k} + \varepsilon^{k\alpha_\varepsilon} H_{\varepsilon,k} \leq a(\varepsilon)) \\
 (35) \quad &= \int_{-\infty}^{\infty} \mathbf{P}_{\varepsilon,z}(\varepsilon^{1-k\alpha_\varepsilon} b(\varepsilon) \leq Y_{\varepsilon,1} + \varepsilon H_{\varepsilon,k} \leq \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon)) \\
 &\quad \times r_x^{(k-1)}(\varepsilon^{(k-1)\alpha_\varepsilon} dz).
 \end{aligned}$$

Let us now prove that there are $c_1, c_2 > 0$ such that for any $L > 0$, we have

$$(36) \quad \sup_{|z| \geq L} \mathbf{P}_{\varepsilon,z}(\varepsilon^{1-k\alpha_\varepsilon} b(\varepsilon) \leq Y_{\varepsilon,1} + \varepsilon H_{\varepsilon,k} \leq \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon)) \leq c_1 e^{-c_2 L^2}.$$

Note that (32) implies

$$(k-1)\alpha_\varepsilon \leq (N-1)\alpha_\varepsilon \leq \theta + \frac{C}{\log \varepsilon^{-1}},$$

so, using that $\varepsilon^{C/\log \varepsilon^{-1}} = e^{-C} = \text{const} > 0$, we have

$$\varepsilon^{1-\alpha_\varepsilon} \geq \varepsilon^{1-k\alpha_\varepsilon + C/\log \varepsilon^{-1}} a(\varepsilon) \geq C \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon),$$

which implies

$$\begin{aligned}
 |Y_{\varepsilon,1} + \varepsilon H_{\varepsilon,k}| &= |Y_\varepsilon(T'_\varepsilon) + \varepsilon H_{\varepsilon,k}| \geq \varepsilon e^{\lambda T'_\varepsilon} (|z| - |U_\varepsilon(T'_\varepsilon)| - |V_\varepsilon(T'_\varepsilon)|) - \varepsilon |H_{\varepsilon,k}| \\
 &\geq \varepsilon^{1-\alpha_\varepsilon} (|z| - |U_\varepsilon(T'_\varepsilon)| - |V_\varepsilon(T'_\varepsilon)| - \varepsilon^{\alpha_\varepsilon} |H_{\varepsilon,k}|) \\
 &\geq C_1 \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon) (|z| - |U_\varepsilon(T'_\varepsilon)| - \varepsilon C_2 - \varepsilon^{\alpha_\varepsilon} |H_{\varepsilon,k}|).
 \end{aligned}$$

Since $b(\varepsilon) = o(a(\varepsilon))$, the left-hand side of (36) can be thus bounded above by

$$\begin{aligned}
 \sup_{|z| \geq L} \mathbf{P}_{\varepsilon,z}(|Y_{\varepsilon,1} + \varepsilon H_{\varepsilon,k}| \leq \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon)) \\
 \leq \mathbf{P}(|H_{\varepsilon,k}| \geq L) + \mathbf{P}\left(|U_\varepsilon(T'_\varepsilon)| \geq L(1 - \varepsilon^{\alpha_\varepsilon}) - \frac{1}{C_1} - \varepsilon C_2\right),
 \end{aligned}$$

and (36) follows from the standard Gaussian tail bound, (33), and Lemma 3.1.

Conditioning on $H_{\varepsilon,k}$, using the independence of $M_\varepsilon(T'_\varepsilon)$ and $H_{\varepsilon,k}$, and using the existence of density of $M_\varepsilon(T'_\varepsilon)$ guaranteed by Proposition 3.3, we obtain

$$\begin{aligned}
 \mathbf{P}_{\varepsilon,z}(\varepsilon^{1-k\alpha_\varepsilon} b(\varepsilon) \leq Y_{\varepsilon,1} + \varepsilon H_{\varepsilon,k} \leq \varepsilon^{1-k\alpha_\varepsilon} a(\varepsilon)) \\
 &= \mathbf{P}_{\varepsilon,z}(\varepsilon^{-(k-1)\alpha_\varepsilon} b(\varepsilon) \leq M_\varepsilon(T'_\varepsilon) + \varepsilon^{\alpha_\varepsilon} H_{\varepsilon,k} \leq \varepsilon^{-(k-1)\alpha_\varepsilon} a(\varepsilon)) \\
 &= \mathbf{E} \int_{b(\varepsilon)\varepsilon^{-(k-1)\alpha_\varepsilon} - \varepsilon^{\alpha_\varepsilon} H_{\varepsilon,k}}^{a(\varepsilon)\varepsilon^{-(k-1)\alpha_\varepsilon} - \varepsilon^{\alpha_\varepsilon} H_{\varepsilon,k}} p_\varepsilon^z(u) du.
 \end{aligned}$$

This, along with the induction hypothesis (34), relations (35)–(36) and Fubini’s theorem, gives

$$(37) \quad P(\varepsilon, x) = \mathbf{E} \int_{b(\varepsilon)\varepsilon^{-(k-1)\alpha_\varepsilon - \varepsilon^{\alpha_\varepsilon}} H_{\varepsilon,k}}^{a(\varepsilon)\varepsilon^{-(k-1)\alpha_\varepsilon - \varepsilon^{\alpha_\varepsilon}} H_{\varepsilon,k}} \int_{-L}^L p_\varepsilon^z(u) r_x^{(k-1)}(\varepsilon^{(k-1)\alpha_\varepsilon} dz) du + \mathcal{O}(e^{-\lambda L^2}) + o(a(\varepsilon)).$$

To use the uniform convergence statement of Proposition 3.3, while simultaneously making the error term decay sufficiently fast, we choose

$$L = L(\varepsilon) = K(\varepsilon) = \sqrt{\frac{2\theta}{\lambda} \log \varepsilon^{-1}}.$$

This allows us to write (37) as

$$P(\varepsilon, x) = \mathbf{E} \int_{b(\varepsilon)\varepsilon^{-(k-1)\alpha_\varepsilon - \varepsilon^{\alpha_\varepsilon}} H_{\varepsilon,k}}^{a(\varepsilon)\varepsilon^{-(k-1)\alpha_\varepsilon - \varepsilon^{\alpha_\varepsilon}} H_{\varepsilon,k}} \int_{-\infty}^{\infty} p^z(u) \mathbf{P}_{\varepsilon x}(\varepsilon^{-(k-1)\alpha_\varepsilon} M_{\varepsilon,k-1} \in dz) du + o(a(\varepsilon)),$$

where we also used the decay rate of $p^z(u)$ at ∞ to restore the integration domain to the entire line.

