SUBEXPONENTIAL LOWER BOUNDS FOR f-ERGODIC MARKOV PROCESSES

MIHA BREŠAR AND ALEKSANDAR MIJATOVIĆ

ABSTRACT. We provide a criterion for establishing lower bounds on the rate of convergence in f-variation of a continuous-time ergodic Markov process to its invariant measure. The criterion consists of novel super- and submartingale conditions for certain functionals of the Markov process. It provides a general approach for proving lower bounds on the tails of the invariant measure and the rate of convergence in f-variation of a Markov process, analogous to the widely used Lyapunov drift conditions for upper bounds. Our key technical innovation produces lower bounds on the tails of the heights and durations of the excursions from bounded sets of a continuous-time Markov process using path-wise arguments.

We apply our theory to elliptic diffusions and Lévy-driven stochastic differential equations with known polynomial/stretched exponential upper bounds on their rates of convergence. Our lower bounds match asymptotically the known upper bounds for these classes of models, thus establishing their rate of convergence to stationarity. The generality of the approach suggests that, analogous to the Lyapunov drift conditions for upper bounds, our methods can be expected to find applications in many other settings.

Key words: subgeometric ergodicity; lower bounds; rate of convergence in f-variation and total variation; invariant measure; return times; Lyapunov functions and Foster-type drift conditions.

AMS Subject Classification 2020: 60J25; 37A25 (Primary); 60J35; 60J60 (Secondary).

1. INTRODUCTION

Quantifying the rate of convergence of a Markov process X towards its invariant measure π is a fundamental problem in probability and its applications (see [11, 26] and the references therein). The literature typically focuses on *upper* bounds on the rates of convergence (see e.g. [9, 13, 28]). One of the most effective approaches to address this problem is via Lyapunov functions and the corresponding drift conditions. This approach has been widely used since the seminal work of Meyn and Tweedie [26, 27] and, due to its broad applicability and robustness, remains one of the most popular methods for quantifying upper bounds on the convergence rates of Markov processes. However, an important limitation of this method, particularly in the case when the upper bounds are subexponential, is the lack of the corresponding lower bounds, which would allow the user to establish the rates of convergence to stationarity. In this paper we develop a theory based on novel

drift conditions that yield *lower* subexponential bounds on the rates of convergence in a general setting. We apply it to models previously studied using Lyapunov functions for upper bounds and establish asymptotically matching lower bounds.

Our results can be summarized as follows: take a Lyapunov function $V : \mathcal{X} \to [1, \infty)$ (where $\mathcal{X}$ is the state space of X), which may (but need not) satisfy the upper bound drift conditions in [9] (also given in (10) below). Construct scalar functions $\varphi : (0, 1] \to \mathbb{R}_+$ and $\Psi : [1, \infty) \to [1, \infty)$ such that, outside of a compact set, the following processes

$$(1) \qquad 1/V(X) - \int_0^{\cdot} \varphi \circ (1/V)(X_s)\mathrm{d}s \qquad \text{and} \qquad \Psi \circ V(X)$$

are a supermartingale and a submartingale, respectively. Then we obtain *lower* bounds on the tails of the return times of X to compact sets, the tails of the invariant measure π of X and the rate of convergence in f-variation (including total variation) distance of the law of X_t to π.

Building on ideas in [13, 15], we provide general Foster-Lyapunov-type drift conditions implying *lower* bounds for *all times* on the convergence rates of ergodic Markov processes. The super- and submartingale properties of the processes in (1) are typically verified using the infinitesimal characteristics of X and Itô's formula. In conjunction with the classical drift conditions for upper bounds [9, 12, 28], our results offer a comprehensive and robust approach to quantifying the convergence rate of ergodic Markov processes.

As pointed out in [13, Sec. 5], deriving a lower bound on the rate of convergence to stationarity typically requires a deep understanding of the tail of the invariant measure, which is available only in very special cases (e.g. Langevin diffusions). The main contribution of this paper consists of providing a robust framework for deriving lower bounds on the convergence rates without prior knowledge of the tail behavior of the invariant measure. For example, in the context of (possibly Lévy-driven) stochastic differential equations with multiplicative noise, our theory makes it possible to establish lower bounds on f-variation for *all times* (as opposed to along a sparse sequence of times going to infinity as in [15]). To the best of our knowledge, such results were previously only available for Langevin diffusions [15, Thm 2.1], cf. Section 2.6 below.

The remainder of the paper is structured as follows. Section 2 states the novel **L**-drift condition and formulates our main theorems for ergodic Markov processes. Section 3 presents the applications of the results of Section 2 to classes of models, studied in [9, 12], exhibiting subexponential (and exponential) ergodicity. Sections 4 and 5 contain the proofs of the results of Section 2. More precisely, Sections 4 states and proves lower bounds for the tails of return times to bounded sets for càdlàg semimartingales. Section 5 applies these results in the context of càdlàg ergodic Markov processes to prove the theorems of Section 2. The proofs of the results, stated in Section 3, for ergodic elliptic diffusions, Lévy-driven stochastic differential equations and hypoelliptic stochastic damping Hamiltonian system are in Section 6. Finally, Section 7 offers concluding remarks, highlights some open questions and describes future directions of research. (See [3] for short YouTube presentations describing the applications and proofs of our main results.)

2. Main results

2.1. **Basic definitions.** Let $X = (X_t)_{t \in \mathbb{R}_+}$, where $\mathbb{R}_+ := [0, \infty)$, be a strong Markov process on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+})$, taking values in a locally compact separable metric space $\mathcal{X}$, endowed with the Borel σ-algebra $\mathcal{B}(\mathcal{X})$. For any $x \in \mathcal{X}$, denote by $\mathbb{P}_x$ and (resp. $\mathbb{E}_x$) the associated probability measure (resp. expectation), satisfying $\mathbb{P}_x(X_0 = x) = 1$, and assume that X has càdlàg (i.e. right-continuous paths with left limits) paths. The process X is ν-irreducible (resp. *Harris recurrent*), where ν is a σ-finite measure on $\mathcal{B}(\mathcal{X})$, if for every $A \in \mathcal{B}(\mathcal{X})$, such that $\nu(A) > 0$, and $x \in \mathcal{X}$ we have $\mathbb{E}_x[\int_0^\infty \mathbb{1}\{X_t \in A\}\mathrm{d}t] > 0$ (resp. $\mathbb{P}_x(\int_0^\infty \mathbb{1}\{X_t \in A\}\mathrm{d}t = \infty) = 1$). The probability measure π on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is an *invariant measure* for X if for all bounded measurable functions $g : \mathcal{X} \to \mathbb{R}_+$ and $t \in \mathbb{R}_+$ we have $\int_{\mathcal{X}} \mathbb{E}_x[g(X_t)]\pi(\mathrm{d}x) = \int_{\mathcal{X}} g(x)\pi(\mathrm{d}x)$. A Harris recurrent process is *positive Harris recurrent* if it admits an invariant measure. For any measurable function $f : \mathcal{X} \to [1, \infty)$, the *$f$-variation* of a signed measure μ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is given by $\|\mu\|_f := \sup\{|\int_{\mathcal{X}} g(x)\mu(\mathrm{d}x)| : g : \mathcal{X} \to \mathbb{R} \text{ measurable}, |g| \leq f\}$. In the special case $f \equiv 1$ we obtain the *total variation* $\|\mu\|_{\mathrm{TV}} := \|\mu\|_f$ of the measure μ. The process X is *ergodic* if $\lim_{t \to \infty} \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} = 0$ for all $x \in \mathcal{X}$.

As we are interested in convergent Markov processes, unless explicitly stated otherwise, the following **standard assumption** holds throughout the paper: *$X = (X_t)_{t \in \mathbb{R}_+}$ is an ergodic, positive Harris recurrent Markov process with càdlàg paths and invariant measure π on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$.*

2.2. **Lower bounds: the tails of π and f-variation.** The lower bounds on the tails of the invariant measure and the rate of convergence of X will be implied by the following **L**-*drift condition*.

Assumption ($\mathbf{L(V, \varphi, \Psi)}$). *Let $V : \mathcal{X} \to [1, \infty)$ be a continuous function, such that, for all $x \in \mathcal{X}$, $\limsup_{t \to \infty} V(X_t) = \infty$ $\mathbb{P}_x$-a.s. Assume also there exists $\ell_0 \in [1, \infty)$ such that (i) and (ii) hold.*
(i) Let $\varphi : (0, 1] \to \mathbb{R}_+$ be non-decreasing[1] and continuous, such that $r \mapsto r\varphi(1/r)$ is decreasing on $[1, \infty)$ and $\lim_{r \to \infty} r\varphi(1/r) = 0$. Assume that for $b \in \mathbb{R}_+$ and any $x \in \mathcal{X}$, the process

$$1/V(X) - \int_0^{\cdot} \varphi(1/V(X_u))\mathrm{d}u - b\int_0^{\cdot} \mathbb{1}\{V(X_u) \leq \ell_0\}\mathrm{d}u \quad \text{is an } (\mathcal{F}_t)\text{-supermartingale under } \mathbb{P}_x.$$

(ii) Let $\Psi : [1, \infty) \to [1, \infty)$ be a differentiable, increasing, submultiplicative[2] function satisfying the following: for any $\ell \in (\ell_0, \infty)$, there exists a constant $C_\ell \in (0, \infty)$ such that

$$(2) \qquad \mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq C_\ell/\Psi(r) \quad \text{for all } r \in (\ell + 1, \infty) \text{ and } x \in \{\ell + 1 \leq V\},$$

where $T^{(r)} := \inf\{t \geq 0 : V(X_t) > r\}$ and $S_{(\ell)} := \inf\{t \geq 0 : V(X_t) < \ell\}$.

In applications, the **L**-drift condition $\mathbf{L(V, \varphi, \Psi)}$ is verified via the infinitesimal generator of X, see Theorem 2.7, Subsection 2.4 and examples in Section 3 below. In particular, the exit probability estimate in (2) is typically implied by a submartingale property of the (appropriately stopped) process $\Psi(V(X))$, see Lemma 2.8 below for details.

---

[1] A non-decreasing function may have intervals of constancy, while an increasing function does not; similarly for non-increasing and decreasing.

[2] A function $\Psi : [1, \infty) \to [1, \infty)$ is *submultiplicative* if it satisfies $\Psi(r_1 + r_2) \leq C\Psi(r_1)\Psi(r_2)$ for some constant $C \in (0, \infty)$ and all $r_1, r_2 \in [1, \infty)$ [34, Def. 25.2].

Theorem 2.1 gives a lower bound on the tail of invariant measure π of the process X under Assumption $\mathbf{L}(V,\varphi,\Psi)$. Our lower bounds are in terms of the functions V, φ, Ψ and a logarithmic correction:[3] for any $\varepsilon, q \in (0, 1)$, define the function

$$(3) \qquad L_{\varepsilon,q}(r) := r\varphi(1/r)\Psi(2r/(1-q))(\log\log r)^\varepsilon, \quad r \in [1, \infty).$$

Theorem 2.1 (Tails of the invariant measure). *Let Assumption $\mathbf{L}(V,\varphi,\Psi)$ hold. Then for any $q, \varepsilon \in (0, 1)$ there exists a constant $c_{\varepsilon,q} \in (0, 1)$ such that*

$$(4) \qquad c_{\varepsilon,q}/L_{\varepsilon,q}(r) \le \pi(x \in \mathcal{X} : V(x) \ge r) \qquad \text{for all } r \in [1, \infty).$$

Theorem 2.2 provides a lower bound on the rate of f-variation convergence of X to its invariant measure π under Assumption $\mathbf{L}(V,\varphi,\Psi)$.

Theorem 2.2 (Lower bounds for f-variation). *Let Assumption $\mathbf{L}(V,\varphi,\Psi)$ hold. Consider a function $f : \mathcal{X} \to [1, \infty)$, satisfying $f = f_\star \circ V$ for some differentiable $f_\star : [1, \infty) \to [1, \infty)$. Assume further that (a) and (b) hold.*

(a) Let a continuous function $h : [1, \infty) \to [1, \infty)$ be such that the function $g := h/f_\star$ is increasing (and without loss of generality $g(1) = 1$) and $\lim_{r\to\infty} g(r) = \infty$. Let $v : \mathcal{X} \times \mathbb{R}_+ \to [1, \infty)$ be increasing in the second argument and satisfy

$$\mathbb{E}_x[h \circ V(X_t)] \le v(x, t) \quad \text{for all } x \in \mathcal{X} \text{ and } t \in \mathbb{R}_+.$$

(b) Pick $\varepsilon, q \in (0, 1)$ and a constant $c_{\varepsilon,q} \in (0, 1)$, such that the inequality in (4) holds with the function $L_{\varepsilon,q}$, and consider a continuous function $a : [1, \infty) \to \mathbb{R}_+$, satisfying

$$a(t) \le f_\star(g^{-1}(t))c_{\varepsilon,q}/L_{\varepsilon,q}(g^{-1}(t)) \quad \text{for all } t \in [1, \infty),$$

where g^{-1} is the inverse of the increasing function g in (a). Suppose also that the function $A(t) := ta(t)$ is increasing, $\lim_{t\to\infty} A(t) = \infty$ and denote its inverse by A^{-1}.

Define the function $r_f : \mathcal{X} \times [1, \infty) \to \mathbb{R}_+$ by $r_f := a \circ A^{-1} \circ (2v)$. Then

$$(5) \qquad r_f(x,t)/2 \le \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_f \quad \text{for all } x \in \mathcal{X} \text{ and } t \in [1, \infty).$$

We obtain the *lower* bound on the rate of convergence in the total variation distance by choosing $f_\star \equiv 1$ (and hence $f = f_\star \circ V \equiv 1$) in the previous theorem.

Corollary 2.3 (Lower bounds for total variation). *Let Assumption $\mathbf{L}(V,\varphi,\Psi)$ hold and set $f_\star \equiv 1$. Assume there exist continuous functions $h, a : [1, \infty) \to \mathbb{R}_+$ and $v : \mathcal{X} \times \mathbb{R}_+ \to [1, \infty)$, satisfying conditions (a) and (b) in Theorem 2.2. Let $r_1 : \mathcal{X} \times [1, \infty) \to \mathbb{R}_+$, $r_1 := a \circ A^{-1} \circ (2v)$, be as in Theorem 2.2. Then the following lower bound holds*

$$(6) \qquad r_1(x,t)/2 \le \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \quad \text{for all } x \in \mathcal{X} \text{ and } t \in [1, \infty).$$

Remark 2.4. (a) The lower bound in Theorem 2.2 is obtained by comparing the tail (with respect to V) of the invariant measure π to the tail of marginal distribution of $V(X_t)$ at time t. The function h in condition (a) of Theorem 2.2 aims to maximise the growth of $r \mapsto r\pi(\{h \circ V \ge r\})$ as $r \to \infty$ with respect to the growth of $t \mapsto \mathbb{E}_x[h \circ V(X_t)]$ as $t \to \infty$, since this expectation controls the tail of $h \circ V(X_t)$ via Markov's inequality (see also Remark 5.2 below for more details).

---

[3]In this paper we adopt the convention that $\log x$ equals 1 for $x \in (0, e)$ and the natural logarithm on $[e, \infty)$.

(b) Given a Lyapunov function V, our methods generate lower bounds on the f-variation distances with the property that the level sets of functions f form a subset of the level sets of V. This is analogous to results concerning the upper bounds using a Lyapunov function V, see e.g. [9, Thm 3.2].

(c) The **L**-drift condition $\mathbf{L}(V,\varphi,\Psi)$ is not restricted to Markov processes with a subexponential invariant measure and rates of convergence. Indeed, in Section 3.1.3 below, Theorems 2.1 and 2.2 yield exponential lower bounds both on the tails of the invariant measures and the convergence rates in a class of elliptic diffusion models where exponential upper bounds are known.

(d) By Theorem 2.1, the function $a(t)$ in Assumption (b) of Theorem 2.2 (with $f \equiv 1$) provides a lower bound on the tail of π. In applications it is often simpler to work with some function a rather than the actual lower bound on the tail of π provided by Theorem 2.1, since the rate of convergence in Theorem 2.2 is given in terms of the inverse of $t \mapsto A(t) = ta(t)$. In particular, under assumptions of Theorem 2.2, inequality (4) implies $0 < f_\star(r)c_{\varepsilon,q}/L_{\varepsilon,q}(r) \leq f_\star(r)\pi(V \geq r) \to 0$ as $r \to \infty$, making the growth of $A(t)$ (as $t \to \infty$) sublinear. This typically leads to subexponential decay of $a \circ A^{-1}$ and thus subexponential lower bounds on f-variation.

(e) The iterated logarithm term in (3), and thus in (5), is an artefact of the proof of Theorem 2.1, where the lower bounds on modulated moments are used to establish lower bounds on the invariant measure π. In all our examples in Section 3, the iterated logarithm term is negligible, suggesting that it will not affect other applications. While we cannot fully remove this term, it is possible to modify the proof so that only arbitrarily many (instead of two) iterations of the logarithm remain.

(f) The submultiplicative function [34, Def. 25.2] (cf. footnote on page 3) Ψ in $\mathbf{L}(V,\varphi,\Psi)$ transforms the process $V(X)$ into a submartingale. The function Ψ features in lower bounds in our theorems through definition (3). Since a product of submultiplicative functions is submultiplicative and, by [34, Prop. 25.4], $r \mapsto g(cr + \gamma)^\alpha$ (where $c, \gamma, \alpha > 0$) is submultiplicative if g is, it follows that $r \mapsto r^a(\log r)^d \exp(br^p)$ is submultiplicative for (i) $b > 0$ and $p \in (0,1)$, (ii) $b = 0$ and $a > 0$ or (iii) $b = a = 0$ and $d \geq 0$. This form is similar to the subgeometric rate functions used in [9, 12]. In fact, the rate functions in [9, 12] are submultiplicative [35, Eq. (5)], a crucial fact used in the proofs of [9, 12].