Recalling that $p^z(u) = p^0(u - z)$, we change variables: $u' = \varepsilon^{(k-1)\alpha_\varepsilon} u$, $z' = \varepsilon^{(k-1)\alpha_\varepsilon} z$, and express the integral on the right-hand side of the last display as

$$\begin{aligned} & \varepsilon^{-(k-1)\alpha_\varepsilon} \mathbf{E} \int_{b(\varepsilon)\varepsilon^{-k\alpha_\varepsilon} H_{\varepsilon,k}}^{a(\varepsilon)\varepsilon^{-k\alpha_\varepsilon} H_{\varepsilon,k}} \int_{-\infty}^{\infty} p^0(\varepsilon^{-(k-1)\alpha_\varepsilon}(u' - \varepsilon^{(k-1)\alpha_\varepsilon} z')) \\ & \times \mathbf{P}_x(M_{\varepsilon,k-1} \in \varepsilon^{(k-1)\alpha_\varepsilon} dz) du' \\ & = \varepsilon^{-(k-1)\alpha_\varepsilon} \mathbf{E} \int_{b(\varepsilon)\varepsilon^{-k\alpha_\varepsilon} H_{\varepsilon,k}}^{a(\varepsilon)\varepsilon^{-k\alpha_\varepsilon} H_{\varepsilon,k}} \int_{-\infty}^{\infty} p^0(\varepsilon^{-(k-1)\alpha_\varepsilon}(u' - z')) \\ & \times \mathbf{P}_x(M_{\varepsilon,k-1} \in dz') du' \\ & = \mathbf{P}_{\varepsilon x}(b(\varepsilon) \leq M_{\varepsilon,k-1} + \varepsilon^{(k-1)\alpha_\varepsilon} \xi + \varepsilon^{k\alpha_\varepsilon} H_{\varepsilon,k} \leq a(\varepsilon)) \\ & = \mathbf{P}_{\varepsilon x}(b(\varepsilon) \leq M_{\varepsilon,k-1} + \varepsilon^{(k-1)\alpha_\varepsilon}(\xi + \varepsilon^{\alpha_\varepsilon} H_{\varepsilon,k}) \leq a(\varepsilon)), \end{aligned}$$

where ξ is a centered Gaussian random variable independent of both $M_{\varepsilon,k-1}$ and $H_{\varepsilon,k}$ with variance $(2\lambda)^{-1}\sigma^2(0)$. This completes the proof of the induction step once one verifies, by a straightforward computation, that $H_{\varepsilon,k-1} := \xi + \varepsilon^{\alpha_\varepsilon} H_{\varepsilon,k}$ has the desired variance given by (33). \square

REMARK 3.4. In the second part of the previous proof, the random variables $\varepsilon^{(k-1)\alpha_\varepsilon} \xi_k$ with

$$\xi_k = H_{\varepsilon,k-1} - \varepsilon^{\alpha_\varepsilon} H_{\varepsilon,k}$$

essentially correspond to the distributional limit of the integral

$$I_k = \int_{(k-1)T'_\varepsilon}^{kT'_\varepsilon} e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(s)) dW(s)$$

conditioned on Y_ε not having exited a neighborhood of the origin of size $\mathcal{O}(\varepsilon)$.

4. Malliavin calculus tools and the proof of Proposition 3.3. In this section, we turn to the proof of Proposition 3.3 by showing that $M'_\varepsilon := M_\varepsilon(T'_\varepsilon)$ has a bounded continuous density with the desired asymptotics. We are going to employ Malliavin calculus tools, using [14] as a basic reference.

Let us first recall some basic notions. Let $H = L^2(\mathbb{R}_+)$ be the separable Hilbert space of square integrable functions on the line and let $B : H \mapsto L^2(\Omega)$ be the *isonormal Gaussian process* on H given by

$$h \mapsto B(h) = \int_0^\infty h(s) dW(s).$$

Let $L^2(\Omega; H)$ be the set of square integrable H -valued random variables with norm $\|u\|_{L^2(\Omega; H)} = \sqrt{\mathbf{E}\|u\|_H^2}$. The *Malliavin derivative operator* $\mathcal{D} : L^2(\Omega) \mapsto L^2(\Omega; H)$ is defined on random variables of the form

$$(38) \quad F = f(B(h_1), \dots, B(h_n)), \quad n \geq 1,$$

by the formula

$$\mathcal{D}F = \sum_{i=1}^n \partial_{x_i} f(B(h_1), \dots, B(h_n)) h_i.$$

It is extended to a closed operator under the graph norm

$$\|F\|_{1,2} = \sqrt{\|F\|_2^2 + \|\mathcal{D}F\|_{2;H}^2},$$

with domain $\mathbf{D}_{1,2}$, where we adopt the notation $\|F\|_2 \equiv \|F\|_{L^2(\Omega)} = \sqrt{\mathbf{E}F^2}$ and $\|u\|_{2;H} = \|u\|_{L^2(\Omega; H)}$.

This construction can be extended to Hilbert space valued random variables producing a closed operator $\mathcal{D} : L^2(\Omega; H) \mapsto L^2(\Omega; H \otimes H)$ under the graph-norm

$$\|u\|_{1,2;H} = \sqrt{\|u\|_{2;H}^2 + \|\mathcal{D}u\|_{2;H \otimes H}^2},$$

with domain $\mathbf{D}_{1,2;H}$, where $\|\mathcal{D}u\|_{2;H \otimes H} = \|u\|_{L^2(\Omega; H \otimes H)}$. The second-order Malliavin derivative is the composition of the operators described above, a closed operator $\mathcal{D}^2 : L^2(\Omega) \rightarrow L^2(\Omega, H \otimes H)$.