2.3. Modulated moments for the process.

A classical approach to the stability of Markov processes relies heavily on decomposing the path of the process as the sum of excursions from some petite set (see definition of a petite set in (28) in Section 5.2 below). Many results, including bounds on the invariant measure, the rate of convergence and moderate deviations have been established using such decomposition (see [10] and the references therein for more details). It is thus natural for quantitative bounds on the tail of the *modulated moments* (i.e. expectations of additive functionals of excursions from petite sets) to contain essential information necessary for bounding the rate of convergence to the invariant measure and related quantities.

The upper bounds on the modulated moments are well understood, see e.g. [9, 10]. In contrast, very little is known about the lower bounds on modulated moments. There are some known results regarding the finiteness of return time moments, e.g. [23], which however are not sufficiently strong to either characterize the tail behavior of the return time or complement the findings of [9, Thm 4.1] for general ergodic Markov processes. In the following theorem, we both generalize and strengthen

these results by accommodating a broader range of processes and providing lower bounds for the tail behavior. Moreover, in Section 3, we show that, for a wide range of models used in applications, our lower bounds on the return times match the upper bounds from [9].

For a set $D \in \mathcal{B}(\mathcal{X})$ and $\delta \in (0, \infty)$, let $\tau_D(\delta) := \inf\{t > \delta : X_t \in D\}$ (with convention $\inf \emptyset := \infty$) be the first hitting time of D after time δ (recall that $\tau_D(\delta)$ is an $(\mathcal{F}_t)$-stopping time by [16, Thm 1.27]). In Theorem 2.5 we provide, under the **L**-drift condition $\mathbf{L}(V, \varphi, \Psi)$, lower bounds on the tails of return times to arbitrary subsets of sublevel sets of the Lyapunov function V. Crucially, in Section 5.2 below, we prove that, under the **L**-drift condition, all petite sets are contained in the sublevel sets of V, making Theorem 2.5 applicable to any petite set of X.

Theorem 2.5. *Let Assumption $\mathbf{L}(V, \varphi, \Psi)$ hold. Consider a set $D \in \mathcal{B}(\mathcal{X})$, with $D \subset \{V \leq m\}$ for some $m \in (1, \infty)$, and fix $q \in (0, 1)$ and $\varepsilon = (1 - q)/2$. Then the following statements hold.*
(a) Let $h : [1, \infty) \to [1, \infty)$ be a non-decreasing continuous function and G_h the inverse of the increasing continuous function $v \mapsto \varepsilon h(v)/(v\varphi(1/v))$ on $[1, \infty)$. Then for every $x \in \mathcal{X}$ there exist constants $C, r_0, \delta \in (0, \infty)$, such that

$$\mathbb{P}_x \left(\int_0^{\tau_D(\delta)} h \circ V(X_s) \mathrm{d}s \geq r \right) \geq \frac{C}{\Psi(2G_h(r)/(1 - q))} \quad \text{for all } r \in (r_0, \infty).$$

(b) Let G_1 be the inverse of the increasing continuous function $v \mapsto \varepsilon/(v\varphi(1/v))$ on $[1, \infty)$. Then for every $x \in \mathcal{X}$ there exist constants $C, \delta, t_0 \in (0, \infty)$, such that

$$\mathbb{P}_x(\tau_D(\delta) \geq t) \geq \frac{C}{\Psi(2G_1(t)/(1 - q))} \quad \text{for all } t \in (t_0, \infty).$$

Remark 2.6. The proof of Theorem 2.5 in Section 5 below in fact shows that the inequalities in Theorem 2.5 hold for every $\delta \in (0, \infty)$, satisfying $\mathbb{P}_x(V(X_\delta) > r) > 0$ for all $x \in \mathcal{X}$ and $r \in [1, \infty)$. This condition holds for all $\delta > 0$ in the models of Section 3 below, because their marginal distributions at positive times have full support with respect to the Lebesgue measure.

2.4. **How are Theorems 2.1, 2.2 and 2.5 applied in practice?** Continuous-time Markov processes, where upper bounds on the rate of convergence have been established, are typically Feller [9,12]. It is thus natural to give sufficient conditions for the assumptions of Theorems 2.1, 2.2 and 2.5 in terms of the infinitesimal characteristics of X expressed via its extended generator. In this section we first provide tools for verifying the assumptions of our main theorems using the generator of the process and then discuss their application in practice.

2.4.1. *Generators and drift conditions.* Following the monograph [8, Ch 1, Def (14.15)], let $\mathcal{D}(\mathcal{A})$ denote the set of measurable functions $g : \mathcal{X} \to \mathbb{R}$ with the following property: there exists a measurable $h : \mathcal{X} \to \mathbb{R}$, such that, for each $x \in \mathcal{X}$, $t \to h(X_t)$ is integrable $\mathbb{P}_x$-a.s. and the process

$$g(X) - g(x) - \int_0^{\cdot} h(X_s) \mathrm{d}s \quad \text{is a } \mathbb{P}_x\text{-local martingale.}$$

Then we write $h = \mathcal{A}g$ and call $(\mathcal{A}, \mathcal{D}(\mathcal{A}))$ the *extended generator* of the process X. Define the left limit at $t \in (0, \infty)$ of the process X by $X_{t-} := \lim_{s \uparrow t} X_s$ and $X_{0-} = X_0$. The following theorem provides a sufficient condition for the validity of Assumptions **(i)** and **(ii)** in the **L**-drift condition $\mathbf{L}(V, \varphi, \Psi)$.

Theorem 2.7. *Let a continuous $V : \mathcal{X} \to [1, \infty)$ satisfy $\limsup_{t \to \infty} V(X_t) = \infty$ $\mathbb{P}_x$-a.s., all $x \in \mathcal{X}$.*

(a) Let $\varphi : (0,1] \to \mathbb{R}_+$ be a non-decreasing, continuous function, such that $r \mapsto 1/(r\varphi(1/r))$ is increasing on $[1, \infty)$ and $\lim_{r \to \infty} 1/(r\varphi(1/r)) = \infty$. If $1/V \in \mathcal{D}(\mathcal{A})$ and there exist $b, \ell_0 \in (0, \infty)$ such that

$$(7) \qquad \mathcal{A}(1/V)(x) \leq \varphi(1/V(x)) + b\mathbb{1}\{V(x) \leq \ell_0\} \quad \text{for all } x \in \mathcal{X},$$

then the process in $\boldsymbol{L(V,\varphi,\Psi)(i)}$ is a supermartingale.

(b) Let $\Psi : [1, \infty) \to [1, \infty)$ be a differentiable, increasing, submultiplicative function. Assume X has bounded jumps: for some constant $d \in \mathbb{R}_+$, we have $\mathbb{P}_x(V(X_t) - V(X_{t-}) \leq d$ for all $t \in \mathbb{R}_+) = 1$ for each $x \in \mathcal{X}$. If $\Psi \circ V \in \mathcal{D}(\mathcal{A})$ and there exist $c, \ell_0 \in (0, \infty)$ such that

$$(8) \qquad \mathcal{A}(\Psi \circ V)(x) \geq -c\mathbb{1}\{V(x) \leq \ell_0\} \quad \text{for all } x \in \mathcal{X},$$

then inequality (2) in $\boldsymbol{L(V,\varphi,\Psi)(ii)}$ holds for every $\ell \in (\ell_0, \infty)$ and some constant $C_\ell \in (0, \infty)$.

If X has bounded jumps, Theorem 2.7 has a natural converse: Assumption $\mathbf{L}(V,\varphi,\Psi)(\mathbf{i})$ and the submartingale property in (9) of Lemma 2.8 below (which implies $\mathbf{L}(V,\varphi,\Psi)(\mathbf{ii})$) yield the inequalities in (7) and (8) involving $\mathcal{A}(1/V)$ and $\mathcal{A}(\Psi \circ V)$, respectively. As this fact is not used in the paper, the details are omitted.

We proceed with Lemma 2.8, which provides a key step in the proof of Theorem 2.7(b). We state it here because it is of independent interest in applications as it gives a sufficient condition for the bound on the exit probability in $\mathbf{L}(V,\varphi,\Psi)(\mathbf{ii})$ in terms of the submartingale condition in (9) below. Denote $t \wedge s := \min\{s, t\}$, $t, s \in \mathbb{R}_+$, and recall $T^{(r)} = \inf\{t \geq 0 : V(X_t) > r\}$, $r \in \mathbb{R}_+$.

Lemma 2.8. *Let a continuous $V : \mathcal{X} \to [1, \infty)$ satisfy $\limsup_{t \to \infty} V(X_t) = \infty$ $\mathbb{P}_x$-a.s., for all $x \in \mathcal{X}$. Let $\Psi : [1, \infty) \to [1, \infty)$ be a differentiable, increasing, submultiplicative function. Assume that for some $d \in \mathbb{R}_+$, we have $\mathbb{P}_x(V(X_t) - V(X_{t-}) \leq d$ for all $t \in \mathbb{R}_+) = 1$ for each $x \in \mathcal{X}$. If for some $\ell_0, c \in (0, \infty)$ and all $r \in (\ell_0, \infty)$, the process*

$$(9) \qquad \Psi \circ V(X_{\cdot \wedge T^{(r)}}) + c \int_0^{\cdot \wedge T^{(r)}} \mathbb{1}\{V(X_u) \leq \ell_0\}\mathrm{d}u \quad \text{is an } (\mathcal{F}_t)\text{-submartingale under } \mathbb{P}_x$$

for all $x \in \mathcal{X}$, then the condition $\boldsymbol{L(V,\varphi,\Psi)(ii)}$ holds with functions V and Ψ.

Remark 2.9. If X has jumps with heavy tails, a submartingale argument of Lemma 2.8 may fail to imply the inequality in Assumption $\mathbf{L}(V,\varphi,\Psi)(\mathbf{ii})$. This is because the overshoot of the process $\Psi \circ V(X)$ need not be integrable (see e.g. the class of models in Section 3.2 below). It is thus crucial that condition (ii) in Assumption $\mathbf{L}(V,\varphi,\Psi)$ is given in terms of the probability $\mathbb{P}_x(T^{(r)} < S_{(\ell)})$ directly, rather than the submartingale property of $\Psi \circ V(X)$. As demonstrated in Section 3.2, in such heavy-tailed cases it is possible to apply path-wise arguments directly to obtain the lower bound on $\mathbb{P}_x(T^{(r)} < S_{(\ell)})$, see Section 6.2 for details.

The expected growth condition assumed in Theorem 2.2(a) is easily verified via Lemma 2.10.

Lemma 2.10. *Let $H : \mathcal{X} \to [1, \infty)$ be continuous with $H \in \mathcal{D}(\mathcal{A})$ and let $\xi : [1, \infty) \to [1, \infty)$ be concave, non-decreasing and differentiable, satisfying $\mathcal{A}H \leq \xi \circ H$ on $\mathcal{X}$. Define the function*

$\Xi(u) := \int_1^u \mathrm{d}s/\xi(s)$ *for* $u \in [1, \infty)$. *Then we have*

$$\mathbb{E}_x[H(X_t)] \leq \Xi^{-1}(\Xi(H(x)) + t) \quad \text{for all } x \in \mathcal{X} \text{ and } t \in \mathbb{R}_+.$$

Lemma 2.10 plays a key role in establishing the upper bound on the expected growth of the process $h \circ V(X)$ in condition (a) of Theorem 2.2. If ξ is constant, the bound on $\mathbb{E}_x[h \circ V(X_t)]$ is linear in time. Crucially Lemma 2.10 permits an unbounded ξ, thus allowing the process $h \circ V(X)$ to exhibiting superlinear expected growth. This is key for establishing matching stretched exponential (see Sections 3.1.2 and 6.1.2 below) and exponential (Sections 3.1.3 and 6.1.3 below) lower bounds on the rate of convergence.

2.4.2. *How to find functions φ, Ψ (in $\boldsymbol{L(V,\varphi,\Psi)}$) and h (in Theorem 2.2(a)) for a given Lyapunov function V?* The classical Lyapunov drift condition [9, 12, 28] requires the process $V(X)$ to satisfy the supermartingale condition in (10) below. If such V is available, verifying $\mathbf{L}(V,\varphi,\Psi)$ reduces to finding scalar functions φ and Ψ, such that $1/V(X) - \int_0^{\cdot} \varphi \circ (1/V)(X_s)\mathrm{d}s$ and $\Psi \circ V(X)$ are a super- and a submartingale, respectively. Since Ψ is increasing, the Lyapunov function $V : \mathcal{X} \to [1, \infty)$ determines the level sets of $\Psi \circ V$, while $\Psi : [1, \infty) \to [1, \infty)$ modulates only the growth of $\Psi \circ V$. Thus, identifying Ψ is typically straightforward if V is given and the task is to find the slowest growing Ψ so that $\Psi \circ V(X)$ is a submartingale. Identifying φ for a given V is also typically a simple task, as it reduces to bounding the drift of $1/V(X)$ using Theorem 2.7(a) in Subsection 2.4.1.

The lower bound on the rate in Theorem 2.2 is based on the comparison of the tails of π and the law of X_t. Recall from Remark 2.4(a) that the role of the function h in assumption (a) of Theorem 2.2 is to balance a lower bound on $r \mapsto r\pi(\{h \circ V \geq r\})$ and an upper bound on the $t \mapsto \mathbb{E}_x[h \circ V(X_t)]$. Typically, a good choice for h is such that $r \mapsto r\pi(\{h \circ V \geq r\})$ grows to infinity as $r \to \infty$ at a polynomial (necessarily sublinear) rate, see Remark 5.2 below for more details. Lemma 2.10, applied to $H = h \circ V$, yields the desired upper bound on $t \mapsto \mathbb{E}_x[h \circ V(X_t)]$ (and hence by Theorem 2.2 a lower bound on the rate of convergence).

2.5. **Sketch of the proofs of the main results: Theorems 2.1, 2.2 and 2.5.** The $\mathbf{L}$-drift condition $\mathbf{L}(V,\varphi,\Psi)$ is the crucial ingredient of all the main theorems. The following implications constitute key steps in their proofs:
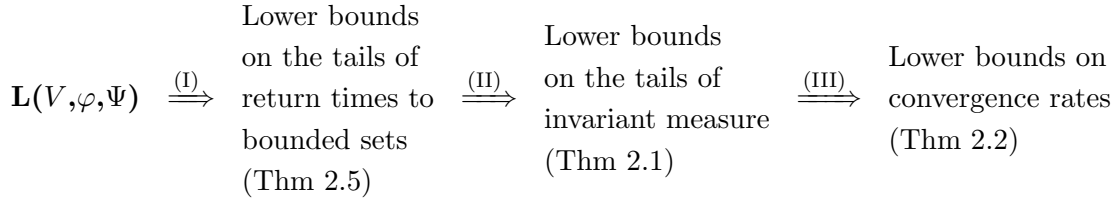
$$\mathbf{L}(V,\varphi,\Psi) \overset{(I)}{\Longrightarrow} \begin{matrix} \text{Lower bounds} \\ \text{on the tails of} \\ \text{return times to} \\ \text{bounded sets} \\ \text{(Thm 2.5)} \end{matrix} \overset{(II)}{\Longrightarrow} \begin{matrix} \text{Lower bounds} \\ \text{on the tails of} \\ \text{invariant measure} \\ \text{(Thm 2.1)} \end{matrix} \overset{(III)}{\Longrightarrow} \begin{matrix} \text{Lower bounds on} \\ \text{convergence rates} \\ \text{(Thm 2.2)} \end{matrix}$$

TABLE 1. The sequence of implications from $\mathbf{L}$-drift conditions to convergence rates.

We now describe informally each of these implications. Implication (I) in Table 1 above requires a lower bound on the return time of X to a sublevel set $\{V \leq \ell\}$ of the Lyapunov function V (see Figure 1). Under the $\mathbf{L}$-drift condition $\mathbf{L}(V,\varphi,\Psi)(\mathbf{i})$, the process $1/V(X) - \int_0^{\cdot} \varphi \circ (1/V)(X_s)\mathrm{d}s$ is a supermartingale when $V(X)$ is above level ℓ, implying an upper bound on the tail of the supremum of $1/V(X)$. This upper bound can be converted into a lower bound on the tail of the return time
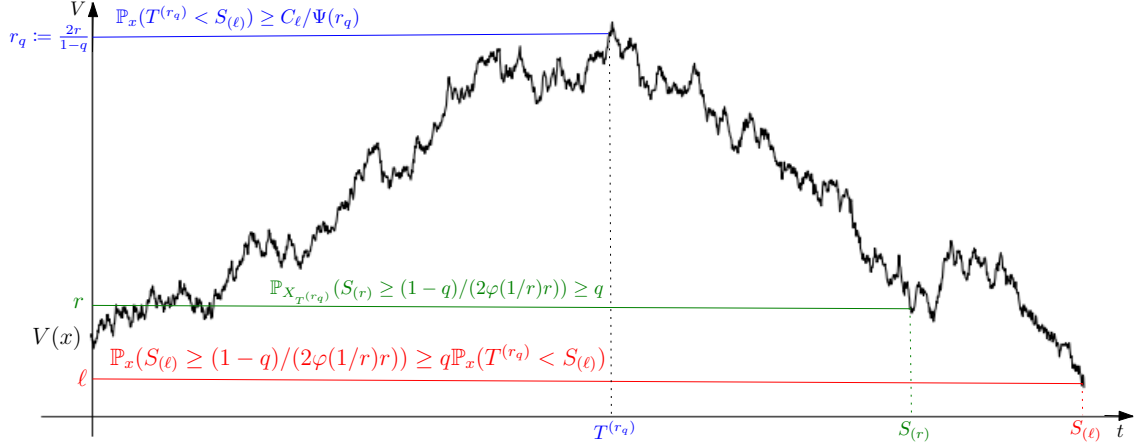
FIGURE 1. Establishing the lower bound in Theorem 2.5 (implication (I)) requires conditions **(i)** and **(ii)** in Assumption $\mathbf{L}(V,\varphi,\Psi)$ to obtain lower bounds on the *duration* $S_{(\ell)}$ and the *height* $\sup_{s\in[0,S_{(\ell)}]} V(X_s)$ of an excursion from the compact set $\{V \leq \ell\}$. The bound on the height (blue inequality) follows directly from $\mathbf{L}(V,\varphi,\Psi)\mathbf{(ii)}$, while the lower bound on the duration (red inequality) requires a more involved argument (in Lemma 4.2 below) using the supermartingale in $\mathbf{L}(V,\varphi,\Psi)\mathbf{(i)}$.

of $V(X)$ below a level ℓ. The argument requires only the supermartingale property, see Lemma 4.2 below for details. In short, we first wait until the process $V(X)$, started from $V(x) > \ell$, reaches a large level $r_q = 2r/(1-q)$ for $r \gg \ell$ and any $q \in (0,1)$, before descending below ℓ. By $\mathbf{L}(V,\varphi,\Psi)\mathbf{(ii)}$ we have the blue lower bound on the exit probability in Figure 1. On this event, once the process $V(X)$ is at $V(X_{T^{(r_q)}})$, we apply the supermartingale property in $\mathbf{L}(V,\varphi,\Psi)\mathbf{(i)}$ to obtain the green lower bound on the tail of the return time $S_{(r)}$, which implies the red lower bound. Since $q \in (0,1)$ is arbitrary, these estimates imply boundedness of any petite set of X (Lemma 5.4 below) as well as lower bounds on the tail probability $\mathbb{P}_x \left(\int_0^{\tau_D(\delta)} h \circ V(X_s)\mathrm{d}s \geq r \right)$ of the additive functional of the excursion from any bounded set D, for $x \in D$, $\delta > 0$ and $\tau_D(\delta) = \inf\{t > \delta : X_t \in D\}$ (Theorem 2.5).

Implication (II) in Table 1 above uses the fact that, under $\mathbf{L}(V,\varphi,\Psi)$, all petite sets of X are bounded and, crucially, that the lower bounds on the probability $\mathbb{P}_x \left(\int_0^{\tau_D(\delta)} h \circ V(X_s)\mathrm{d}s \geq r \right)$ hold for *all* non-decreasing functions $h : [1,\infty) \to [1,\infty)$. These facts, together with the well-known characterisation of Meyn and Tweedie in [25, Thm 1.2(b)] of the integrability with respect to π, yield a lower bound on the tail of the invariant measure π in Theorem 2.1.

Implication (III) in Table 1 above requires a comparison of the lower bound on the tail of the invariant measure π, obtained via implication (II), with the upper bound on the tail of the law of $V(X_t)$, controlled by the expected growth of the process $V(X)$. Finally, an application of Lemma 5.1, which is generalisation to f-variation norms of [15, Thm 3.6] (see also [13, Thm 5.1]), yields our lower bound on the rate of convergence in Theorem 2.2.

2.6. **Related literature.**

Lyapunov functions and lower bounds. A general Lyapunov-function approach to lower bounds on the total variation distance along a sequence of times tending to infinity is developed in [15]. A key step [15, Thm 3.6] consists of converting lower bounds on the tails of the invariant measure to the lower bounds on the total variation, an idea also exploited in the present paper (see its generalisation from total to f-variation in Lemma 5.1 below). Crucially, in [15, Thm 3.2], Hairer gives criteria (based on ideas from [38]) for establishing lower bounds on total variation along a sequence of times tending to infinity when the invariant measure is not explicitly known. These criteria can be summarised briefly as follows: let W_1 and W_2 be C^2-functions on $\mathbb{R}^n$ such that $W_1/W_2 \to 0$ as $W_1 \to \infty$ and the inequalities $\mathcal{A}W_1 \geq 0$ and $\mathcal{A}W_2 \leq F$ hold outside a compact set for some function F, where $\mathcal{A}$ denotes the generator of X. Under these conditions [15, Thm 3.2] implies $\int F \mathrm{d}\pi = \infty$ and, together with [15, Thm 3.6], yields a sequence of times $t_n \to \infty$ at which a lower bound on the total variation (based on non-integrability of F) holds.

The critical step in [15] consists of finding a non-integrable function F, which essentially amounts to the function $\mathcal{A}W_2$ being non-integrable: $\int \mathcal{A}W_2 \mathrm{d}\pi = \infty$. Consequently, the lower bounds on the tails of π hold only along a (possibly sparse) sequence of levels, yielding a (possibly sparse) sequence of times $t_n \to \infty$ at which the lower bound on the total variation can be established. In contrast, our approach provides lower bounds for all times by essentially analysing lower bounds as upper bounds of $1/V(X)$ (as usual V denotes a Lypuanov function for X). In the language of [15] this involves estimating the drift $\mathcal{A}(1/V)$, and then using *path-wise* arguments for semimartingales (Lemma 4.2 below) to characterise $\int h \circ V \mathrm{d}\pi = \infty$ for *all* increasing scalar functions h. This idea allows us to establish lower bounds on the tails of π for all levels and hence get lower bounds on the convergence rate in f-variation for all times. Moreover, this suggests that establishing lower bounds for all times using the approach in [15] directly would at the very least require a sufficiently rich family of functions $W_2^{(\kappa)}$, satisfying the assumptions of [15, Thm 3.2] (and hence $\int \mathcal{A}W_2^{(\kappa)} \mathrm{d}\pi = \infty$) for all values of the parameter κ.

In the context of elliptic diffusions studied in Section 3.1 below, our results may be viewed as a generalisation of the bounds along $t_n \to \infty$, obtained via [15, Thm 3.2], to all times $t \in [1, \infty)$. If the process X has jumps with heavy tails as in Section 3.2 below, [15, Thm 3.2] appears to be difficult to apply: for the Lévy-driven stochastic differential equation of Section 3.2, W_1 is either not integrable with respect to the jump measure of X or, if it is, we have $\mathcal{A}W_1 < 0$ outside of a compact (see the formula for $\mathcal{A}$ in equation (44) and Remark 6.6 below for details). This makes the condition $0 \leq \mathcal{A}W_1 < \infty$ hard to satisfy outside of any compact set. Finally, it is feasible that our methods could yield novel insights for the hypoelliptic Hamiltonian system studied in [15] and possibly extend the result on lower bounds in [15, Thm 1.1] to all times. The function W_1 in [15, eq. (5.15)], satisfying $\mathcal{A}W_1 \geq 0$ and thus making $W_1(X)$ a submartingale, is a natural candidate for $\Psi \circ V$ in the **L**-drift condition **L(V,φ,Ψ)(ii)**. Determining the function V and estimating the asymptotic behaviour of $\mathcal{A}(1/V)$, so that **L(V,φ,Ψ)(i)** holds, would enable the application of our results. This is left for future research.

Lyapunov functions and upper bounds. The rate of convergence to invariant measures of ergodic Markov processes has been studied extensively. The majority of the modern literature, based on a probabilistic approach using Lyapunov functions, dates back to the seminal work of Meyn and Tweedie in the 1990s [25, 27, 28] and primarily focuses on the upper bound estimates.

These results have since been further improved and generalised. Notable contemporary versions can be found in [2, 9] and Hairer's lecture notes [13], see also monograph [11] and the references therein. Briefly, if for some continuous $V : \mathcal{X} \to [1, \infty)$, increasing, differentiable, concave $\phi : [1, \infty) \to [1, \infty)$, a closed petite set (defined in (28) below) $A \in \mathcal{B}(\mathcal{X})$ and a constant $b \in \mathbb{R}_+$, the process

$$(10) \qquad V(X) + \int_0^{\cdot} \phi \circ V(X_u) \mathrm{d}u - b \int_0^{\cdot} \mathbb{1}\{X_u \in A\} \mathrm{d}u \qquad \text{is a supermartingale,}$$

the following statements hold (see e.g. [9]):

(a) the process X is Harris recurrent with invariant measure π and $\int_{\mathcal{X}} \phi \circ V(x) \pi(\mathrm{d}x) < \infty$;

(b) $r_*(t) \| \mathbb{P}_x(X_t \in \cdot) - \pi(\cdot) \|_{\mathrm{TV}} \leq C V(x)$, where $r_*(s) := \phi \circ H_\phi^{-1}(s)$ and $H_\phi(u) = \int_1^u \frac{\mathrm{d}s}{\phi(s)}$, $u \geq 1$.

Since this approach relies on transforming the state space with a Lyapunov function V, key information may be lost, potentially resulting in poor upper bound estimates on the convergence rate (see the motivating example in Section 3.1 below; see also Example 3.5). This naturally motivates a general study of *lower* bounds in the context of Lyapunov functions presented in this paper, particularly in the case when the upper bounds on the rate of convergence are subexponential. As explained in Section 2.4, our results naturally augment the existing Lyapunov function approach for the stability of Markov processes and provide a robust method for checking the quality of upper bound estimates obtained via a given Lyapunov function.

Poincaré inequalities and the rates of convergence. These functional analytic techniques typically work directly with the infinitesimal generator and are thus not dependent on the potentially suboptimal choice of a Lyapunov function (see e.g. [2, 30]). However, due to their analytical nature, such methods often require more restrictive (typically global rather than local) assumptions on the behavior of the transition operator of the process. Additionally, even under strong assumptions (e.g. reversibility), the literature addressing lower bounds remains sparse.

Potential theory and Lévy-driven Ornstein-Uhlenbeck (OU) processes. A potential theoretic approach for the stability of a class of Lévy-driven OU-processes, arising as limits of multiclass many-server queues, has been developed in [1, 32]. In contrast to our results, the lower bounds [32, Thm 1.2] for the general case of their models hold along a sequence of times $t_n \to \infty$. The theory in [32] gives no information on the sparsity of the sequence $(t_n)_{n \in \mathbb{N}}$. However, in some special cases of the models in [1, 32], matching lower and upper bounds for all $t \in [1, \infty)$ are established using Hairer's result [13, Thm 5.1]. In these special cases, the precise decay of the tails of the invariant measure of the model in [1] is established using analytical methods.

Mixing coefficients of Rosenblatt and Kolmogorov. Convergence to stationarity of an ergodic Markov process can also be quantified via Rosenblatt's and Kolmogorov's mixing coefficients. Veretennikov [36] provides upper bounds on the coefficients for polynomialy ergodic elliptic diffusions, cf. Section 3.1.1 below. In the special case when the noise is additive, the corresponding lower bounds on the coefficients are given in [37] (for discrete-time additive-noise case see [18]).

Markov Chain Monte Carlo. The vast majority of the convergence theory for Markov Chain Monte Carlo algorithms concerns upper bounds, see e.g. [11] and the references therein. In comparison, lower bounds on the rate of convergence are sparse. Some recent model dependent results

are given in [5,6]. However, the tools used appear to be hard to extend to general ergodic Markov processes discussed in this paper.

Ergodic theory. Convergence of a Markov process to its invariant measure is closely related to the decay of correlations and the mixing of dynamical systems in ergodic theory. In this setting, seminal papers [22,33] establish asymptotically matching polynomial upper and lower bounds on the mixing rates using *operator renewal theory*. This approach extends a classical probabilistic result of Rogozin, bounding the remainder term in the renewal theorem [31].

It would be natural to consider extending Rogozin's approach to obtain lower bounds on the convergence rates of Markov processes towards their invariant measures. While this might be achievable, as suggested in [24], employing renewal theory would likely require much stronger assumptions on the process than in the present paper. Moreover, in this context, lower bounds on the rates of convergence can be established only if both upper and lower bounds on the model parameters are assumed to match. In contrast, as demonstrated by the examples in Section 3, our approach yields asymptotically matching lower bounds on the rate of convergence even for the models with oscillating parameters. Adapting the techniques from [22,33] to the context of ergodic Markov processes studied in this paper would appear to lead to weaker results.

3. Applications to elliptic diffusions, Lévy-driven SDEs and hypoelliptic stochastic damping Hamiltonian systems

We now apply the results of Section 2 to stochastic models used across probability and its application. We mostly consider models where Lyapunov drift conditions have been developed to obtain upper bounds on the rates of convergence. The examples are chosen to demonstrate the robustness and the complementary nature of our approach to the existing theory for upper bounds: we cover the classes of models studied in [9, Sec. 5] and [12, Sec. 3]. In particular, we analyse models exhibiting polynomial, stretched exponential and exponential ergodicity and in all cases provide lower bounds, which asymptotically match known upper bounds. More precisely, we give lower bounds in f-variation for polynomially ergodic elliptic and hypoelliptic diffusions, studied in [12, Sec. 3] and [9,39], respectively. For brevity, in all other examples we consider total variation (even though our methods could handle f-variation in these cases) and show that our lower bounds match asymptotically the upper bounds in [9,12].

The proofs of the results in this section, contained in Section 6 below, typically involve verifying Assumption $\mathbf{L}(V,\varphi,\Psi)$ via Theorem 2.7 and the condition in (a) of Theorem 2.2 via Lemma 2.10. In all examples of this section, the dependence of the multiplicative constant on the initial position $x \in \mathcal{X}$ can be obtained explicitly from our main estimate in (5) of Theorem 2.2. However, for ease of presentation, the explicit dependence on the starting point has been omitted.

3.1. **Elliptic diffusions.** In this section we apply our results to obtain lower bounds on the rate of convergence to stationarity of elliptic diffusions. We first introduce the general form of the model and give a simple motivating example. In Sections 3.1.1 and 3.1.2 we discuss the polynomial and stretched exponential cases, respectively. We stress that our assumptions allow multiplicative noise with unbounded instantaneous variance of the process. Example 3.5 of Section 3.1.1 demonstrates the necessity of two-sided asymptotic assumptions on the coefficients (used in this section) for

obtaining the actual rate of convergence to stationarity. It also shows that upper bounds on the rate of convergence, established in [9,12,36] using asymptotic upper bounds on the coefficients only, may be orders of magnitude larger than the actual convergence rate.

Section 3.1 concludes with the class of elliptic diffusions exhibiting exponential ergodicity (see Section 3.1.3 below). Even though our main focus is on subexponential ergodicity, Section 3.1.3 shows that our methods for lower bounds are also applicable to the exponentially ergodic case.

For $n \in \mathbb{N}$, let the process $X = (X_t)_{t \in \mathbb{R}_+}$ with state space $\mathcal{X} := \mathbb{R}^n$ be the unique strong solution of the stochastic differential equation (SDE)

$$(11) \qquad X_t = X_0 + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}B_s,$$

where the functions $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ are locally Lipschitz (i.e. for every $l > 0$ there exists a finite constant c_l such that $|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq c_l |x - y|$ for all $x, y \in \mathbb{R}^n$ with $|x|, |y| < l$), and $(B_t)_{t \in \mathbb{R}_+}$ a standard n-dimensional Brownian motion. Here and throughout we denote by $|\cdot|$ and $\langle \cdot, \cdot \rangle$ the Euclidean norm and standard scalar product on $\mathbb{R}^n$, respectively. In particular, $|x|^2 = \langle x, x \rangle$ for all $x \in \mathbb{R}^n$. Let $\Sigma := \sigma\sigma^\mathsf{T}$, where σ^T is the transpose of the matrix σ, be uniformly elliptic: $\langle \Sigma(x)y, y \rangle \geq \delta_a |y|^2$ for some $\delta_a > 0$ and all $y, x \in \mathbb{R}^n$. The diffusion X in (11) is assumed to be ergodic with an invariant measure π.

Motivating example: Consider the following toy example of an elliptic SDE (11) with $n = 1$ (i.e. $\mathcal{X} = \mathbb{R}$), $\sigma \equiv 1$, and $b(x) = -x/|x|$ for all $|x| > 1$. Applying the generator $\mathcal{A}$ of X to the Lyapunov function $V(x) = 1 + x^2$ yields $\mathcal{A}V(x) \leq -V(x)^{1/2}$ for all $x \in \mathbb{R}$ with $|x|$ large. Then, for every initial condition $x \in \mathbb{R}$, an application of the drift condition from [9, Thms 3.2 and 3.4] (see (10) above) with $V(x) = 1 + x^2$ (and $\phi(r)$ proportional to $\sqrt{r}$ as $r \to \infty$) implies the following upper bound: $\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq C_x/t$ for all $t \in [1, \infty)$ and some constant $C_x \in (0, \infty)$. Given that the process X is in fact exponentially ergodic (use [9, Thms 3.2 and 3.4] with $V(x) = \exp(|x|)$ for large $|x|$), it is evident that the upper bounds obtained through Lyapunov drift condition (10) may be exceedingly inaccurate. Furthermore, this example demonstrates that the classical Lyapunov-function theory for upper bounds is not sufficient for characterising the processes that are subexponentially ergodic. We now apply the results in Section 2 to address this problem by establishing the actual subexponential rate of convergence using our Lyapunov-function drift conditions in $\mathbf{L}(V, \varphi, \Psi)$.

3.1.1. *Polynomial tails.* The following assumption ensures polynomial upper bounds on the tails of the invariant measure and the rate of convergence in total variation [12, Sec 3.2], [36]. Recall that a function $g : \mathbb{R}^n \to \mathbb{R}$ satisfies $g(x) = o(1)$ as $|x| \to \infty$ if $\lim_{|x| \to \infty} \sup_{u \in \mathbb{R}^n, |u|=1} g(u|x|) = 0$.