The *divergence operator* or *Skorokhod integral* $\delta : L^2(\Omega; H) \mapsto L^2(\Omega)$ is defined as the $L^2(\Omega; H)$ adjoint of \mathcal{D} , that is,

$$\mathbf{E}(\mathcal{D}F, u)_{L^2(\Omega; H)} = \mathbf{E}[F\delta(u)], \quad u \in \text{Dom } \delta,$$

where $\text{Dom } \delta$ is defined as the set of those $u \in L^2(\Omega; H)$ such that the left-hand side of the indicated identity defines a bounded functional as a function of F . In particular, H can be embedded naturally into $\text{Dom } \delta$ and $\delta(h) = B(h)$ for any $h \in H$.

PROPOSITION 4.1 ([14], Proposition 2.2.1). *Let $F \in L^2(\Omega) \in D_{1,2}$ and assume that*

$$(39) \quad \frac{\mathcal{D}F}{\|\mathcal{D}F\|_H^2} \in \text{Dom } \delta.$$

Then F has a bounded continuous density given by

$$(40) \quad p(z) = \mathbf{E} \left[\mathbf{1}_{\{F > z\}} \delta \left(\frac{\mathcal{D}F}{\|\mathcal{D}F\|_H^2} \right) \right].$$

We are going to use this result with $F = M'_\varepsilon$ under the measure $\mathbf{P} = \mathbf{P}_{\varepsilon x}$. We will write $p_\varepsilon^x(z)$ for its density given by (40). We are going to compare M'_ε with the limiting random variable

$$(41) \quad I = x + \sigma(0) \int_0^\infty e^{-\lambda t} dW(t),$$

which is a centered Gaussian random variable with density $p^x(z)$. In what follows, C will denote a positive finite constant, independent of ε , x and t , which may change on each appearance.

LEMMA 4.2. *Let $Y_\varepsilon(0) = \varepsilon x$ with $|x| \leq K(\varepsilon)$, where $K(\varepsilon)$ satisfies (6). Then for any $\gamma \in (0, 2)$ and sufficiently small ε , we have*

$$\mathbf{E}|I - M'_\varepsilon|^2 \leq C\varepsilon^\gamma.$$

PROOF. We start by estimating

$$(42) \quad \begin{aligned} \mathbf{E}[Y_\varepsilon(t)]^2 &= \varepsilon^2 e^{2\lambda t} \mathbf{E}[M_\varepsilon(t)]^2 = \varepsilon^2 e^{2\lambda t} \mathbf{E}[x + U_\varepsilon(t) + V_\varepsilon(t)]^2 \\ &\leq C\varepsilon^2 e^{2\lambda t} (|x|^2 + \mathbf{E}[U_\varepsilon(t)]^2 + \mathbf{E}[V_\varepsilon(t)]^2) \\ &\leq C\varepsilon^2 e^{2\lambda t} (1 + |x|^2), \end{aligned}$$

where the first inequality is due to the elementary $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$, while the second one follows from Lemma 3.1 and (23).

Next, we write

$$(43) \quad \begin{aligned} I - M'_\varepsilon &= \int_0^{T'_\varepsilon} e^{-\lambda t} [\sigma(0) - \tilde{\sigma}(Y_\varepsilon(t))] dW(t) \\ &\quad + \sigma(0) \int_{T'_\varepsilon}^\infty e^{-\lambda t} dW(t) - V_\varepsilon(T'_\varepsilon) \\ &= J_1(\varepsilon) + J_2(\varepsilon) + J_3(\varepsilon). \end{aligned}$$

Using the boundedness of $\tilde{\sigma}'(x)$ and the identity $\tilde{\sigma}(0) = \sigma(0)$ [which is due to $f'(0) = 1$], we can write

$$\mathbf{E}J_1^2(\varepsilon) \leq C \int_0^{T'_\varepsilon} e^{-2\lambda t} \mathbf{E}[Y_\varepsilon(t)]^2 dt \leq C\varepsilon^2(1 + K^2(\varepsilon))T'_\varepsilon \leq C\varepsilon^\gamma$$

for any $\gamma \in (0, 2)$ and small ε , where the first inequality is due to (42) and $|x| \leq K(\varepsilon)$. By (26),

$$\mathbf{E}J_2^2(\varepsilon) \leq \sigma^2(0)\mathbf{E}\left[\int_{T'_\varepsilon}^\infty e^{-\lambda t} dW(t)\right]^2 = \frac{\sigma^2(0)}{2\lambda}e^{-2\lambda T'_\varepsilon} \leq C\varepsilon^2$$

for sufficiently small ε . Also, $\mathbf{E}J_3^2(\varepsilon) \leq C\varepsilon^2$ by (23). Combining (43) with the elementary inequality $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$ and the above estimates, we complete the proof. \square

To use Proposition 4.1, we need to control the divergence and have certain estimates on the Malliavin derivatives of M'_ε . This is the content of the next two results.

PROPOSITION 4.3 ([14], Proposition 1.5.8). *There is $C > 0$ such that*

$$\|\delta(u)\|_2 \leq C(\|\mathbf{E}u\|_H + \|\mathcal{D}u\|_{2;H\otimes H}),$$

where it is implicit that the finiteness of the right-hand side implies $u \in \text{Dom } \delta$.

LEMMA 4.4. *Let $Y_\varepsilon(0) = \varepsilon x$ and let $K(\varepsilon)$ satisfy (6). Then*

$$(44) \quad \lim_{\varepsilon \downarrow 0} \sup_{|x| \leq K(\varepsilon)} \|\mathcal{D}M'_\varepsilon - DI\|_{2;H} = 0$$

and, for every $m \geq 1$,

$$(45) \quad \limsup_{\varepsilon \rightarrow 0} \sup_{|x| \leq K(\varepsilon)} \mathbf{E}\|\mathcal{D}M'_\varepsilon\|_H^{-m} < \infty,$$

$$(46) \quad \lim_{\varepsilon \downarrow 0} \sup_{|x| \leq K(\varepsilon)} \mathbf{E}\|\mathcal{D}^2M'_\varepsilon\|_{2;H\otimes H}^m = 0.$$

The proof of Lemma 4.4 will be given in Section 5. We proceed now with the proof of Proposition 3.3.