Assumption $\mathbf{A}_p$. *There exist $\alpha, \beta, \gamma \in (0, \infty)$ and $\ell \in [0, 2)$, such that $2 - \ell < 2 + (2\alpha - \gamma)/\beta =: m_c$ and the coefficients b and $\Sigma = \sigma\sigma^\mathsf{T}$ of (11) satisfy (as $|x| \to \infty$)*

$$\langle b(x), x/|x| \rangle/|x|^{\ell-1} = -\alpha + o(1), \quad \langle \Sigma(x)x/|x|, x/|x| \rangle/|x|^\ell = \beta + o(1), \quad \mathrm{Tr}(\Sigma(x))/|x|^\ell = \gamma + o(1).$$

As defined in [12], Langevin tempered diffusion on $\mathbb{R}^n$, given by a smooth $\widetilde{\pi} : \mathbb{R}^n \to (0, \infty)$ (proportional to the density of π on $\mathbb{R}^n$), satisfies SDE (11) with coefficients $\Sigma(x) = I_n/\widetilde{\pi}(x)^{2d}$, where I_n is the identity matrix and $d \geq 0$, and $b(x) = (1 - 2d)\nabla(\log \widetilde{\pi})(x)/(2\widetilde{\pi}(x)^{2d})$, where ∇ denotes the gradient. If $d = 0$, we get the classical Langevin diffusion with bounded volatility. This is an important class of diffusions because, by construction, their invariant measure equals π.

They satisfy Assumption A_p when $1/\widetilde{\pi}(x)^{2d}$ is proportional to $|x|^\ell$, with $\ell \in [0, 2)$, and $\langle \nabla\widetilde{\pi}(x), x \rangle$ is proportional to $\widetilde{\pi}(x)$ as $|x| \to \infty$.

When π has polynomial tails, upper bounds for the convergence in total variation for Langevin tempered diffusions on $\mathbb{R}^n$ (as well as their generalisation in Assumption A_p) were studied in [12, Sec 3.2]. In this section we apply our methods under Assumption A_p to obtain matching lower bounds for the results in [12], establishing the rate of convergence to stationarity.

Theorem 3.1. *Let Assumption A_p hold. Pick $k \in [0, \ell + (2\alpha - \gamma)/\beta)$ and fix the critical exponent $\alpha_k := m_c/(2-\ell) - 1 - k/(2-\ell) = (2(\alpha+\beta) - \gamma)/((2-\ell)\beta) - 1 - k/(2-\ell)$. Then for the function $f_k(x) := 1 + |x|^k$, a starting point $x \in \mathbb{R}^n$ and $\varepsilon > 0$, there exists a constant $c_{k,\varepsilon} \in (0, \infty)$ such that*

$$c_{k,\varepsilon}/t^{\alpha_k + \varepsilon} \le \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{f_k} \quad \text{for all } t \in [1, \infty).$$

The next result provides lower bounds on the tail of the invariant measure π and the return time $\tau_D(\delta) = \inf\{t > \delta : X_t \in D\}$ of the diffusion X to a bounded set $D \in \mathcal{B}(\mathbb{R}^n)$ after time $\delta > 0$.

Theorem 3.2. *Let Assumption A_p hold and recall $0 < 2 - \ell < m_c = 2 + (2\alpha - \gamma)/\beta$.*

(a) For every $\varepsilon > 0$, there exists $c_\pi \in (0, \infty)$ such that

$$c_\pi/r^{m_c + \ell - 2 + \varepsilon} \le \pi(\{|x| \ge r\}) \qquad \text{for all } r \in [1, \infty).$$

(b) For any $x \in \mathbb{R}^n$, any bounded set $D \in \mathcal{B}(\mathbb{R}^n)$ and arbitrary $\varepsilon, \delta \in (0, \infty)$, there exists a constant $c_\tau \in (0, \infty)$, such that

$$c_\tau/t^{m_c/(2-\ell) + \varepsilon} \le \mathbb{P}_x(\tau_D(\delta) \ge t) \qquad \text{for all } t \in [1, \infty).$$

The proofs of Theorems 3.1 and 3.2 use the **L**-drift condition $\mathbf{L}(V, \varphi, \Psi)$, where V, φ and Ψ exhibit polynomial growth. The functions φ and Ψ are obtained directly from the generator inequalities of Theorem 2.7 and are also polynomial, see Section 6.1.1 below for details.

Remark 3.3 (**matching rates**). As mentioned above, our lower bounds in Theorem 3.1 on the f-variation distance matches the upper bounds in [12]. Recall from Theorem 3.1 the parameters k, α_k and the function f_k. Then, for every $\varepsilon > 0$, there exist constants $c_{k,\varepsilon}, C_{k,\varepsilon} \in (0, \infty)$ such that

$$(12) \qquad c_{k,\varepsilon}/t^{\alpha_k + \varepsilon} \le \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{f_k} \le C_{k,\varepsilon}/t^{\alpha_k - \varepsilon} \quad \text{for all } t \in [1, \infty).$$

Analogous bounds could be established for the tails of the invariant measure π and the return time $\tau_D(\delta)$. Put differently, in conjunction with [12], Theorems 3.1 and 3.2 imply the convergence rate to stationarity and the decay of the tails of the invariant measure and the return times in logarithmic scale:

$$\lim_{t \to \infty} \frac{\log \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot))\|_{f_k}}{\log t} = \alpha_k, \qquad \lim_{t \to \infty} \frac{\log \mathbb{P}_x(\tau_D(\delta) > t)}{\log t} = m_c/(2-\ell),$$

$$\lim_{t \to \infty} \frac{\log \pi(\{|x| \ge r\})}{\log r} = m_c + \ell - 2.$$

Remark 3.4 (**oscillating coefficients**). Note that Assumption A_p requires matching upper and lower bounds on the asymptotic behaviour (as $|x| \to \infty$) of the coefficients of SDE (11). The

following is a natural generalisation of Assumption A_p, allowing oscillating coefficients: for all $x \in \mathbb{R}^n$ outside of some compact set it holds that $\gamma_L \leq \mathrm{Tr}(\Sigma(x))/|x|^\ell \leq \gamma_U$,

$$-\alpha_L \leq \langle b(x), x/|x| \rangle/|x|^{\ell-1} \leq -\alpha_U \qquad \& \qquad \beta_L \leq \langle \Sigma(x)x/|x|, x/|x| \rangle/|x|^\ell \leq \beta_U$$

for some constants $\gamma_L, \gamma_U, \alpha_L, \alpha_U, \beta_L, \beta_U \in (0, \infty)$. Under this assumption, the lower bounds in Theorem 3.1 and 3.2 remain valid with with $m_c^L = 2 + (2\alpha_L - \gamma_L)/\beta_L$ (instead of m_c). While a slight modification of the proofs of Theorems 3.1 and 3.2 is necessary, the same polynomial class of Lyapunov functions can be used. The details are omitted for brevity. The upper bounds depend on $m_c^U = 2 + (2\alpha_U - \gamma_U)/\beta_U$ and can be obtained from the classical theory [12]. The discrepancy between $m_c^L > m_c^U$ leads to an asymptotic gap between the lower and upper bounds on the rate of convergence. As demonstrated by the following example, this gap is not an artefact of our methods. Example 3.5 below also shows that upper bounds on the coefficients alone, assumed in [12, Remark after Thm 16], are insufficient to deduce the rate of convergence in the sense of (12).

Example 3.5. Let the function $\pi : \mathbb{R} \to (0, \infty)$ be a positive thrice continuously differentiable density on $\mathbb{R}$ (up to a normalizing constant). The one-dimensional SDE

$$(13) \qquad\qquad dX_t = (\log \pi)'(X_t)dt + \sqrt{2}dB_t.$$

is an example of an elliptic diffusion of the form (11) in $\mathbb{R}$.

Assumption A_p above is given in terms of the limits of the coefficients of SDE (11) as $|x| \to \infty$. This differs from the assumptions in [12, Sec 3.2], where only the asymptotic upper bounds on the model parameters are assumed. In particular, in the context of SDE (13), the results from [12, Sec 3.2] imply the following: if $\limsup_{|x| \to \infty} x(\log \pi)'(x) = -\alpha$ for some $\alpha \in (1, \infty)$, then for every $x \in \mathbb{R}$ and $\varepsilon > 0$ there exists a constant $C > 0$, such that $\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq C/t^{(\alpha-1)/2-\varepsilon}$ holds for all $t \in [1, \infty)$. However, as we shall now see, an upper bound on the drift in (13) is not sufficient to determine the actual rate of convergence to the invariant measure.

Fix $\alpha \in (1, \infty)$. Then for every $k \in (\alpha, \infty)$, there exists an invariant density π, such that the following statements hold: $\limsup_{|x| \to \infty} x(\log \pi)'(x) = -\alpha$ and, for every $x \in \mathbb{R}$ and $\varepsilon > 0$, there exist constants $c', C' \in (0, \infty)$ such that

$$(14) \qquad c'/t^{(k-1)/2} \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq C'/t^{(k-1)/2-\varepsilon} \quad \text{for all } t \in [1, \infty),$$

where the process X follows SDE (13) with $X_0 = x$. The proof of (14) is in Section 6.1.1 below.

This example demonstrates that, assuming only upper bounds on the coefficients of the diffusion in (11) (as in [12, Sec 3.2]), may result in the upper bound on the rate being greater than the actual rate of convergence by any polynomial order. In contrast, if we assume matching lower and upper bounds on the drift of (13) (i.e. $(\log \pi)'(x) = -\alpha/x(1 + o(1))$ as $|x| \to \infty$), the bounds in (12) imply (14) with $k = \alpha$.

Remark 3.6 (**Can the drift in** (11) **point away from the origin?**). In Assumption A_p we stipulate that the drift b, while possibly vanishing at infinity, assymptotically points towards the origin. It is natural to ask whether this assumption can be relaxed by allowing the drift b to oscillate between pointing towards and away from the origin as $|x| \to \infty$.[4] An example of such a process is

---

[4]We thank the anonymous referee for drawing our attention to this natural question.

given in [20, Example 2.5] by a Langevin diffusion: for a constant $c \in (0, 1)$, let X follow the SDE

(15) $\mathrm{d}X_t = b(X_t)\mathrm{d}t + \mathrm{d}B_t$,   where $b(x) = -(\cos(x) + c)x/|x|$ for all $x \in \mathbb{R}$ with large $|x|$.

In [20] it is proved that $\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq C_\alpha t^{\alpha/(1-\alpha)}$ for any $\alpha \in (0, 1)$ and all $t \in [1, \infty)$ (and some constant $C_\alpha > 0$). It is natural to enquire whether a subexponential lower bound can be obtained for such a process.

This question is highly non-trivial in general since a sharp upper bound on the rate of convergence is not clear. The answer to it will depend on the details of the model. More specifically, the diffusion in (15) is in fact exponentially ergodic,

(16) $\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq C_0 \exp(-C_1 t)$ for constants $C_0, C_1 \in (0, \infty)$ and all $t \in [1, \infty)$,

as can be seen by a direct calculation using an appropriate Lyapunov function (see end of Section 6.1.1 below), related to the one used in Example 3.5 above. In particular, the diffusion in (15) demonstrates that achieving (exponential) ergodicity does not require the drift to be directed towards the origin everywhere outside of a compact set. As the diffusion X in (15) exhibits exponential ergodicity, subexponential lower bounds cannot be established in this example.

3.1.2. *Stretched exponential tails.* Under the following assumption, the process X exhibits stretched exponential ergodicity.

Assumption $\mathbf{A}_{se}$. *There exist constants $p \in (0, 1)$ and $\ell \in [0, 2p)$, such that the coefficients b and $\Sigma = \sigma\sigma^{\mathsf{T}}$ in (11) satisfy the following asymptotic assumptions:* $\limsup_{|x| \to \infty} \mathrm{Tr}(\Sigma(x))/|x|^\ell < \infty$ *and*

$$-\alpha_L \leq \langle b(x), x/|x|\rangle/|x|^{\ell-p} \leq -\alpha_U \qquad \& \qquad \beta_L \leq \langle \Sigma(x)x/|x|, x/|x|\rangle/|x|^\ell \leq \beta_U,$$

for all x outside of some compact set and some constants $\alpha_L, \alpha_U, \beta_L, \beta_U \in (0, \infty)$.

Assumption A_{se} covers tempered Langevin diffusions exhibiting stretched exponential ergodicity, studied in [9, Sec 5.2]. In particular, parameters $\beta \in (0, 1)$ and $d \in [0, 1/\beta - 1)$ (in the notation of [9, Sec 5.2]), are included in the cases $p = \beta - 1$ and $\ell = 2\beta d$. Moreover, the case $\ell = 0$ includes Langevin diffusions.

The main result, Theorem 3.7 below, establishes lower bounds on the tail of the invariant measure π, the rate of convergence towards the invariant measure in the TV-distance and the tail of the return time $\tau_D(\delta) = \inf\{t > \delta : X_t \in D\}$ of X to a bounded set $D \in \mathcal{B}(\mathbb{R}^n)$ after time $\delta > 0$.

Theorem 3.7. *Let Assumption A_{se} hold.*
(a) There exist $c_\pi, u_\pi \in (0, \infty)$ such that

$$c_\pi/\exp(u_\pi r^{1-p})) \leq \pi(\{|x| \geq r\}) \quad \text{for all } r \in [1, \infty).$$

(b) For every $x \in \mathbb{R}^n$, $\delta > 0$ and a bounded set $D \in \mathcal{B}(\mathbb{R}^n)$, there exist $c_\tau, u_\tau \in (0, \infty)$ such that

$$c_\tau/\exp(u_\tau t^{(1-p)/(1+p-\ell)}) \leq \mathbb{P}_x(\tau_D(\delta) \geq t) \quad \text{for all } t \in [1, \infty).$$

(c) For every $x \in \mathbb{R}^n$ there exist $c_{\mathrm{TV}}, u_{\mathrm{TV}} \in (0, \infty)$ such that

$$c_{\mathrm{TV}}/\exp(u_{\mathrm{TV}} t^{(1-p)/(1+p-\ell)}) \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \quad \text{for all } t \in [1, \infty).$$

In the proof of Theorem 3.7 we work with the **L**-drift condition $\mathbf{L}(V,\varphi,\Psi)$, where the Lyapunov function V grows polynomially, while Ψ grows at the stretched exponential rate with the "stretching" parameter $1-p$. The function φ is again polynomial and is obtained directly from the generator inequalities in Theorem 2.7. Like Ψ, the function h in Assumption (a) of Theorem 2.2 grows at the stretched exponential rate with parameter $1-p$, see Section 6.1.2 below for the details of the proof of Theorem 3.7.

Remark 3.8 (**matching rates**). The class of models defined by Assumption A_{se} contains elliptic and tempered Langevin diffusions studied in the context of upper bounds in [9, Sections 5.1 and 5.2]. The techniques of [9] can be applied to obtain upper bounds for all models covered by Assumption A_{se}. Our lower bounds in Theorem 3.7 match these upper bounds. More precisely, there exist constants $c_{\mathrm{TV}}, C_{\mathrm{TV}}, u_{\mathrm{TV}}, U_{\mathrm{TV}} \in (0,\infty)$ such that

$$(17) \qquad c_{\mathrm{TV}}/\exp(u_{\mathrm{TV}} t^{(1-p)/(1+p-\ell)}) \le \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \le C_{\mathrm{TV}}/\exp(U_{\mathrm{TV}} t^{(1-p)/(1+p-\ell)})$$

holds for all $t \in [1,\infty)$. Analogous bounds hold for the tails of the invariant measure π and the return time $\tau_D(\delta)$.

Remark 3.9 (**matching constants**). Assumption A_{se} cannot ensure that constants u_π, u_τ, u_{TV} in the exponents in Theorem 3.7 are only ε-away from optimal constants in upper bounds. This is because, under Assumption A_{se}, the coefficients of the SDE in (11) may oscillate asymptotically as $|x| \to \infty$. However, assuming that

$$\langle b(x), x/|x|\rangle/|x|^{-p} = -\alpha + o(|x|^{p-1}), \quad \langle \Sigma(x)x/|x|, x/|x|\rangle = \beta + o(|x|^{p-1}) \quad \& \quad \mathrm{Tr}(\Sigma(x)) = \gamma + o(1),$$

as $|x| \to \infty$, for some constants $p \in (0,1)$, $\alpha, \beta, \gamma \in (0,\infty)$, it is possible to prove

$$\lim_{r\to\infty} r^{(1-p)} \log \pi(\{|x| \ge r\}) = 2\alpha/(\beta(1-p)) = \alpha^{-(1-p)/(1+p)} \lim_{t\to\infty} t^{(1-p)/(1+p-\ell)} \log \mathbb{P}_x(\tau_D(\delta) \ge t).$$

Limit inferiors in these limits follow from the theory developed in this paper. However, they require a more involved Lyapunov function (i.e. a product of a polynomial and stretched exponential functions) than the one in the proof of Theorem 3.7. Upper bounds in these limits follow from [9, Thm 5.3]. Moreover, for the constants in the exponents for the total variation distance in (17), we can show $U_{\mathrm{TV}}/u_{\mathrm{TV}} < 2 + \varepsilon$. We omit the details for brevity.

3.1.3. *Exponential tails.* In this section we demonstrate that our methods may also be used to derive lower bounds on the rate of convergence of certain exponentially ergodic processes.

Assumption A_e. *The coefficients b and $\Sigma = \sigma\sigma^\mathsf{T}$ in (11) satisfy the following asymptotic assumptions:* $\limsup_{|x|\to\infty} \mathrm{Tr}(\Sigma(x)) < \infty$ *and*

$$-\alpha_L \le \langle b(x), x/|x|\rangle \le -\alpha_U \qquad \& \qquad \beta_L \le \langle \Sigma(x)x/|x|, x/|x|\rangle \le \beta_U$$

for all x outside of a compact set and some constants $\alpha_L, \alpha_U, \beta_L, \beta_U \in (0,\infty)$.

Theorem 3.10. *Let Assumption A_e hold.*

(a) There exist $c_\pi, u_\pi \in (0,\infty)$ such that

$$c_\pi/\exp(u_\pi r) \le \pi(\{|x| \ge r\}) \quad \text{for all } r \in [1,\infty).$$

(b) *For every $x \in \mathbb{R}^n$, $\delta \in (0, \infty)$ and a bounded set $D \in \mathcal{B}(\mathbb{R}^n)$ there exist $c_\tau, u_\tau \in (0, \infty)$ such that*

$$c_\tau / \exp(u_\tau t) \leq \mathbb{P}_x(\tau_D(\delta) \geq t) \quad \text{for all } t \in [1, \infty).$$

(c) *For every $x \in \mathbb{R}^n$ there exist $c_{\text{TV}}, u_{\text{TV}} \in (0, \infty)$ such that*

$$c_{\text{TV}} / \exp(u_{\text{TV}} t) \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\text{TV}} \quad \text{for all } t \in [1, \infty).$$

The **L**-drift condition **L**(V, φ, Ψ) in the proof of Theorem 3.10 uses V and Ψ, which exhibit polynomial and exponential growth, respectively. As in the examples of Sections 3.1.1 and 3.1.2, the function φ is polynomial and follows directly from the generator inequalities in Theorem 2.7. The details of the proof of Theorem 3.10 are in Section 6.1.3 below.

Remark 3.11 (**matching rates**). Exponential upper bounds on the results in Theorem 3.10 can be obtained, for example, by applying results from [9, Sec 3] and estimates in Proposition 6.3 below. In particular, there exist constants $C_{\text{TV}}, U_{\text{TV}} \in (0, \infty)$ such that

$$c_{\text{TV}} / \exp(u_{\text{TV}} t) \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\text{TV}} \leq C_{\text{TV}} / \exp(U_{\text{TV}} t) \quad \text{for all } t \in [1, \infty).$$

Similar bounds hold for the tails of the invariant measure π and the return time $\tau_D(\delta)$. Moreover, ensuring that the constants $u_\pi, u_\tau, u_{\text{TV}}$ in the exponents in Theorem 3.10 are close to their optimal constants in the upper bounds would require stronger assumptions and a more involved choice of Lyapunov functions. As the focus of the present paper is on subexponential ergodicity, the details are omitted.

3.2. **Lévy-driven SDE.** In this section we apply the results of Section 2 to a solution X of a Lévy-driven SDE taking values in $\mathcal{X} = \mathbb{R}$. More precisely, let X follow the SDE

$$(18) \qquad \qquad \mathrm{d}X_t = -\mu X_t \mathrm{d}t + \sigma(X_{t-}) \mathrm{d}L_t,$$

where $\mu \in (0, \infty)$, $\sigma : \mathbb{R} \to \mathbb{R}_+$ is Lipschitz with $0 < \inf_{x \in \mathbb{R}} \sigma(x) \leq \sup_{x \in \mathbb{R}} \sigma(x) < \infty$ and L is a pure-jump Lévy process with Lévy measure ν (cf. [34, Ch. 1]). Such X exists and is unique for any starting point $X_0 = x \in \mathbb{R}$ [29, Thm V.6]. The drift and the Gaussian component of L are zero as they do not influence the asymptotic behaviour of X when ν has heavy tails and σ is bounded. Adding them would require only a minor modification of the proof of Theorem 3.12 below.

Assumption A$_{\text{L}}$. *Assume that for some $m_c \in (1, \infty)$, the Lévy measure ν satisfies*

$$0 < \liminf_{r \to \infty} (\log r)^{m_c} \nu([r, \infty)) \leq \limsup_{r \to \infty} (\log r)^{m_c} \nu([r, \infty)) < \infty \quad \text{and}$$
$$0 = \limsup_{r \to -\infty} \nu((-\infty, r]) (\log |r|)^{m_c + 1}.$$

As we shall see, when the jumps of L have very heavy tails (as in Assumption A$_{\text{L}}$), X does not exhibit exponential ergodicity. Recall that $\tau_D(\delta) = \inf\{t > \delta : X_t \in D\}$ is the return time of the process X to a bounded set $D \in \mathcal{B}(\mathbb{R}^n)$ after time $\delta > 0$.

Theorem 3.12. *Let Assumption A_L hold and the process X satisfy SDE (18) above.*

(a) *For any $\varepsilon > 0$ there exists a constant $c_\pi \in (0, \infty)$ such that*

$$c_\pi / (\log r)^{m_c - 1 + \varepsilon} \leq \pi([r, \infty)) \quad \text{for all } r \in [1, \infty).$$

(b) For every $x \in \mathbb{R}$, $\varepsilon, \delta > 0$ and bounded set $D \in \mathcal{B}(\mathbb{R})$, there exists $c_\tau \in (0, \infty)$ such that

$$c_\tau / t^{m_c + \varepsilon} \leq \mathbb{P}_x(\tau_D(\delta) \geq t) \quad \text{for all } t \in [1, \infty).$$

(c) For any $x \in \mathbb{R}$ and $\varepsilon > 0$ there exists a constant $c_{\mathrm{TV}} \in (0, \infty)$ such that

$$c_{\mathrm{TV}} / t^{m_c - 1 + \varepsilon} \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi\|_{\mathrm{TV}} \quad \text{for all } t \in [1, \infty).$$

The **L**-drift condition $\mathbf{L}(V, \varphi, \Psi)$ in the proof of Theorem 3.12 uses the Lyapunov function V with logarithmic growth. This is necessary to ensure the integrability of the marginals of the process $V(X)$. The function Ψ has polynomial growth as does φ (the latter is again obtained from the generator inequalities in Theorem 2.7). The function h in Assumption (a) of Theorem 2.2 is taken to be polynomial with the "largest" growth rate, such that the marginals of the process $h \circ V(X)$ remain integrable. The proof of Theorem 3.12 is in Section 6.2 below.

Remark 3.13 (**matching rates**). Theorem 3.12 is applicable to general pure-jump Lévy drivers with two-sided jumps and arbitrary path variation. In particular, it covers the simpler model studied in [12] with additive noise, where it is assumed that $\sigma \equiv 1$ and $\nu((-\infty, 1]) = 0$ (making L a compound Poisson process with positive jumps and, consequently, $\mathcal{X} = \mathbb{R}_+$ for $X_0 = x \in \mathbb{R}_+$). Our lower bounds match the upper bounds from [12], i.e. for some $c_{\mathrm{TV}}, C_{\mathrm{TV}} \in (0, \infty)$ we have

$$c_{\mathrm{TV}} / t^{m_c - 1 + \varepsilon} \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi\|_{\mathrm{TV}} \leq C_{\mathrm{TV}} / t^{m_c - 1 - \varepsilon} \quad \text{for all } t \in [1, \infty).$$

Analogous bound holds for the tail of the invariant measure π and the return time $\tau_D(\delta)$.

Remark 3.14. Theorem 3.12 and Remark 3.13 show that the process X in (18) converges to stationarity at a polynomial rate even though its invariant measure has a logarithmically heavy positive tail. Since Assumption $\mathrm{A_L}$ stipulates that the negative tail of the Lévy measure ν is orders of magnitude thinner than its positive tail, we prove Theorem 3.12 in Section 6.2 below using a Lyapunov function (45), which is bounded on the negative half-line.

The asymmetry in Assumption $\mathrm{A_L}$ is also visible in the invariant measure π. If, for example, $\nu((-\infty, r])$ decays polynomially, then the form of the extended generator of X in (44) below, applied to an appropriate polynomial Lyapunov function V, tends to $+\infty$ (as $r \to -\infty$) at the same rate as V itself. This follows by a similar argument to the one in the proof of Proposition 6.4(a) below. In this case, by [9, Prop. 3.1] and the Markov inequality, we obtain a polynomial upper bound on the tail $\pi((-\infty, r])$ as $r \to -\infty$. Moreover, by the same argument, if $\nu((-\infty, r])$ decays exponentially as $r \to -\infty$, so does $\pi((-\infty, r])$.

3.3. Stochastic damping Hamiltonian system.

Consider the hypoelliptic diffusion $X = (Z, Y)$, where Z_t (resp. Y_t) is the position (resp. velocity) at time t of a physical system moving in $\mathbb{R}^n$, satisfying the following SDE: $\mathrm{d}Z_t = Y_t \mathrm{d}t$ and $\mathrm{d}Y_t = \sigma(Z_t, Y_t)\mathrm{d}B_t - (c(Z_t, Y_t)Y_t + \nabla U(Z_t))\mathrm{d}t$. Here the gradient $\nabla U(Z_t)$ is a friction force, $c(Z_t, Y_t)Y_t$ a damping force and $\sigma(Z_t, Y_t)\mathrm{d}B_t$ a random force, with B being the standard n-dimensional Brownian motion, acting on the system at (Z_t, Y_t). Exponential convergence to the invariant measure of this class of models has been studied extensively, see e.g. [21, 39].

Our primary focus is on deriving lower bounds in cases where subexponential upper bounds on the rate of convergence have been established. We thus consider a one-dimensional example previously studied in the context of subexponential upper bounds [9]:

$$\text{(19)} \qquad \mathrm{d}Z_t = Y_t\mathrm{d}t, \qquad \mathrm{d}Y_t = \sigma\mathrm{d}B_t - (cY_t + U'(Z_t))\mathrm{d}t,$$

where $U : \mathbb{R} \to \mathbb{R}$ is in $C^2(\mathbb{R})$ and $\sigma, c \in (0, \infty)$ are positive constants. We are interested in the case when the invariant measure of the process $X = (Z, Y)$ has polynomial tails.

Assumption A_H. *Let the function $U \in C^2(\mathbb{R})$ be such that $zU'(z) = a + o(1)$ as $|z| \to \infty$ for some $a \in (0, \infty)$. Assume also that the constants $\sigma, c \in (0, \infty)$ in (19) satisfy $ac/\sigma^2 > 1/2$.*

The next result provides matching polynomial lower and upper bounds on the rate of convergence to stationarity in f-variation of the hypoelliptic diffusion in (19) (proof is in Section 6.3 below). The parameter $m = 0$ in Theorem 3.15 corresponds to total variation.

Theorem 3.15. *Let Assumption A_H hold and pick $m \in [0, 2ac/\sigma^2 - 1)$. Then for the function $f_m(z, y) := 1 + |z|^m$, any $x = (z, y) \in \mathbb{R}^n$ and $\varepsilon > 0$, there exist $c_{m,\varepsilon}, C_{m,\varepsilon} \in (0, \infty)$ such that*

$$c_{m,\varepsilon}/t^{ac/\sigma^2 - 1/2 - m/2 + \varepsilon} \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{f_m} \leq C_{m,\varepsilon}/t^{ac/\sigma^2 - 1/2 - m/2 - \varepsilon} \quad \text{for all } t \in [1, \infty).$$

The upper bound on the rate of convergence in Theorem 3.15 is obtained by applying the drift condition in [9] (see (10) above) to an appropriate Lyapunov function (inspired by [39]). Note that the model in Theorem 3.15 has not been analysed in [9], where a hypoelliptic diffusion in (19) with stretched exponential tails is considered. Following [15], the matching lower bound in Theorem 3.15 is obtained by comparing the tails of X_t and π via Lemma 5.1. This lower bound does not require a verification of the **L**-drift condition $\mathbf{L}(V,\varphi,\Psi)$, because the invariant measure π of the process in (19) has a known density proportional to $(z, y) \mapsto \exp(-2c/\sigma^2(y^2/2 + U(z)))$.

The Lyapunov function V_u used to obtain the upper bound in Theorem 3.15 also yields polynomial upper bounds on the tail of return times. By analogy with all other models discussed in Section 3, it would be natural to use the same V_u to establish the **L**-drift condition $\mathbf{L}(V,\varphi,\Psi)$. However, a function φ, which makes the process $1/V_u(X)$ into a supermartingale, cannot satisfy the growth conditions in $\mathbf{L}(V,\varphi,\Psi)\mathbf{(i)}$. Thus, unlike in the other models of this section, a different Lyapunov function V_l for lower bounds is needed. Such a V_l exists but it only yields exponential lower bounds on the tails of return times. The reason for this discrepancy is that V_u necessarily mainly depends on the heavy-tailed component Z, while V_l has to mostly depend on the light-tailed component Y.

4. Return times to bounded sets for semimartingales

This section develops a general theory for the analysis of return times of continuous-time semimartingales to bounded sets. The main result in this section, Lemma 4.2, is a far reaching generalisation of the approach dealing with return times initiated in [4, Sec 3] to stochastic processes with jumps and/or unbounded variance. Throughout this section we fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a right-continuous filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. We begin with an elementary maximal inequality.

Proposition 4.1 (Maximal inequality). *Let $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ be a right-continuous filtration and $\xi = (\xi_t)_{t\in\mathbb{R}_+}$ an $(\mathcal{F}_t)$-adapted process with càdlàg paths taking values in $[0, 1]$. Define an $(\mathcal{F}_t)$-stopping time $\tau_r := \inf\{t \in \mathbb{R}_+ : \xi_t > r\}$ (recall $\inf \emptyset = \infty$) and assume that, for some $r > 0$ and a locally*

bounded measurable function $f : \mathbb{R}_+ \times [0,1] \to \mathbb{R}_+$, the process $(\xi_{t \wedge \tau_r} - \int_0^{t \wedge \tau_r} f(u, \xi_u) \mathrm{d}u)_{t \in \mathbb{R}_+}$ is an $(\mathcal{F}_t)$-supermartingale. Then, for any $s \in (0, \infty)$, we have

$$\mathbb{P}\left(\sup_{0 \le u < s} \xi_u > r \Big| \mathcal{F}_0 \right) \le r^{-1} \left(\xi_0 + \mathbb{E}\left[\int_0^{s \wedge \tau_r} f(u, \xi_u) \mathrm{d}u \Big| \mathcal{F}_0 \right] \right) \quad a.s.$$

Proof. Pick any $s \in (0, \infty)$ and consider an $(\mathcal{F}_t)$-stopping time $\tau_r \wedge s$, bounded above by s. Note that $\sup_{0 \le u < \tau_r \wedge s} \xi_u \le r$. Since f is bounded on the compact set $[0, s] \times [0, r]$, there exists a constant $C \in (0, \infty)$ such that $\sup_{0 \le u < \tau_r \wedge s} f(u, \xi_u) \le C$ a.s. Thus we obtain

$$(20) \qquad\qquad 0 \le \mathbb{E}\left[\int_0^{\tau_r \wedge s} f(u, \xi_u) \mathrm{d}u \Big| \mathcal{F}_0 \right] \le Cs < \infty \quad a.s.$$

Since $(\xi_{t \wedge \tau_r} - \int_0^{t \wedge \tau_r} f(u, \xi_u) \mathrm{d}u)_{t \in \mathbb{R}_+}$ is an $(\mathcal{F}_t)$-supermartingale, $\xi_{\tau_r \wedge s} - \int_0^{\tau_r \wedge s} f(u, \xi_u) \mathrm{d}u$ is integrable and $\mathbb{E}[\xi_{\tau_r \wedge s} - \int_0^{\tau_r \wedge s} f(u, \xi_u) \mathrm{d}u | \mathcal{F}_0] \le \xi_0$. The inequality in (20) and the fact that ξ is non-negative imply the following:

$$\mathbb{E}[\xi_{\tau_r \wedge s} | \mathcal{F}_0] = \mathbb{E}\left[\xi_{\tau_r \wedge s} - \int_0^{\tau_r \wedge s} f(u, \xi_u) \mathrm{d}u \Big| \mathcal{F}_0 \right] + \mathbb{E}\left[\int_0^{\tau_r \wedge s} f(u, \xi_u) \mathrm{d}u \Big| \mathcal{F}_0 \right]$$

$$(21) \qquad\qquad \le \xi_0 + \mathbb{E}\left[\int_0^{\tau_r \wedge s} f(u, \xi_u) \mathrm{d}u \Big| \mathcal{F}_0 \right].$$

Moreover, by the definition of τ_r in the proposition we have $\{\sup_{u \in [0,s)} \xi_u > r\} = \{\tau_r < s\}$ a.s. Since ξ is càdlàg, on the event $\{\tau_r < s\}$ we have $\xi_{\tau_r \wedge s} = \xi_{\tau_r} \ge r$ a.s. Thus, by (21), we have

$$\mathbb{P}\left(\sup_{0 \le u < s} \xi_u > r \Big| \mathcal{F}_0 \right) = \mathbb{P}\left(\tau_r < s \Big| \mathcal{F}_0 \right) \le r^{-1} \mathbb{E}[\xi_{\tau_r \wedge s} \mathbb{1}\{\tau_r < s\} | \mathcal{F}_0] \le r^{-1} \mathbb{E}[\xi_{\tau_r \wedge s} | \mathcal{F}_0]$$

$$\le r^{-1} \left(\xi_0 + \mathbb{E}\left[\int_0^{\tau_r \wedge s} f(u, \xi_u) \mathrm{d}u \Big| \mathcal{F}_0 \right] \right),$$

implying the proposition. $\square$

Proposition 4.1 will be applied in the proof of Lemma 4.2 with a continuous function f. To state the lemma, consider an $(\mathcal{F}_t)$-adapted process $\kappa := (\kappa_t)_{t \in \mathbb{R}_+}$ with càdlàg paths, taking values in $[1, \infty)$. Let $\mathcal{T}$ denote the set of all $[0, \infty]$-valued stopping times with respect to $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$. For any $\ell, r \in \mathbb{R}_+$ and stopping time $T \in \mathcal{T}$, define the first entry times (after T) by

$$(22) \qquad\qquad \lambda_{\ell, T} := T + \inf\{s \in \mathbb{R}_+ : T < \infty, \ \kappa_{T+s} < \ell\},$$

$$(23) \qquad\qquad \rho_{r, T} := T + \inf\{s \in \mathbb{R}_+ : T < \infty, \ \kappa_{T+s} > r\},$$

where $\inf \emptyset = \infty$. If $T = 0$, we write $\lambda_\ell := \lambda_{\ell, 0}$ and $\rho_r := \rho_{r, 0}$.