PROOF OF PROPOSITION 3.3. Let us first assume that $F = M'_\varepsilon$ satisfies (39) and, as before, write $p_\varepsilon^x(z)$ for its density. We are going to compare M'_ε with the limiting random variable I [introduced in (41)], which is a centered Gaussian random variable with density $p^x(z)$.

We start by writing

$$\begin{aligned}
 |p_\varepsilon^x(z) - p^x(z)| &= \left| \mathbf{E} \left[\mathbf{1}_{\{M'_\varepsilon > z\}} \delta \left(\frac{\mathcal{D}M'_\varepsilon}{\|\mathcal{D}M'_\varepsilon\|_H^2} \right) - \mathbf{1}_{\{I > z\}} \delta \left(\frac{\mathcal{D}I}{\|\mathcal{D}I\|_H^2} \right) \right] \right| \\
 &= \left| \mathbf{E} \left[\mathbf{1}_{\{M'_\varepsilon > z\}} \delta \left(\frac{\mathcal{D}M'_\varepsilon}{\|\mathcal{D}M'_\varepsilon\|_H^2} - \frac{\mathcal{D}I}{\|\mathcal{D}I\|_H^2} \right) \right] \right| \\
 (47) \quad &+ \left| \mathbf{E} \left[(\mathbf{1}_{\{M'_\varepsilon > z\}} - \mathbf{1}_{\{I > z\}}) \delta \left(\frac{\mathcal{D}I}{\|\mathcal{D}I\|_H^2} \right) \right] \right| \\
 &\leq \left\| \delta \left(\frac{\mathcal{D}M'_\varepsilon}{\|\mathcal{D}M'_\varepsilon\|_H^2} - \frac{\mathcal{D}I}{\|\mathcal{D}I\|_H^2} \right) \right\|_2 \\
 &\quad + \|\mathbf{1}_{\{M'_\varepsilon > z\}} - \mathbf{1}_{\{I > z\}}\|_2 \left\| \delta \left(\frac{\mathcal{D}I}{\|\mathcal{D}I\|_H^2} \right) \right\|_2,
 \end{aligned}$$

where we used the Cauchy–Schwarz inequality in the last step.

Let us first deal with the second term on the right-hand side of (47). Observe

$$(48) \quad \mathcal{D}_t I = e^{-\lambda t} \sigma(0), \quad \|\mathcal{D}I\|_H^2 = \frac{\sigma^2(0)}{2\lambda},$$

so $\mathcal{D}I$ is deterministic and

$$(49) \quad \left\| \delta \left(\frac{\mathcal{D}I}{\|\mathcal{D}I\|_H^2} \right) \right\|_2 = \frac{2\lambda}{\sigma^2(0)} \left\| \int_0^\infty e^{-\lambda t} \sigma(0) dW(t) \right\|_2 = \frac{\sigma(0)}{\sqrt{2\lambda}} < \infty.$$

The other factor can be estimated as follows: for any $\eta > 0$,

$$\begin{aligned}
 \mathbf{E}[\mathbf{1}_{\{M'_\varepsilon > z\}} - \mathbf{1}_{\{I > z\}}]^2 &= \mathbf{P}\{M'_\varepsilon \leq z < I \text{ or } I \leq z < M'_\varepsilon\} \\
 (50) \quad &\leq \mathbf{P}(|I - z| \leq \eta) + \mathbf{P}(|I - M'_\varepsilon| > \eta)
 \end{aligned}$$

$$(51) \quad \leq C\eta + \frac{\mathbf{E}|I - M'_\varepsilon|^2}{\eta^2},$$

where in the last step we used the Markov inequality and that I has a bounded density. Choosing $\eta = \varepsilon^{\gamma'}$ for any $\gamma' \in (0, 1/2)$ and invoking Lemma 4.2 with $\gamma = 2\gamma'$ shows that the right-hand side of (50) converges to zero as $\varepsilon \downarrow 0$. This and (49) imply that the second term on the right-hand side of (47) converges to zero as well, uniformly in $|x| \leq K(\varepsilon)$ and $z \in \mathbb{R}$.

Next, we have to estimate the first term on the right-hand side of (47). We wish to apply Proposition 4.3 to the H -valued random variable u_ε given by

$$\begin{aligned}
 (52) \quad u_\varepsilon(t) &= \frac{\mathcal{D}_t M'_\varepsilon}{\|\mathcal{D}M'_\varepsilon\|_H^2} - \frac{\mathcal{D}_t I}{\|\mathcal{D}I\|_H^2} \\
 &= \mathcal{D}_t M'_\varepsilon \frac{\|\mathcal{D}I\|_H^2 - \|\mathcal{D}M'_\varepsilon\|_H^2}{\|\mathcal{D}M'_\varepsilon\|_H^2 \|\mathcal{D}I\|_H^2} + \frac{\mathcal{D}_t M'_\varepsilon - \mathcal{D}_t I}{\|\mathcal{D}I\|_H^2}.
 \end{aligned}$$

The H -norm of u_ε can be estimated using the triangle inequality and (48):

$$\begin{aligned}
 \|u_\varepsilon\|_H &\leq C \left(\frac{\|\mathcal{D}I\|_H^2 - \|\mathcal{D}M'_\varepsilon\|_H^2}{\|\mathcal{D}M'_\varepsilon\|_H} + \|\mathcal{D}_t M'_\varepsilon - \mathcal{D}_t I\|_H \right) \\
 &\leq C \left(\frac{(\|\mathcal{D}I\|_H^2 - \|\mathcal{D}M'_\varepsilon\|_H)(\|\mathcal{D}I\|_H^2 + \|\mathcal{D}M'_\varepsilon\|_H)}{\|\mathcal{D}M'_\varepsilon\|_H} \right. \\
 (53) \quad &\quad \left. + \|\mathcal{D}_t M'_\varepsilon - \mathcal{D}_t I\|_H \right) \\
 &\leq C \|\mathcal{D}M'_\varepsilon - \mathcal{D}I\|_H \left[1 + \frac{1}{\|\mathcal{D}M'_\varepsilon\|_H} \right].
 \end{aligned}$$

Using Jensen’s inequality, (53) and the Cauchy–Schwarz inequality yields

$$\begin{aligned}
 \|\mathbf{E}u_\varepsilon\|_H &\leq \mathbf{E}\|u_\varepsilon\|_H \leq C \left[\mathbf{E}\|\mathcal{D}M'_\varepsilon - \mathcal{D}I\|_H + \mathbf{E}\frac{\|\mathcal{D}M'_\varepsilon - \mathcal{D}I\|_H}{\|\mathcal{D}M'_\varepsilon\|_H} \right] \\
 (54) \quad &\leq C \|\mathcal{D}M'_\varepsilon - \mathcal{D}I\|_{2;H} \left[1 + \sqrt{\mathbf{E}\|\mathcal{D}M'_\varepsilon\|_H^{-2}} \right] \rightarrow 0
 \end{aligned}$$

as $\varepsilon \downarrow 0$, where the convergence is due to Lemma 4.4.