Lemma 4.2. *Let $\kappa = (\kappa_t)_{t \in \mathbb{R}_+}$ be a $[1, \infty)$-valued $(\mathcal{F}_t)$-adapted process with càdlàg paths, satisfying $\limsup_{t \to \infty} \kappa_t = \infty$ a.s. Suppose that there exist a level $\ell \in (1, \infty)$ and a non-decreasing continuous function $\varphi : (0, 1] \to \mathbb{R}_+$, such that the process*

$$\left(1/\kappa_{(\rho_{r_q}+t) \wedge \lambda_{r, \rho_{r_q}}} - \int_{\rho_{r_q}}^{(\rho_{r_q}+t) \wedge \lambda_{r, \rho_{r_q}}} \varphi(1/\kappa_u) \mathrm{d}u \right)_{t \in \mathbb{R}_+} \qquad \textit{is an } (\mathcal{F}_{\rho_{r_q}+t})\textit{-supermartingale}$$

for every $q \in (0,1)$ and $r \in (\ell, \infty)$, where $r_q := 2r/(1-q)$. Pick any $q \in (0,1)$, $\varepsilon \in (0, (1-q)/2]$ and a non-decreasing function $f : \mathbb{R}_+ \to \mathbb{R}_+$. Then for every $r \in (\ell, \infty)$ we have

$$(24) \qquad \mathbb{P}\left(\int_0^{\lambda_\ell} f(\kappa_s)\mathrm{d}s \geq f(r)\varepsilon/(r\varphi(1/r)) \Big| \mathcal{F}_0 \right) \geq q\mathbb{P}(\rho_{r_q} < \lambda_\ell | \mathcal{F}_0) \quad a.s.$$

Remark 4.3. (I) The assumption $\limsup_{t\to\infty} \kappa_t = \infty$ a.s. in Lemma 4.2 implies $\mathbb{P}(\rho_r < \infty) = 1$ for all $r \in [1, \infty)$, making $\kappa_{\rho_{r_q}}$ well defined. The main step in the proof of inequality (24) in Lemma 4.2 consists of establishing the following: with probability at least q, after reaching the level r_q, the process κ spends more than $\varepsilon/(r\varphi(1/r))$ units of time before returning below the level r.

(II) Note that $f \equiv 1$ in Lemma 4.2 yields a lower bound on the tail probability $\mathbb{P}(\lambda_\ell \geq t | \mathcal{F}_0)$. In applications of Lemma 4.2 it is crucial that q can be taken arbitrarily close to 1. This allows us to conclude that for any fixed time t and starting point κ_0 larger than the level $2G_f(t)/(1-q)$, the process does not leave (ℓ, ∞) before time t with probability q. In particular, this will imply that petite sets of a Markov process satisfying $\mathbf{L}(V, \varphi, \Psi)$ are necessarily bounded (Lemma 5.4 below).

(III) The null set where the inequality in (24) fails to hold may vary with r. However, when applying Lemma 4.2 in this paper, we only require the case where $\mathcal{F}_0$ is a trivial σ-algebra, making (24) hold for all $r \in (\ell, \infty)$ simultaneously.

Proof of Lemma 4.2. Pick $q \in (0,1)$ and $r \in (\ell, \infty)$. Note that the inequality in (24) holds for all $\varepsilon \in (0, (1-q)/2]$ if it holds for $\varepsilon = (1-q)/2$ (the right-hand side of (24) does not depend on ε, while the probability on the left-hand side is decreasing in ε). We may thus fix $\varepsilon = (1-q)/2$.

We start by showing that, once the process κ reaches the level $r_q = 2r/(1-q)$, with probability at least q it takes $\varepsilon/(r\varphi(1/r))$ units of time for κ to return to the interval $[1, r)$. More precisely, we now establish the following inequality:

$$(25) \qquad \mathbb{P}(\lambda_{r, \rho_{r_q}} \geq \rho_{r_q} + \varepsilon/(r\varphi(1/r)) | \mathcal{F}_{\rho_{r_q}}) \geq q \quad \text{a.s.}$$

By the non-confinement assumption $\limsup_{t\to\infty} \kappa_t = \infty$, we have $\rho_{r_q} < \infty$ a.s. Define the càdlàg process $(\xi_t)_{t\in\mathbb{R}_+}$ by $\xi_t := 1/\kappa_{\rho_{r_q}+t}$. Note that $\tau_{1/r} = \inf\{t > 0 : \xi_t > 1/r\} = \lambda_{r,\rho_{r_q}} - \rho_{r_q}$ by (22) and hence $t \wedge \tau_{1/r} = ((\rho_{r_q}+t) \wedge \lambda_{r,\rho_{r_q}}) - \rho_{r_q}$. Moreover, by the definition of $\tau_{1/r}$, we have $\{\sup_{u\in[0,s]} \xi_u > 1/r\} = \{\tau_{1/r} < s\}$ a.s. for any $s \in \mathbb{R}_+$. By assumption, the process $(\xi_{t\wedge\tau_{1/r}} - \int_0^{t\wedge\tau_{1/r}} \varphi(\xi_u)\mathrm{d}u)_{t\in\mathbb{R}_+}$ is an $(\mathcal{F}_{\rho_{r_q}+t})$-supermartingale. Moreover, since $\varphi : (0,1] \to \mathbb{R}_+$ is non-decreasing and continuous, it has a unique extension (via its right-limit at 0) to a continuous function $\varphi : [0,1] \to \mathbb{R}_+$. Applying Proposition 4.1 (with a continuous function $f(u,s) = \varphi(s)$) to ξ and the stopping time $\tau_{1/r}$ yields

$$\mathbb{P}(\lambda_{r,\rho_{r_q}} < \rho_{r_q} + t | \mathcal{F}_{\rho_{r_q}}) = \mathbb{P}(\tau_{1/r} < t | \mathcal{F}_{\rho_{r_q}}) = \mathbb{P}(\sup_{0 \leq u < t} \xi_u > 1/r | \mathcal{F}_{\rho_{r_q}})$$

$$\leq r\left(\xi_0 + \mathbb{E}\left[\int_0^{t\wedge\tau_{1/r}} \varphi(\xi_u)\mathrm{d}u \Big| \mathcal{F}_{\rho_{r_q}} \right] \right)$$

$$= r\left(1/\kappa_{\rho_{r_q}} + \mathbb{E}\left[\int_{\rho_{r_q}}^{(\rho_{r_q}+t)\wedge\lambda_{r,\rho_{r_q}}} \varphi(1/\kappa_u)\mathrm{d}u \Big| \mathcal{F}_{\rho_{r_q}} \right] \right)$$

$$\leq r\left(1/r_q + \varphi(1/r)\,\mathbb{E}\left[(\rho_{r_q}+t) \wedge \lambda_{r,\rho_{r_q}} - \rho_{r_q} | \mathcal{F}_{\rho_{r_q}} \right] \right)$$

$$\leq r(1/r_q + \varphi(1/r)t) = (1-q)/2 + r\varphi(1/r)t, \quad t \in (0,\infty),$$

where the second inequality holds by the following facts: φ is a non-decreasing function and the inequality $1/\kappa_u \leq 1/r$ is valid on the event $\{\rho_{r_q} < u < (\rho_{r_q} + t) \wedge \lambda_{r,\rho_{r_q}}\}$. The third inequality is a consequence of the fact $((\rho_{r_q} + t) \wedge \lambda_{r,\rho_{r_q}}) - \rho_{r_q} = \tau_{1/r} \wedge t \leq t$, while the last equality follows from the definition of r_q. By taking complements, we get

$$\mathbb{P}(\lambda_{r,\rho_{r_q}} \geq \rho_{r_q} + t | \mathcal{F}_{\rho_{r_q}}) \geq 1 - ((1-q)/2 + r\varphi(1/r)t).$$

Setting $t = \varepsilon/(r\varphi(1/r))$ and recalling $\varepsilon = (1-q)/2$, we obtain (25).

Note that on the event $\{\lambda_{r,\rho_{r_q}} \geq \rho_{r_q} + \varepsilon/(r\varphi(1/r))\}$, for any non-decreasing function f, we have $f(\kappa_{\rho_{r_q}+t}) \geq f(r)$ for all $t \in [0, \varepsilon/(r\varphi(1/r))]$. Since $r > \ell$, on the event $\{\rho_{r_q} < \lambda_\ell\}$, the inequality $\lambda_\ell \geq \lambda_{r,\rho_{r_q}}$ holds, implying the following inclusion:

$$\left\{ \int_0^{\lambda_\ell} f(\kappa_t)\mathrm{d}t \geq f(r)\varepsilon/(r\varphi(1/r)) \right\} \supset \{\rho_{r_q} < \lambda_\ell\} \cap \{\lambda_{r,\rho_{r_q}} \geq \rho_{r_q} + \varepsilon/(r\varphi(1/r))\}.$$

By the inequality in (25), we thus obtain the inequality in (24):

$$\mathbb{P}\left(\int_0^{\lambda_\ell} f(\kappa_t)\mathrm{d}t \geq f(r)\varepsilon/(r\varphi(1/r)) \Big| \mathcal{F}_0 \right)$$
$$\geq \mathbb{E}\left[\mathbb{1}\{\rho_{r_q} < \lambda_\ell\}\mathbb{P}\left(\lambda_{r,\rho_{r_q}} > \rho_{r_q} + \varepsilon/(r\varphi(1/r)) | \mathcal{F}_{\rho_r} \right) \Big| \mathcal{F}_0 \right] \geq q\mathbb{P}(\rho_{r_q} < \lambda_\ell | \mathcal{F}_0) \quad \text{a.s.} \quad \square$$

5. LOWER BOUNDS ON THE ERGODICITY OF MARKOV PROCESSES

5.1. A lower bound on the f-variation rate of a Markov process.
In this subsection we consider a strong Markov process $X = (X_t)_{t \in \mathbb{R}_+}$ on a general metric space $\mathcal{X}$ with invariant measure π on $\mathcal{B}(\mathcal{X})$ (see Section 2.1 for definitions). The following lemma generalizes to f-variation the lower bound in [15, Thm 3.6] on the *total variation* between π and the law of X_t. The key assumption in Lemma 5.1 is the lower bound on the decay of the tail of the integral of f with respect to π. We stress that Lemma 5.1 does not require the **L**-drift condition $\mathbf{L}(V, \varphi, \Psi)$.

Lemma 5.1. *Let X be a Markov process with an invariant measure π on the state space $\mathcal{X}$. Let functions $H, f, G : \mathcal{X} \to [1, \infty)$ be such that $f(x)G(x) = H(x)$ for all $x \in \mathcal{X}$ and (a) & (b) hold.*

(a) There exists a function $a : [1, \infty) \to (0, 1]$ such that the function $A(r) := ra(r)$ is increasing, $\lim_{r\uparrow\infty} A(r) = \infty$ and $\int_{\{G \geq r\}} f(x)\pi(\mathrm{d}x) \geq a(r)$ for all $r \in [1, \infty)$.

(b) There exists a function $v : \mathcal{X} \times \mathbb{R}_+ \to [1, \infty)$, increasing in the second argument and satisfying $\mathbb{E}_x[H(X_t)] \leq v(x, t)$ for all $x \in \mathcal{X}$ and $t \in [1, \infty)$.

Then the following bound holds for every $t \in [1, \infty)$ and $x \in \mathcal{X}$:

$$\|\pi(\cdot) - \mathbb{P}_x(X_t \in \cdot)\|_f \geq \left(a \circ A^{-1} \circ (2v) \right)(x, t)/2.$$

Proof. It follows from the definition of f-variation distance and Markov inequality that, for every $t \in \mathbb{R}_+$ and every $r \geq 1$, one has the lower bound

$$\|\pi(\cdot) - \mathbb{P}_x(X_t \in \cdot)\|_f \geq \int_{\{G \geq r\}} f(x)\pi(\mathrm{d}x) - \mathbb{E}_x[f(X_t)\mathbb{1}\{G(X_t) \geq r\}]$$

$$\geq a(r) - \frac{1}{r}\mathbb{E}_x[f(X_t)G(X_t)\mathbb{1}\{G(X_t) \geq r\}]$$

$$\geq a(r) - \frac{1}{r}\mathbb{E}_x[H(X_t)] \geq a(r) - \frac{v(x,t)}{r}.$$

Let $r = r(t)$ be the unique solution to the equation $ra(r) = 2v(x,t)$. Put differently we have $r(t) = A^{-1}(2v(x,t))$ for all $t \in [1,\infty)$ and $v(x,t)/r(t) = a(r(t))/2$. Thus we obtain

$$a(r(t)) - v(x,t)/r(t) = a(r(t))/2 = a(A^{-1}(2v(x,t)))/2,$$

which, combined with the previous display, concludes the proof. $\square$

Remark 5.2. In applications of Lemma 5.1 in practice, a good choice of $H = f \cdot G$ (recall that f is given by the variation norm) requires balancing (I) and (II) below.

(I) It is beneficial to choose H so that the lower bound $a(r)$ on the tail $\pi(\{G \geq r\})$ is such that $A(r) = ra(r)$ tends to infinity polynomially. This is because a slower (logarithmic) growth in A would imply a faster (stretched exponential) growth in v of $A^{-1} \circ (2v)$, making the lower bound $a \circ A^{-1} \circ (2v)$ smaller (recall that $a(r) \to 0$ as $r \to \infty$). In particular, this requires H to grow sufficiently fast.

(II) The growth of $v(x,t)$ is often obtained via the application of Lemma 2.10. In particular this lemma relies on bounding $\mathcal{A}H$ (differently put, the derivatives of H) by a concave function of H, introducing a restriction on the growth of H.

To see how the choice of H plays out in specific models, see applications of Theorem 2.2 and Corollary 2.3 in Section 6 below, where $H = h \circ V$, V is the Lyapunov function and h an arbitrary function chosen with (I) and (II) above in mind.

5.2. **Return time estimates and petite sets.** The main estimate required in the proofs of our main theorems, stated in Section 2 above, is given in Proposition 5.3. It essentially bounds from below the tail of the return time $S_{(\ell)}$ of the process X into the set $\{V < \ell\}$.

Proposition 5.3. *Let Assumption* **L(V,φ,Ψ)** *hold. Then there exists $\ell_0 \in [1,\infty)$ such that the following holds: for any $\ell \in (\ell_0,\infty)$ there exists $C_\ell \in (0,\infty)$, such that for any $x \in \{\ell + 1 \leq V\}$, a non-decreasing continuous function $h : \mathbb{R}_+ \to \mathbb{R}_+$, $q \in (0,1)$ and $\varepsilon \in (0,(1-q)/2]$, inequality (26) holds for all times $t > h(\ell)\varepsilon/(\ell\varphi(1/\ell))$,*

$$(26) \quad \mathbb{P}_x\left(\int_0^{S_{(\ell)}} h \circ V(X_s)\mathrm{d}s \geq t\right)$$

$$\geq q\mathbb{1}\{V(x) < 2G_h(t)/(1-q)\}\frac{C_\ell}{\Psi(2G_h(t)/(1-q))} + q\mathbb{1}\{V(x) \geq 2G_h(t)/(1-q)\}.$$

In (26), *$G_h : (h(\ell)\varepsilon/(\ell\varphi(1/\ell)),\infty) \to (\ell,\infty)$ is the inverse of $r \mapsto h(r)\varepsilon/(r\varphi(1/r))$ on (ℓ,∞).*

The proof of Proposition 5.3 is based on Lemma 4.2 and requires us to show that Assumption **L(V,φ,Ψ)** implies the assumptions of Lemma 4.2 for the process $\kappa = V(X)$.

Proof. Consider the process $\kappa = V(X)$. Recall the definition of the return time $\lambda_{\ell,T}$ and the first-passage time $\rho_{r,T}$ (where $\ell, r \in (0, \infty)$ and T an $(\mathcal{F}_t)$-stopping time) for the process κ in (22) and (23) respectively. As in Section 4, we denote $\lambda_\ell = \lambda_{\ell,0}$ and $\rho_r = \rho_{r,0}$. Note that, by Assumption $\mathbf{L}(V,\varphi,\Psi)$, we have $S_{(\ell)} = \lambda_\ell$ and $T^{(r)} = \rho_r$ and the process κ satisfies the non-confinement property for every starting point $x \in \mathcal{X}$, i.e. $\mathbb{P}_x(\limsup_{t\to\infty} \kappa_t = \infty) = 1$. Moreover, by $\mathbf{L}(V,\varphi,\Psi)\mathbf{(i)}$, there exists $\ell_0 \in [1, \infty)$ such that, for every $r \in (\ell_0, \infty)$, the process

$$\left(1/\kappa_{(\rho_{r_q}+t)\wedge\lambda_{r,\rho_{r_q}}} - \int_{\rho_{r_q}}^{(\rho_{r_q}+t)\wedge\lambda_{r,\rho_{r_q}}} \varphi(1/\kappa_u)\mathrm{d}u \right)_{t\in\mathbb{R}_+}$$

is an $(\mathcal{F}_{\rho_{r_q}+t})$-supermartingale under $\mathbb{P}_x$ (since $\int_{\rho_{r_q}}^{\lambda_{r,\rho_{r_q}}} \mathbb{1}\{\kappa_u \leq \ell_0\}\mathrm{d}u = 0$ $\mathbb{P}_x$-a.s.) for every $x \in \mathcal{X}$, where $r_q := 2r/(1-q)$. Thus, by the inequality in (24) of Lemma 4.2, for any $\ell \in [\ell_0, \infty)$, $r \in (\ell, \infty)$, non-decreasing function $h : \mathbb{R}_+ \to \mathbb{R}_+$, $q \in (0,1)$ and $\varepsilon \in (0, (1-q)/2]$ we have

$$(27) \qquad \mathbb{P}_x\left(\int_0^{\lambda_\ell} h(\kappa_s)\mathrm{d}s \geq \varepsilon h(r)/(r\varphi(1/r)) \right) \geq q\mathbb{P}_x(\rho_{r_q} < \lambda_\ell).$$

Recall that by $\mathbf{L}(V,\varphi,\Psi)\mathbf{(i)}$ the function $r \mapsto h(r)\varepsilon/(r\varphi(1/r))$ is continuous and increasing on $[1, \infty)$ (and thus invertible on (ℓ_0, ∞)), with inverse G_h is defined on $t \in (h(\ell_0)\varepsilon/(\ell_0\varphi(1/\ell_0)), \infty)$. For any $t > h(\ell)\varepsilon/(\ell\varphi(1/\ell))$, set $r = G_h(t) > \ell$ and note $r_q = 2G_h(t)/(1-q) > \ell+1$. The inequality in (26) follows from (27) and inequality (2) in $\mathbf{L}(V,\varphi,\Psi)\mathbf{(ii)}$ for all $x \in \{\ell+1 \leq V\}$, since on the subset $x \in \{r_q \leq V\}$ we have $\mathbb{P}_x(\rho_{r_q} < \lambda_\ell) = \mathbb{P}_x(T^{(r_q)} < S_{(\ell)}) = 1$ by definition. $\qquad\square$

A non-empty measurable set $B \in \mathcal{B}(\mathcal{X})$ is *petite* (for the Markov process X) if there exist a probability measure a on $\mathcal{B}(\mathbb{R}_+)$ and a finite measure ν_a on $\mathcal{B}(\mathcal{X})$ with $\nu_a(\mathcal{X}) > 0$, satisfying

$$(28) \qquad \int_0^\infty \mathbb{P}_x(X_t \in \cdot)a(\mathrm{d}t) \geq \nu_a(\cdot) \quad \text{for all } x \in B.$$

The following lemma shows that, under Assumptions $\mathbf{L}(V,\varphi,\Psi)$, every petite set for X belongs to a sublevel set of the Lyapunov function V.

Lemma 5.4 (Under $\mathbf{L}$-drift condition, petite sets are bounded). *Let $\mathbf{L}(V,\varphi,\Psi)$ hold. Assume that a set $B \in \mathcal{B}(\mathcal{X})$ is petite for the process X. Then there exists $r_0 \in (1, \infty)$ such that $B \subset \{V \leq r_0\}$.*

The proof of this lemma is based on a simple idea, which we first explain informally. Since ν_a in (28) is a non-zero measure, we have $\nu_a(D) > 0$ for some compact set D. Denote by $\tau_D(0) := \inf\{t > 0 : X_t \in D\}$ the first time X is in D and let τ_a be an independent random time with law a (in (28)). Pick $t_0 \in (0, \infty)$ such that $\mathbb{P}(\tau_a > t_0) \leq \nu_a(D)/2$ and note $\{X_{\tau_a} \in D, \tau_a \leq t_0\} \subset \{\tau_D(0) \leq t_0\}$. Since, by (28), it holds

$$\nu_a(D) \leq \mathbb{P}_x(X_{\tau_a} \in D) \leq \nu_a(D)/2 + \mathbb{P}_x(X_{\tau_a} \in D, \tau_a \leq t_0) \leq \nu_a(D)/2 + \mathbb{P}_x(\tau_D(0) \leq t_0),$$

we get $0 < \nu_a(D)/2 \leq \mathbb{P}_x(\tau_D(0) \leq t_0)$ for all starting points x in the petite set B. However, under the $\mathbf{L}$-drift condition $\mathbf{L}(V,\varphi,\Psi)$ (by Proposition 5.3) we have $\mathbb{P}_x(\tau_D(0) \leq t_0) \leq 1 - q$ for any $q \in (0,1)$ and all x with $V(x)$ sufficiently large. Hence B must be contained in a sublevel set of V.

Proof of Lemma 5.4. Let B be an arbitrary petite set with a probability measure a on $\mathcal{B}(\mathbb{R}_+)$ and a non-zero measure ν_a on $\mathcal{B}(\mathcal{X})$ such that (28) holds. Since $\cup_{\ell=1}^\infty \{V \leq \ell\} = \mathcal{X}$ and, by (28), $1 \geq \nu_a(\mathcal{X}) > 0$, there exists $\ell_1 \in (1, \infty)$ such that $c := \nu_a(\{V \leq \ell_1\}) \in (0, 1]$. By Proposition 5.3

(with $h \equiv 1$), there exist $\ell_0 \in [\ell_1, \infty)$ such that for every $q \in (0,1)$, $\varepsilon = (1-q)/2$ and $x \in \mathcal{X}$ we have

$$(29) \qquad \mathbb{P}_x(S_{(\ell_0)} \geq t) \geq q \quad \text{for all } t \in (\varepsilon/(\ell_0\varphi(1/\ell_0)), \infty) \text{ and } x \in \{V \geq 2G_1(t)/(1-q)\},$$

where $G_1 : (\varepsilon/(\ell_0\varphi(1/\ell_0)), \infty) \to (\ell_0, \infty)$ is the inverse of the function $r \mapsto \varepsilon/(r\varphi(1/r))$. Since a is a probability measure on $\mathcal{B}(\mathbb{R}_+)$, there exists $t_1 \in (\varepsilon/(\ell_0\varphi(1/\ell_0)), \infty)$ with $a([t_1, \infty)) < c/2$.

Pick $q \in (1 - c/2, 1)$ and define $r_0 := 2G_1(t_1)/(1-q)$. Since G_1 is increasing, we have $r_0 > G_1(t_1) > \ell_0 \geq \ell_1$. Moreover, since the return times satisfy $S_{(\ell_0)} \leq S_{(\ell_1)}$, for any $x \in \{V \geq r_0\}$ the inequality in (29) yields $\mathbb{P}_x(S_{(\ell_1)} < t_1) \leq \mathbb{P}_x(S_{(\ell_0)} < t_1) < 1 - q < c/2$.

For $x \in \{V \geq r_0\}$, the inequalities $\mathbb{P}_x(V(X_t) \leq \ell_1) \leq \mathbb{P}_x(S_{(\ell_1)} < t) \leq \mathbb{P}_x(S_{(\ell_1)} < t_1) < c/2$ hold for all $t \in [0, t_1]$. Since $a([t_1, \infty)) < c/2$, by (28) the following inequalities hold for all $x \in B \cap \{V \geq r_0\}$,

$$c = \nu_a(\{V \leq \ell_1\}) \leq \int_0^\infty \mathbb{P}_x(V(X_t) \leq \ell_1)a(\mathrm{d}t) \leq \int_0^{t_1} \mathbb{P}_x(V(X_t) \leq \ell_1)a(\mathrm{d}t) + a([t_1, \infty)) < c,$$

implying $B \cap \{V \geq r_0\} = \emptyset$. Put differently, $B \subset \{V < r_0\}$ and the lemma follows. $\qquad\square$

5.3. **Proofs of the main results.** We begin with the proof of the lower bounds on modulated moments stated in Theorem 2.5 above. This theorem will play a crucial role in the analysis of the stability of X and, more specifically, in the proof of Theorem 2.1.

Proof of Theorem 2.5. Fix a set $D \in \mathcal{B}(\mathcal{X})$, contained in $\{V \leq m\}$ for some $m \in (1, \infty)$, $q \in (0,1)$ and $\varepsilon = (1-q)/2$. Since the function $h : [1, \infty) \to [1, \infty)$ in Theorem 2.5 is continuous and non-decreasing by assumption, Proposition 5.3 implies that there exist $\ell_0 \in (m, \infty)$ and $C_{\ell_0} \in (0,1)$, such that the inequality in (26) holds for h, $\ell = \ell_0$ and all $r \in (\ell_0, \infty)$.

By $\mathbf{L(V,\varphi,\Psi)}$, the function $r \mapsto \varepsilon h(r)/(r\varphi(1/r))$ on (ℓ_0, ∞) is increasing and tends to infinity. Define $r_0 := \varepsilon h(\ell_0)/(\ell_0\varphi(1/\ell_0))$ and denote by $G_h : (r_0, \infty) \to (\ell_0, \infty)$ its increasing inverse. Since $C_{\ell_0} \in (0,1)$ and $\Psi : [1, \infty) \to [1, \infty)$, for all $x \in \{V \geq \ell_0 + 1\}$ and $r \in (r_0, \infty)$, the inequality in (26) yields

$$(30) \qquad \mathbb{P}_x\left(\int_0^{S_{(\ell_0)}} h \circ V(X_s)\mathrm{d}s \geq r\right) \geq \frac{qC_{\ell_0}}{\Psi(2G_h(r)/(1-q))}.$$

Note that $\mathbb{E}_x[\int_0^\infty \mathbb{1}\{V(X_s) > \ell_0 + 1\}\mathrm{d}s] > 0$. Indeed, if $\mathbb{P}_x(V(X_s) > \ell_0 + 1) = 0$ for Lebesgue almost every $s \in \mathbb{R}_+$, the right-continuity of X would imply $\sup_{s \in \mathbb{R}_+} V(X_s) \leq \ell_0 + 1$ $\mathbb{P}_x$-a.s., contradicting the assumption $\limsup_{t \to \infty} V(X_t) = \infty$ $\mathbb{P}_x$-a.s. In particular, since the expectation is positive, there exists $\delta > 0$ satisfying $\mathbb{P}_x(V(X_\delta) > \ell_0 + 1) > 0$.

Recall that $\tau_D(\delta) = \inf\{t > \delta : X_t \in D\}$ is the first time, after time $\delta \geq 0$, the process X hits the set $D \in \mathcal{B}(\mathcal{X})$ fixed above. By conditioning at time δ, applying the Markov property of X and

the inequality in (30), we obtain the following lower bound

$$
\mathbb{P}_x\left(\int_0^{\tau_D(\delta)} h \circ V(X_s)\mathrm{d}s \geq r\right) \geq \mathbb{P}_x\left(\int_\delta^{\tau_D(\delta)} h \circ V(X_s)\mathrm{d}s \geq r, V(X_\delta) > \ell_0 + 1\right)
$$

$$
\geq \mathbb{E}_x\left[\mathbb{1}\{V(X_\delta) > \ell_0 + 1\} \cdot \mathbb{P}_{X_\delta}\left(\int_0^{\tau_D(0)} h \circ V(X_s)\mathrm{d}s \geq r\right)\right]
$$

$$
\geq \mathbb{E}_x\left[\mathbb{1}\{V(X_\delta) > \ell_0 + 1\} \cdot \mathbb{P}_{X_\delta}\left(\int_0^{S_{(\ell_0)}} h \circ V(X_s)\mathrm{d}s \geq r\right)\right]
$$

$$
\geq qC_{\ell_0}\mathbb{P}_x(V(X_\delta) > \ell_0 + 1)/\Psi(2G_h(r)/(1-q)) \quad \text{for } r \in (r_0, \infty),
$$

where the third inequality follows from the fact that, since $D \subset \{V \leq m\}$ and $m \leq \ell_0$, starting from any point in $\{V > \ell_0 + 1\}$ the first hitting time $\tau_D(0)$ satisfies $S_{(\ell_0)} \leq \tau_D(0)$. Since δ was chosen so that $\mathbb{P}_x(V(X_\delta) > \ell_0 + 1) > 0$, setting $C := qC_{\ell_0}\mathbb{P}_x(V(X_\delta) > \ell_0 + 1)$ concludes the proof of part (a). Part (b) is a special case of part (a) for the function $h \equiv 1$. $\square$

The following corollary combines the lower bounds of Theorem 2.5 with the fact that, under Assumption $\mathbf{L}(V,\varphi,\Psi)$, any petite set of X is contained in a sublevel set of the Lyapunov function V (see Lemma 5.4 above). The result provides a sufficient condition (in the form of an integral test) for the divergence of the expectation with respect to invariant measure π of a non-decreasing function composed with V.

Corollary 5.5. *Let Assumption $\mathbf{L}(V,\varphi,\Psi)$ hold. Then for every $q \in (0,1)$ and a non-decreasing function $h : [1,\infty) \to [1,\infty)$, the following implication holds:*

$$
(31) \qquad \exists r' \in (0,\infty) \ s.t. \quad \int_{r'}^{\infty} \frac{1}{\Psi(2G_h(r)/(1-q))}\mathrm{d}r = \infty \quad \Longrightarrow \quad \int_{\mathcal{X}} h \circ V(x)\pi(\mathrm{d}x) = \infty,
$$

where G_h is the inverse of the increasing function $r \mapsto (1-q)h(r)/(2r\varphi(1/r))$.

Proof. By the standard assumption in the paper, stated above Assumption $\mathbf{L}(V,\varphi,\Psi)$, the process X is positive Harris recurrent. The seminal result [25, Thm 1.2(b)] implies that a measurable $h : [1,\infty) \to [1,\infty)$ satisfies the implication:

$$
\int_{\mathcal{X}} h \circ V(x)\pi(\mathrm{d}x) < \infty \quad \Longrightarrow \quad \exists \text{ closed petite } D \text{ s.t. } \forall \delta > 0, \ \sup_{x \in D} \mathbb{E}_x\left[\int_0^{\tau_D(\delta)} h \circ V(X_s)\mathrm{d}s\right] < \infty.
$$

By Lemma 5.4, *every* petite set D for X satisfies $D \subset \{V \leq r_0\}$ for some $r_0 \in [1,\infty)$. Thus, Theorem 2.5 implies that for a non-decreasing $h : [1,\infty) \to [1,\infty)$, every closed petite set D and any $x \in D$, there exist $\delta > 0$ and $C \in (0,\infty)$ such that

$$
\mathbb{E}_x\left[\int_0^{\tau_D(\delta)} h \circ V(X_s)\mathrm{d}s\right] = \int_0^{\infty} \mathbb{P}_x\left(\int_0^{\tau_D(\delta)} h \circ V(X_s)\mathrm{d}s \geq r\right)\mathrm{d}r \geq \int_{r_0}^{\infty} \frac{C}{\Psi(2G_h(r)/(1-q))}\mathrm{d}r
$$

for some sufficiently large $r_0 \in (0,\infty)$, where G_h is the inverse of the increasing function $r \mapsto (1-q)h(r)/(2r\varphi(1/r))$. If the assumption in the implication in (31) holds, then the last integral in the previous display must also be infinite because the function $r \mapsto 1/\Psi(2G_h(r)/(1-q))$ is continuous and thus locally bounded. The criterion in [25, Thm 1.2(b)] stated above thus yields the conclusion of the implication in (31). $\square$

The implication in (31) in Corollary 5.5 is at the core of the proof of Theorem 2.1. It is key that integral test (31) covers all non-decreasing functions h, not only the polynomial ones.

Proof of Theorem 2.1. Pick $q, \varepsilon \in (0, 1)$ and note that the statement in display (4) of the theorem is equivalent to the following:

$$\exists r_0 \in (0, \infty) \text{ such that,} \quad 1/L_{\varepsilon,q}(r) \leq \pi(\{x \in \mathcal{X} : V(x) \geq r\}) \quad \text{for all } r \in [r_0, \infty),$$

where $L_{\varepsilon,q}(r) = r\varphi(1/r)\Psi(2r/(1-q))(\log \log r)^\varepsilon$. Assume $\mathbf{L}(V, \varphi, \Psi)$ holds.

The proof is by contradiction. Assume that there exists $\varepsilon > 0$, such that *for every* $r_0 \in (0, \infty)$ there exists $r_1 \in [r_0, \infty)$ satisfying $1/L_{\varepsilon,q}(r_1) > \pi(\{x \in \mathcal{X} : V(x) \geq r_1\})$. We may pick $r_0 > 1$ and $r_1 > \exp(\exp(\exp(1)))r_0$. Recursively we can define an increasing sequence $(r_n)_{n \in \mathbb{N}}$, satisfying $r_{n+1} > \exp(\exp(\exp(n+1)))r_n$ and $1/L_{\varepsilon,q}(r_n) > \pi(\{x \in \mathcal{X} : V(x) \geq r_n\})$ for all $n \in \mathbb{N}$. In particular, since $r_0 > 1$, we have

$$(32) \qquad \log \log r_n > \exp(n) \qquad \text{for all } n \in \mathbb{N}.$$

Using the sequence $(r_n)_{n \in \mathbb{N}}$, we construct a non-decreasing function $h : [1, \infty) \to [1, \infty)$, satisfying $\int_{\mathcal{X}} h \circ V(x)\pi(\mathrm{d}x) < \infty$ *and* the assumption of the implication in (31).

Define the function $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ by $\mu(r) := 1$ for $r \in [0, r_1)$ and $\mu(r) := 1/L_{\varepsilon,q}(r_n)$ for $r \in [r_n, r_{n+1})$, $n \in \mathbb{N}$. Since the function $r \mapsto \pi(\{V \geq r\})$ is non-increasing, we have $\pi(\{V \geq r\}) \leq \mu(r)$ for all $r \in \mathbb{R}_+$. Let $h : [1, \infty) \to [1, \infty)$ be a differentiable function such that $h(r) = 1$ for $r \in [1, r_1)$. For $n \in \mathbb{N} \setminus \{1\}$ and $r \in [r_n, r_{n+1})$ we define the derivative of h by

$$(33) \qquad h'(r) = \begin{cases} r\varphi(1/r)\Psi(2(r_n+1)/(1-q))(\log \log r_n)^{\varepsilon/2}, & r \in [r_n, r_n + 1); \\ 1/(r_n(r_{n+1} - r_n)), & r \in [r_n + 1, r_{n+1}). \end{cases}$$

Since, by Assumption $\mathbf{L}(V, \varphi, \Psi)$, $r \mapsto r\varphi(1/r)$ is decreasing and Ψ is differentiable, increasing and submultiplicative (i.e. $\Psi(2(r_n+1)/(1-q)) \leq \Psi(2r_n/(1-q))\Psi(2/(1-q))$ for all $r_n \in [1, \infty)$; without loss of generality we assume here that the constant C in definition of a submultiplicative function in footnote on page 3 equals one, since we may substitute Ψ in $\mathbf{L}(V, \varphi, \Psi)\mathbf{(ii)}$ with $C\Psi$ if $C > 1$), we have

$$h'(r)\mu(r) \leq \begin{cases} \Psi(2/(1-q))(\log \log r_n)^{-\varepsilon/2}, & r \in [r_n, r_n + 1); \\ 1/(r_n(r_{n+1} - r_n)), & r \in [r_n + 1, r_{n+1}). \end{cases}$$

The identity $1 + \int_1^{V(x)} h'(r)\mathrm{d}r = h(V(x))$ for all $x \in \mathcal{X}$ and Fubini's theorem imply the equality $\int_{\mathcal{X}} h(V(x))\pi(\mathrm{d}x) = 1 + \int_1^\infty h'(r)\pi(\{V \geq r\})\mathrm{d}r$. Recall $\pi(\{V \geq r\}) \leq \mu(r)$ for $r \in \mathbb{R}_+$ and note

$$\int_{\mathcal{X}} h(V(x))\pi(\mathrm{d}x) = 1 + \int_{r_1}^\infty h'(r)\pi(\{V \geq r\})\mathrm{d}r \leq 1 + \int_{r_1}^\infty h'(r)\mu(r)\mathrm{d}r$$

$$= 1 + \sum_{n=1}^\infty \left(\int_{r_n}^{1+r_n} h'(r)\mu(r)\mathrm{d}r + \int_{1+r_n}^{r_{n+1}} h'(r)\mu(r)\mathrm{d}r \right)$$

$$(34) \qquad \leq 1 + \Psi(2/(1-q)) \sum_{n=1}^\infty (\log \log r_n)^{-\varepsilon/2} + \sum_{n=1}^\infty 1/r_n < \infty,$$

where the final inequality follows from (32), which makes both sums in (34) clearly finite.