To treat the second term given by Proposition 4.3, we observe that $\mathcal{D}I$ being deterministic implies $\mathcal{D}^2I = 0$, and thus

$$\mathcal{D}_s u_\varepsilon(t) = \mathcal{D}_s \left[\frac{\mathcal{D}_t M'_\varepsilon}{\|\mathcal{D}M'_\varepsilon\|_H^2} \right],$$

where the right-hand side is estimated by (2.5) in [14] implying

$$\|\mathcal{D}u_\varepsilon\|_{H \otimes H} \leq \frac{3\|\mathcal{D}^2 M'_\varepsilon\|_{H \otimes H}}{\|\mathcal{D}M'_\varepsilon\|_H^2},$$

so we can use the Cauchy–Schwarz inequality to obtain

$$(55) \quad \|\mathcal{D}u_\varepsilon\|_{2;H \otimes H} \leq C(\mathbf{E}[\|\mathcal{D}M'_\varepsilon\|_H^{-8}])^{\frac{1}{4}} (\mathbf{E}\|\mathcal{D}^2 M'_\varepsilon\|_{H \otimes H}^4)^{\frac{1}{4}} \rightarrow 0$$

as $\varepsilon \downarrow 0$ by Lemma 4.4. Combining (54) and (55) with Proposition 4.3 gives

$$\lim_{\varepsilon \downarrow 0} \left\| \delta \left(\frac{\mathcal{D}M'_\varepsilon}{\|\mathcal{D}M'_\varepsilon\|_H^2} - \frac{\mathcal{D}I}{\|\mathcal{D}I\|_H^2} \right) \right\|_2 = 0$$

uniformly in $|x| \leq K(\varepsilon)$ completing the verification of

$$\lim_{\varepsilon \downarrow 0} \sup_{\substack{|x| \leq K(\varepsilon) \\ z \in \mathbb{R}}} |p_\varepsilon^x(z) - p^x(z)| = 0.$$

To complete the proof, it remains to show that $\tilde{u}_\varepsilon = \mathcal{D}M'_\varepsilon / \|\mathcal{D}M'_\varepsilon\|_H^2 \in \text{Dom } \delta$. Clearly, Jensen’s inequality and Lemma 4.4 imply

$$\|\mathbf{E}\tilde{u}_\varepsilon\|_H \leq \mathbf{E}\|\tilde{u}_\varepsilon\|_H = \mathbf{E}\|\mathcal{D}M'_\varepsilon\|_H^{-1} < \infty,$$

while $\mathcal{D}I = 0$ and (55) give us

$$\|\mathcal{D}\tilde{u}_\varepsilon\|_{2; H \otimes H} = \|\mathcal{D}u_\varepsilon\|_{2; H \otimes H} < \infty.$$

The desired conclusion now follows from Proposition 4.3 and the last two displays. \square

5. Proof of Lemma 4.4. In this section, we complete the proof of Theorem 1.1 by proving Lemma 4.4. As in the previous section, C will denote a positive finite constant, independent of ε , x , and t , which may change on each appearance and we will use the elementary inequality $(\sum_{i=1}^n a_i)^m \leq n^{m-1} \sum_{i=1}^n a_i^m$, Fubini’s theorem, and the boundedness of $\tilde{\sigma}$, h and all their derivatives without further mention.

We start by recalling that the Malliavin derivative $\mathcal{D}_t M_\varepsilon(u)$ satisfies a stochastic integral equation ([14], Theorem 2.2.1)

$$\begin{aligned} (56) \quad \mathcal{D}_t M_\varepsilon(u) &= e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(t)) + \int_t^u e^{-\lambda s} \tilde{\sigma}'(Y_\varepsilon(s)) \mathcal{D}_t Y_\varepsilon(s) dW(s) \\ &\quad + \frac{\varepsilon}{2} \int_t^u e^{-\lambda s} h'(Y_\varepsilon(s)) \mathcal{D}_t Y_\varepsilon(s) ds \\ &= e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(t)) + \varepsilon \int_t^u \tilde{\sigma}'(Y_\varepsilon(s)) \mathcal{D}_t M'_\varepsilon(s) dW(s) \end{aligned}$$

$$(57) \quad + \frac{\varepsilon^2}{2} \int_t^u h'(Y_\varepsilon(s)) \mathcal{D}_t M'_\varepsilon(s) ds$$

for $u \geq t$, while $\mathcal{D}_t M_\varepsilon(u) = 0$ otherwise. Here, we used (22) and Proposition 1.3.8 from [14] [see also formula (1.65) in this reference] in the first step, and (22) in the second one. This integral equation is equivalent to a stochastic differential equation in the variable u :

$$(58) \quad d[\mathcal{D}_t M_\varepsilon(u)] = \mathcal{D}_t M_\varepsilon(u) dZ_\varepsilon(u), \quad \mathcal{D}_t M_\varepsilon(t) = e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(t)),$$

where we introduced the semimartingale

$$Z_\varepsilon(u) = \varepsilon \int_t^u \tilde{\sigma}'(Y_\varepsilon(s)) dW(s) + \frac{\varepsilon^2}{2} \int_t^u h'(Y_\varepsilon(s)) ds.$$

It is well known (see, e.g., [10], Section 5.6) that the solution of the linear equation (58) is given by the Doléans-Dade exponential

$$\begin{aligned} (59) \quad \mathcal{D}_t M_\varepsilon(u) &= e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(t)) \exp\left(Z_\varepsilon(u) - \frac{\langle Z_\varepsilon \rangle_u}{2}\right) \\ &= e^{-\lambda t} \tilde{\sigma}(Y_\varepsilon(t)) \exp\left(\varepsilon \int_t^u \tilde{\sigma}'(Y_\varepsilon(s)) dW(s) \right. \end{aligned}$$

$$(60) \quad \left. + \frac{\varepsilon^2}{2} \int_t^u (h'(Y_\varepsilon(s)) - (\tilde{\sigma}'(Y_\varepsilon(s)))^2) ds\right).$$

LEMMA 5.1. *For any $m \geq 1$, we have*

$$\sup_{u \in [0, T'_\varepsilon]} \mathbf{E} |\mathcal{D}_t M_\varepsilon(u)|^m \leq C e^{-m\lambda t}.$$

PROOF. We can use the explicit formula (59) to estimate

$$\begin{aligned} \mathbf{E} |\mathcal{D}_t M_\varepsilon(u)|^m &\leq C e^{-m\lambda t} \mathbf{E} \left[\exp(\varepsilon m \int_t^u \tilde{\sigma}'(Y_\varepsilon(s)) dW(s) \right. \\ &\quad \left. + \frac{\varepsilon^2 m}{2} \int_t^u (h'(Y_\varepsilon(s)) - (\tilde{\sigma}'(Y_\varepsilon(s))^2) ds) \right] \\ &\leq C e^{-m\lambda t} \mathbf{E} \left[\exp(\varepsilon m \int_t^u \tilde{\sigma}'(Y_\varepsilon(s)) dW(s) \right. \\ &\quad \left. - \frac{\varepsilon^2 m^2}{2} \int_t^u (\tilde{\sigma}'(Y_\varepsilon(s))^2 ds) \right] = C e^{-m\lambda t}, \end{aligned}$$

where we used $\varepsilon^2(u - t) \leq \varepsilon^2 T'_\varepsilon \rightarrow 0$, as $\varepsilon \downarrow 0$, in the second inequality to manipulate the Lebesgue integrals in the exponent, while the last step is due to the martingale property of the process under the expectation. \square

Let us now prove the claims of Lemma 4.4 one by one.

PROOF OF (44). Let $I(u) = \sigma(0) \int_0^u e^{-\lambda s} dW(s)$. Then

$$\mathcal{D}_t I(u) = \sigma(0) e^{-\lambda t} \mathbf{1}_{\{t \leq u\}}.$$

By the triangle inequality, we have

$$|\mathcal{D}_t M'_\varepsilon - \mathcal{D}_t I| \leq |\mathcal{D}_t M_\varepsilon(T'_\varepsilon) - \mathcal{D}_t I(T'_\varepsilon)| + e^{-\lambda t} \sigma(0) \mathbf{1}_{\{t > T'_\varepsilon\}},$$

so

$$\begin{aligned} \|\mathcal{D}M_\varepsilon - \mathcal{D}I\|_{2;H}^2 &= \int_0^\infty \mathbf{E} |\mathcal{D}_t M_\varepsilon - \mathcal{D}_t I|^2 dt \\ (61) \quad &\leq C \int_0^\infty \mathbf{E} |\mathcal{D}_t M_\varepsilon(T'_\varepsilon) - \mathcal{D}_t I(T'_\varepsilon)|^2 dt + C \sigma^2(0) e^{-2\lambda T'_\varepsilon}. \end{aligned}$$

The second term on the right-hand side converges to zero simply by $T'_\varepsilon \rightarrow \infty$. To estimate the first term, we use (56) and the triangle inequality:

$$(62) \quad |\mathcal{D}_t M_\varepsilon(T'_\varepsilon) - \mathcal{D}_t I(T'_\varepsilon)| \leq \mathbf{1}_{\{t \leq T'_\varepsilon\}} C e^{-\lambda t} |\tilde{\sigma}(Y_\varepsilon(t)) - \sigma(0)| + |R_\varepsilon(t)|,$$

where

$$R_\varepsilon(t) = \varepsilon \int_t^{T'_\varepsilon} \tilde{\sigma}'(Y_\varepsilon(s)) \mathcal{D}_t M'_\varepsilon(s) dW(s) + \frac{\varepsilon^2}{2} \int_t^{T'_\varepsilon} h'(Y_\varepsilon(s)) \mathcal{D}_t M'_\varepsilon(s) ds.$$

The second moment of this latter quantity can be bounded as

$$\mathbf{E}|R_\varepsilon(t)|^2 \leq C\varepsilon^2(1 + \varepsilon^2 T'_\varepsilon) \int_t^{T'_\varepsilon} \mathbf{E}|\mathcal{D}_t M'_\varepsilon(s)|^2 ds \leq C\varepsilon^2 e^{-2\lambda t},$$

where the first inequality is due to Jensen’s inequality, while we used Lemma 5.1 and $\lim_{\varepsilon \downarrow 0} \varepsilon^2 T'_\varepsilon = 0$ in the second one. Combining this with (62) and (42), we obtain

$$\begin{aligned} \mathbf{E}|\mathcal{D}_t M_\varepsilon(T'_\varepsilon) - \mathcal{D}_t I(T'_\varepsilon)|^2 &\leq C(\mathbf{1}_{\{t \leq T'_\varepsilon\}} e^{-2\lambda t} \mathbf{E}|Y_\varepsilon(t)|^2 + \mathbf{E}|R_\varepsilon(t)|^2) \\ &\leq C(\varepsilon^2(1 + K^2(\varepsilon))\mathbf{1}_{\{t \leq T'_\varepsilon\}} + \varepsilon^2 e^{-2\lambda t}), \end{aligned}$$

and integrating with respect to t gives us

$$\int_0^\infty \mathbf{E}|\mathcal{D}_t M_{\varepsilon,\lambda}(T'_\varepsilon) - \mathcal{D}_t I(T'_\varepsilon)|^2 dt \leq C\varepsilon^2 T'_\varepsilon(1 + K^2(\varepsilon)) \rightarrow 0$$

as $\varepsilon \downarrow 0$. Therefore the right-hand side of (61) converges to zero completing the proof. \square

PROOF OF (45). Applying Jensen’s inequality for the convex function $x \mapsto |x|^{-m/2}$ and the measure on $[0, T'_\varepsilon]$ with density

$$\frac{e^{-2\lambda t}}{\int_0^{T'_\varepsilon} e^{-2\lambda s} ds} = \frac{2\lambda e^{-2\lambda t}}{1 - e^{-2\lambda T'_\varepsilon}},$$

we obtain

$$\begin{aligned} \mathbf{E}\|DM'_\varepsilon\|_H^{-m} &= \mathbf{E}\left(\int_0^{T'_\varepsilon} e^{-2\lambda t} e^{2\lambda t} |\mathcal{D}_t M'_\varepsilon|^2 dt\right)^{-m/2} \\ &\leq \left(\frac{2\lambda}{1 - e^{-2\lambda T'_\varepsilon}}\right)^{\frac{m}{2}+1} \int_0^{T'_\varepsilon} e^{-2\lambda t} e^{-m\lambda t} \mathbf{E}|\mathcal{D}_t M'_\varepsilon|^{-m} dt. \end{aligned}$$

Using the explicit formula (59), we can estimate

$$\begin{aligned} e^{-m\lambda t} \mathbf{E}|\mathcal{D}_t M'_\varepsilon|^{-m} &\leq C\mathbf{E} \exp(-\varepsilon m \int_t^{T'_\varepsilon} \tilde{\sigma}'(Y_\varepsilon(s)) dW(s) \\ &\quad - \frac{\varepsilon^2 m}{2} \int_t^{T'_\varepsilon} (h'(Y_\varepsilon(s)) + (\tilde{\sigma}'(Y_\varepsilon(s)))^2) ds) \\ &\leq C\mathbf{E} \left[\exp(\varepsilon m \int_t^{T'_\varepsilon} \tilde{\sigma}'(Y_\varepsilon(s)) dW(s) \right. \\ &\quad \left. - \frac{\varepsilon^2 m^2}{2} \int_t^{T'_\varepsilon} (\tilde{\sigma}'(Y_\varepsilon(s)))^2 ds) \right] = C, \end{aligned}$$

where we used $\lim_{\varepsilon \downarrow 0} \varepsilon^2 T'_\varepsilon = 0$ in the second inequality to manipulate the Lebesgue integrals, while the last step is due to the martingale property of the exponential. Combining the last two displays completes the proof. \square

PROOF OF (46). It is sufficient to prove the claim for $m \geq 2$ since convergence for $m \in [1, 2)$ will follow from Lyapunov’s inequality. Taking the Malliavin derivative of both sides of (56) and using Proposition 1.3.8 in [14], we obtain that the second derivative $\mathcal{D}_{r,t}^2 M'_\varepsilon(u)$ satisfies the integral equation

$$\begin{aligned} \mathcal{D}_{r,t}^2 M'_\varepsilon(u) &= \varepsilon \tilde{\sigma}'(Y_\varepsilon(t)) \mathcal{D}_r M'_\varepsilon(t) + \varepsilon \tilde{\sigma}'(Y_\varepsilon(r)) \mathcal{D}_t M'_\varepsilon(r) \\ &\quad + \varepsilon^2 \int_{r \vee t}^u e^{\lambda s} \tilde{\sigma}''(Y_\varepsilon(s)) \mathcal{D}_r M'_\varepsilon(s) \mathcal{D}_t M'_\varepsilon(s) dW(s) \\ &\quad + \varepsilon \int_{r \vee t}^u \tilde{\sigma}'(Y_\varepsilon(s)) \mathcal{D}_{r,t}^2 M_\varepsilon(s) dW(s) \\ &\quad + \frac{\varepsilon^3}{2} \int_{r \vee t}^u e^{\lambda s} h''(Y_\varepsilon(s)) \mathcal{D}_r M'_\varepsilon(s) \mathcal{D}_t M'_\varepsilon(s) ds \\ &\quad + \frac{\varepsilon^2}{2} \int_{r \vee t}^u h'(Y_\varepsilon(s)) \mathcal{D}_{r,t}^2 M'_\varepsilon(s) ds, \end{aligned}$$

for $u \geq r \vee t$, while $\mathcal{D}_{r,t}^2 M'_\varepsilon(u) = 0$ if $u < r \vee t$.

Taking m th moments, we can apply the a priori bound in Lemma 5.1, the BDG inequality, and Jensen’s inequality to estimate

$$\begin{aligned} \mathbf{E}|\mathcal{D}_{r,t}^2 M'_\varepsilon(u)|^m &\leq C\varepsilon^2(e^{-m\lambda t} + e^{-m\lambda r}) \\ &\quad + C\varepsilon^{2m}(T'_\varepsilon)^{\frac{m}{2}-1} \int_{r \vee t}^u e^{m\lambda s} \mathbf{E}|\mathcal{D}_r M'_\varepsilon(s) \mathcal{D}_t M'_\varepsilon(s)|^m ds \\ &\quad + C\varepsilon^m(T'_\varepsilon)^{m-1} \int_{r \vee t}^u \mathbf{E}|\mathcal{D}_{r,t}^2 M'_\varepsilon(s)|^m ds. \end{aligned}$$

The Cauchy–Schwarz inequality and Lemma 5.1 again imply

$$\begin{aligned} \mathbf{E}|\mathcal{D}_r M'_\varepsilon(s) \mathcal{D}_t M'_\varepsilon(s)|^m &\leq \sqrt{\mathbf{E}|\mathcal{D}_r M'_\varepsilon(s)|^{2m} \mathbf{E}|\mathcal{D}_t M'_\varepsilon(s)|^{2m}} \\ &\leq C\sqrt{e^{-2m\lambda r} e^{-2m\lambda t}} = Ce^{-m\lambda(r+t)}. \end{aligned}$$

Combining the last two displays, we obtain

$$\begin{aligned} \mathbf{E}|\mathcal{D}_{r,t}^2 M'_\varepsilon(u)|^m &\leq C\varepsilon^m(e^{-m\lambda r} + e^{-m\lambda t} + \varepsilon^m(T'_\varepsilon)^{\frac{m}{2}-1} e^{m\lambda(u-r-t)}) \\ &\quad + C\varepsilon^m(T'_\varepsilon)^{m-1} \int_{r \vee t}^u \mathbf{E}|\mathcal{D}_{r,t}^m M'_\varepsilon(s)|^2 ds. \end{aligned}$$

We recall Gronwall’s inequality: if nonnegative functions h, g, k satisfy

$$h(u) \leq g(u) + \int_a^u k(s)h(s) ds, \quad u \geq a,$$

and g is nondecreasing, then

$$h(u) \leq g(u)e^{\int_a^u k(s) ds}, \quad u \geq a.$$

Using this with $h(u) = \mathbf{E}|\mathcal{D}_{r,t}^2 M'_\varepsilon(u)|^m$, $a = r \vee t$,

$$g(u) = C\varepsilon^m(e^{-m\lambda r} + e^{-m\lambda t} + \varepsilon^m(T'_\varepsilon)^{\frac{m}{2}-1}e^{m\lambda(u-r-t)}),$$

where $k(u) = C\varepsilon^m(T'_\varepsilon)^{m-1}$, and setting $u = T'_\varepsilon$, we obtain

$$\begin{aligned} \mathbf{E}|\mathcal{D}_{r,t}^2 M'_\varepsilon|^m &\leq C\varepsilon^m(e^{-m\lambda r} + e^{-m\lambda t} + \varepsilon^m(T'_\varepsilon)^{\frac{m}{2}-1}e^{m\lambda(T'_\varepsilon-r-t)})e^{C\varepsilon^m(T'_\varepsilon)^{m-1}(T'_\varepsilon-r\vee t)} \\ &\leq C\varepsilon^m(e^{-m\lambda r} + e^{-m\lambda t} + \varepsilon^m(T'_\varepsilon)^{\frac{m}{2}-1}e^{m\lambda(T'_\varepsilon-r-t)})e^{C\varepsilon^m(T'_\varepsilon)^m} \\ &\leq C(\varepsilon^m(e^{-m\lambda r} + e^{-m\lambda t}) + \varepsilon^{2m}(T'_\varepsilon)^{\frac{m}{2}-1}e^{m\lambda(T'_\varepsilon-r-t)}) \\ &\leq C(\varepsilon^m(e^{-m\lambda r} + e^{-m\lambda t}) + \varepsilon^{2m}\varepsilon^{-\rho}e^{m\lambda(T'_\varepsilon-r-t)}) \end{aligned}$$

for any $\rho > 0$ and sufficiently small ε , where we used $\lim_{\varepsilon \downarrow 0} \varepsilon^a(T'_\varepsilon)^b = 0$ for all $a, b > 0$. Using this along with Jensen’s inequality, we obtain

$$\begin{aligned} \mathbf{E}\|\mathcal{D}^2 M'_\varepsilon\|_{H \otimes H}^m &= \mathbf{E}\left(\int_0^{T'_\varepsilon} \int_0^{T'_\varepsilon} |\mathcal{D}_{r,t}^2 M'_\varepsilon|^2 dr dt\right)^{\frac{m}{2}} \\ &\leq (T'_\varepsilon)^{m-2} \int_0^{T'_\varepsilon} \int_0^{T'_\varepsilon} \mathbf{E}|\mathcal{D}_{r,t}^2 M'_\varepsilon|^m dr dt \\ &\leq C(\varepsilon^m(T'_\varepsilon)^m + \varepsilon^{2m-\rho}e^{m\lambda T'_\varepsilon}) \leq C(\varepsilon^m(T'_\varepsilon)^m + \varepsilon^{m\kappa-\rho}), \end{aligned}$$

where we used (26) in the last step. Choosing $\rho \in (0, m\kappa)$ to ensure that the right-hand side of the previous display converges to zero, we complete the proof. \square

REFERENCES

[1] ALMADA MONTER, S. A. and BAKHTIN, Y. (2011). Normal forms approach to diffusion near hyperbolic equilibria. *Nonlinearity* **24** 1883–1907. [MR2802310](#)

[2] BAKHTIN, Y. (2008). Exit asymptotics for small diffusion about an unstable equilibrium. *Stochastic Process. Appl.* **118** 839–851. [MR2411523](#)

[3] BAKHTIN, Y. (2010). Small noise limit for diffusions near heteroclinic networks. *Dyn. Syst.* **25** 413–431. [MR2731621](#)

[4] BAKHTIN, Y. (2011). Noisy heteroclinic networks. *Probab. Theory Related Fields* **150** 1–42. [MR2800902](#)

[5] BASS, R. F. (2011). *Stochastic Processes. Cambridge Series in Statistical and Probabilistic Mathematics* **33**. Cambridge Univ. Press, Cambridge. [MR2856623](#)

[6] CHAMPAGNAT, N. and VILLEMONAIS, D. (2016). Exponential convergence to quasi-stationary distribution and Q -process. *Probab. Theory Related Fields* **164** 243–283. [MR3449390](#)

[7] DAY, M. V. (1995). On the exit law from saddle points. *Stochastic Process. Appl.* **60** 287–311. [MR1376805](#)

[8] EIZENBERG, A. (1984). The exit distributions for small random perturbations of dynamical systems with a repulsive type stationary point. *Stochastics* **12** 251–275. [MR0749377](#)

- [9] FREIDLIN, M. I. and WENTZELL, A. D. (2012). *Random Perturbations of Dynamical Systems*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **260**. Springer, Heidelberg. Translated from the 1979 Russian original by Joseph Szücs. [MR2953753](#)
- [10] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. [MR1121940](#)
- [11] KIFER, Y. (1981). The exit problem for small random perturbations of dynamical systems with a hyperbolic fixed point. *Israel J. Math.* **40** 74–96. [MR0636908](#)
- [12] MIKAMI, T. (1995). Large deviations for the first exit time on small random perturbations of dynamical systems with a hyperbolic equilibrium point. *Hokkaido Math. J.* **24** 491–525. [MR1357028](#)
- [13] NOURDIN, I. and VIENS, F. G. (2009). Density formula and concentration inequalities with Malliavin calculus. *Electron. J. Probab.* **14** 2287–2309. [MR2556018](#)
- [14] NUALART, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. Springer, Berlin. [MR2200233](#)

COURANT INSTITUTE OF
MATHEMATICAL SCIENCES
NEW YORK UNIVERSITY
251 MERCER STREET
NEW YORK, NEW YORK 10012
USA
E-MAIL: bakhtin@cims.nyu.edu
zsolt@cims.nyu.edu