Recall that the function $u \mapsto (1-q)h(u)/(2u\varphi(1/u))$ is increasing on $[1, \infty)$ by $\mathbf{L}(V, \varphi, \Psi)$ and define $r' := (1-q)h(1)/(2\varphi(1)) > 0$ (in fact $h(1) = 1$). Denote by $G_h : [r', \infty) \to [1, \infty)$ the inverse and introduce the substitution $r = (1-q)h(u)/(2u\varphi(1/u))$ into the following integral:

$$\int_{r'}^{\infty} 2/((1-q)\Psi(2G_h(r)/(1-q)))\mathrm{d}r = \int_1^{\infty} (h(u)/(u\varphi(1/u)))'/\Psi(2u/(1-q))\mathrm{d}u$$

$$= \int_1^{\infty} (h'(u)/(u\varphi(1/u)) + (1/(u\varphi(1/u)))'h(u))/\Psi(2u/(1-q))\mathrm{d}u$$

$$\geq \sum_{n=n_0}^{\infty} \int_{r_n}^{1+r_n} \frac{u\varphi(1/u)\Psi(2(r_n+1)/(1-q))(\log\log r_n)^{\varepsilon/2}}{\Psi(2u/(1-q))u\varphi(1/u)}\mathrm{d}u$$

$$(35) \qquad \geq \sum_{n=n_0}^{\infty} \int_{r_n}^{1+r_n} (\log\log r_n)^{\varepsilon/2}\mathrm{d}u = \sum_{n=n_0}^{\infty} (\log\log r_n)^{\varepsilon/2} = \infty.$$

The first inequality follows from the definition of h' given in (33) above and the fact that the function $u \mapsto 1/(u\varphi(1/u))$ is continuous and increasing and hence almost everywhere differentiable with a non-negative derivative. The second inequality in the previous display follows from the fact that Ψ is increasing by Assumption $\mathbf{L}(V, \varphi, \Psi)$. The divergence of the sum is a consequence of the inequality in (32). By Corollary 5.5, the inequality in (35) implies $\int_{\mathcal{X}} h \circ V(x)\pi(\mathrm{d}x) = \infty$, which contradicts (34) and concludes the proof of Theorem 2.1. $\qquad\square$

The drift condition on the Lyapunov function V in Assumption $\mathbf{L}(V, \varphi, \Psi)$ and the lower bound on the invariant measure π from Theorem 2.1 are the key ingredients in the proof of the lower bound on the rate of convergence in total variation. For the f-variation distance we require the following corollary of Theorem 2.1.

Corollary 5.6. *Let Assumption $\mathbf{L}(V, \varphi, \Psi)$ hold. Let $f_\star : [1, \infty) \to [1, \infty)$ be a differentiable function and consider an increasing continuous $g : [1, \infty) \to [1, \infty)$, satisfying $\lim_{r\to\infty} g(r) = \infty$. Then, for every $\varepsilon, q \in (0, 1)$ and the function $L_{\varepsilon,q}$ in (3), there exists $c_{\varepsilon,q} \in (0, \infty)$ such that*

$$(36) \qquad \int_{\{g \circ V \geq r\}} f_\star \circ V(x)\pi(\mathrm{d}x) \geq c_{\varepsilon,q}f_\star(g^{-1}(r))/L_{\varepsilon,q}(g^{-1}(r)) \quad \text{for all } r \in [1, \infty).$$

Proof. Pick $\varepsilon, q \in (0, 1)$. Then, by Theorem 2.1, there exists a constant $c_{\varepsilon,q} \in (0, \infty)$ such that

$$\pi(\{V \geq g^{-1}(r)\}) \geq c_{\varepsilon,q}/L_{\varepsilon,q}(g^{-1}(r)) \quad \text{for all } r \in [1, \infty).$$

Using the inequality above along with the facts that f is differentiable and g is continuous increasing and thus has an inverse g^{-1}, we obtain

$$\int_{\{g \circ V \geq r\}} f_\star \circ V(x)\pi(\mathrm{d}x) = \int_{\{V \geq g^{-1}(r)\}} \left(f_\star(1) + \int_1^{V(x)} f_\star'(y)\mathrm{d}y\right)\pi(\mathrm{d}x)$$

$$= f_\star(1)\pi(\{V \geq g^{-1}(r)\}) + \int_1^{\infty} f_\star'(y)\pi(\{V \geq \max\{g^{-1}(r), y\}\})\mathrm{d}y$$

$$\geq f_\star(1)\pi(\{V \geq g^{-1}(r)\}) + \int_1^{g^{-1}(r)} f_\star'(y)\pi(\{V \geq g^{-1}(r)\})\mathrm{d}y$$

$$= f_\star(g^{-1}(r))\pi(\{V \geq g^{-1}(r)\}) \geq c_{\varepsilon,q}f_\star(g^{-1}(r))/L_{\varepsilon,q}(g^{-1}(r)). \qquad\square$$

We now establish the lower bound on the convergence in f-variation in Theorem 2.2.

Proof of the Theorem 2.2. By Assumption (a) in Theorem 2.2, a differentiable $f_\star : [1, \infty) \to [1, \infty)$ and continuous $h, g : [1, \infty) \to [1, \infty)$ satisfy $g = h/f_\star$ on $[1, \infty)$, with g increasing and $\lim_{r \to \infty} g(r) = \infty$. Moreover, there exists a function $v : \mathcal{X} \times \mathbb{R}_+ \to [1, \infty)$, increasing in t, such that $\mathbb{E}_x[h \circ V(X_t)] \leq v(x, t)$ for all $x \in \mathcal{X}$ and $t \in \mathbb{R}_+$. By Corollary 5.6, for any $\varepsilon, q \in (0, 1)$ there exists a constant $c_{\varepsilon,q} \in (0, \infty)$, such that the inequality in (36) holds.

Any function $a : [1, \infty) \to \mathbb{R}_+$, satisfying Assumption (b) in Theorem 2.2 (i.e. the inequality $a(r) \leq c_{\varepsilon,q} f_\star(g^{-1}(r))/L_{\varepsilon,q}(g^{-1}(r))$ holds for $r \in [1, \infty)$ and the function $r \mapsto ra(r)$ is increasing with $\lim_{r \to \infty} ra(r) = \infty$), by (36) also satisfies Assumption (a) of Lemma 5.1 with $f = f_\star \circ V$ and $G := g \circ V$. As observed in the previous paragraph, the functions $H := h \circ V$ and $v(x, t)$ satisfy the condition in Assumption (b) of Lemma 5.1: $\mathbb{E}_x[H(X_t)] \leq v(x, t)$ for all $x \in \mathcal{X}$ and $t \in \mathbb{R}_+$. An application of Lemma 5.1 concludes the proof of the theorem. $\square$

Proof of Lemma 2.8. Pick $\ell \in (\ell_0, \infty)$, $r \in (\ell + 1, \infty)$ and $x \in \{\ell + 1 \leq V < r\}$, and recall the definitions $T^{(r)} := \inf\{t \geq 0 : V(X_t) > r\}$ and $S_{(\ell)} := \inf\{t \geq 0 : V(X_t) < \ell\}$. Assumption of the lemma implies that for some $d \in [1, \infty)$, we have $V(X_{T^{(r)}}) - V(X_{T^{(r)}-}) \leq d$ $\mathbb{P}_x$-a.s, and since Ψ is increasing we obtain $\Psi \circ V(X_{t \wedge S_{(\ell)} \wedge T^{(r)}}) \leq \Psi(r + d)$ for all $t \in \mathbb{R}_+$ $\mathbb{P}_x$-a.s. Moreover, given that $\ell \in (\ell_0, \infty)$ it follows that $\int_0^{\cdot \wedge S_{(\ell)} \wedge T^{(r)}} \mathbb{1}\{V(X_u) \leq \ell_0\} \equiv 0$ $\mathbb{P}_x$-a.s. Thus, by the assumption of the lemma and the optional sampling theorem the process $\Psi \circ V(X_{\cdot \wedge S_{(\ell)} \wedge T^{(r)}})$ is an $(\mathcal{F}_t)$-submartingale under $\mathbb{P}_x$.

We establish a lower bound on $\mathbb{P}_x(T^{(r)} < S_{(\ell)})$ as follows. By assumption in Lemma 2.8 we have $\limsup_{t \to \infty} V(X_t) = \infty$ $\mathbb{P}_x$-a.s., which implies $T^{(r)} \wedge S_{(\ell)} \leq T^{(r)} < \infty$ $\mathbb{P}_x$-a.s. The dominated convergence theorem and the monotonicity of Ψ yield

$$\Psi(V(x)) \leq \lim_{t \to \infty} \mathbb{E}_x[\Psi \circ V(X_{t \wedge S_{(\ell)} \wedge T^{(r)}})] = \mathbb{E}_x[\Psi \circ V(X_{S_{(\ell)} \wedge T^{(r)}})] \leq \Psi(\ell) + \mathbb{P}_x(T^{(r)} < S_{(\ell)})\Psi(r + d).$$

Thus, $\mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq (\Psi(V(x)) - \Psi(\ell))/\Psi(r + d) \geq (\Psi(V(x)) - \Psi(\ell))/(C\Psi(r)\Psi(d)) \geq C_\ell/\Psi(r)$, where $C_\ell := (\Psi(\ell + 1) - \Psi(\ell))/(C\Psi(d))$. The second inequality holds since Ψ is submultiplicative (with a constant $C > 0$) and the third holds because Ψ is increasing (both properties are assumed in the lemma). Noting $C_\ell > 0$ concludes the proof of the inequality in $\mathbf{L}(V, \varphi, \Psi)\mathbf{(ii)}$. $\square$

Proof of Theorem 2.7. Supermartingale condition (a). Fix arbitrary $x \in \mathcal{X}$ and let $\ell_0 \in (1, \infty)$ be such that the inequality in (7) holds. By assumption in Theorem 2.7(a), we have $1/V \in \mathcal{D}(\mathcal{A})$. Thus, by [8, Ch 1, Def (14.15)], there exists an increasing sequence $\{T_n : n \in \mathbb{N}\}$ of $(\mathcal{F}_t)$-stopping times, satisfying $T_n \uparrow \infty$ as $n \to \infty$ $\mathbb{P}_x$-a.s. and the localised process

$$1/V(X_{\cdot \wedge T_n}) - 1/V(x) - \int_0^{\cdot \wedge T_n} \mathcal{A}(1/V)(X_s)\mathrm{d}s \quad \text{is an } (\mathcal{F}_t)\text{-martingale under } \mathbb{P}_x \text{ for all } n \in \mathbb{N}.$$

By assumptions of the theorem, we have $0 < 1/V(x') \leq 1$ for $x' \in \mathcal{X}$ and $0 \leq \varphi(u) \leq \varphi(1)$ for $u \in (0, 1]$, implying $\mathbb{E}_x[1/V(X_{t \wedge T_n})] < \infty$ and $\mathbb{E}_x[\int_0^t \varphi(1/V(X_s))\mathrm{d}s] < \infty$ for all $t \in \mathbb{R}_+$,

respectively. By the inequality in (7), there exists $b \in \mathbb{R}_+$, such that for every $n \in \mathbb{N}$ we have

$$\mathbb{E}_x[1/V(X_{t \wedge T_n})] - \mathbb{E}_x\left[\int_0^{t \wedge T_n} \varphi(1/V(X_s))\mathrm{d}s\right]$$

$$= 1/V(x) + \mathbb{E}_x\left[\int_0^{t \wedge T_n} (\mathcal{A}(1/V)(X_s) - \varphi(1/V(X_s)))\mathrm{d}s\right]$$

$$\leq 1/V(x) + b\,\mathbb{E}_x\left[\int_0^{t \wedge T_n} \mathbb{1}\{V(X_s) \leq \ell_0\}\mathrm{d}s\right] \qquad \text{for all } t \in \mathbb{R}_+.$$

This inequality, Fatou's lemma and the monotone convergence theorem, yield

$$\mathbb{E}_x[1/V(X_t)] = \mathbb{E}_x[\liminf_{n \to \infty} 1/V(X_{t \wedge T_n})] \leq \liminf_{n \to \infty} \mathbb{E}_x[1/V(X_{t \wedge T_n})]$$

$$\leq \liminf_{n \to \infty} \left(1/V(x) + \mathbb{E}_x\left[\int_0^{t \wedge T_n} \varphi(1/V(X_s))\mathrm{d}s\right] + b\,\mathbb{E}_x\left[\int_0^{t \wedge T_n} \mathbb{1}\{V(X_s) \leq \ell_0\}\mathrm{d}s\right]\right)$$

$$= 1/V(x) + \mathbb{E}_x\left[\int_0^t \varphi(1/V(X_s))\mathrm{d}s\right] + b\,\mathbb{E}_x\left[\int_0^t \mathbb{1}\{V(X_s) \leq \ell_0\}\mathrm{d}s\right].$$

This proves the condition **(i)** in Assumption $\mathbf{L}(V,\varphi,\Psi)$.

Exit probability condition (b). We employ analogous arguments to show that for some $\ell_0, c \in (0, \infty)$ and all $r \in (\ell_0, \infty)$, the process

$$\Psi \circ V(X_{\cdot \wedge T^{(r)}}) + c\int_0^{\cdot \wedge T^{(r)}} \mathbb{1}\{V(X_u) \leq \ell_0\}\mathrm{d}u, \quad \text{is an } (\mathcal{F}_t)\text{-submartingale under } \mathbb{P}_x \text{ for all } x \in \mathcal{X}.$$

The condition in $\mathbf{L}(V,\varphi,\Psi)\mathbf{(ii)}$ then follows from the application of Lemma 2.8.

Fix arbitrary $x \in \mathcal{X}$ and let $\ell_0 \in (1, \infty)$ be such that the inequality (8) hold. Since $\Psi \circ V \in \mathcal{D}(\mathcal{A})$ (by assumption in Theorem 2.7(b)), as before there exists a localising sequence of $(\mathcal{F}_t)$-stopping times $\{T_n : n \in \mathbb{N}\}$, such that $(M_t^n)_{t \in \mathbb{R}_+}$, where

$$M_t^n := \Psi \circ V(X_{t \wedge T_n}) - \Psi \circ V(x) - \int_0^{t \wedge T_n} \mathcal{A}(\Psi \circ V)(X_s)\mathrm{d}s,$$

is an $(\mathcal{F}_t)$-martingale under $\mathbb{P}_x$. Thus, for any $r \in (\ell_0, \infty)$, the stopped process $(M_{t \wedge T^{(r)}}^n)_{t \in \mathbb{R}_+}$ is also an $(\mathcal{F}_t)$-martingale under $\mathbb{P}_x$ (recall $T^{(r)} = \inf\{t \geq 0 : V(X_t) > r\}$). Moreover, by assumption in Theorem 2.7(b) there exists $d \in [1, \infty)$ such that $V(X_{t \wedge T^{(r)}}) - V(X_{t \wedge T^{(r)}-}) \leq r + d$ for all $t \in \mathbb{R}_+$ and $r \in (\ell_0, \infty)$ $\mathbb{P}_x$-a.s. The fact that Ψ is increasing implies $\Psi \circ V(X_{t \wedge T_n \wedge T^{(r)}}) \leq \Psi(r + d)$ for all $t \in \mathbb{R}_+$, $r \in (\ell_0, \infty)$ and $n \in \mathbb{N}$ $\mathbb{P}_x$-a.s. Since $\mathbb{E}_x[M_{t \wedge T^{(r)}}^n] = 0$, by the inequality in (8), there exists $c \in \mathbb{R}_+$, such that for every $n \in \mathbb{N}$,

$$\mathbb{E}_x[\Psi \circ V(X_{t \wedge T_n \wedge T^{(r)}})] = \Psi \circ V(x) + \mathbb{E}_x\left[\int_0^{t \wedge T_n \wedge T^{(r)}} \mathcal{A}(\Psi \circ V)(X_s)\mathrm{d}s\right]$$

$$\geq \Psi \circ V(x) - c\,\mathbb{E}_x\left[\int_0^{t \wedge T_n \wedge T^{(r)}} \mathbb{1}\{V(X_s) \leq \ell_0\}\mathrm{d}s\right], \ t \in \mathbb{R}_+, r \in (\ell_0, \infty).$$

The dominated convergence theorem (as $n \to \infty$), applied to both sides of the inequality, yields the submartingale condition, and by Lemma 2.8, Assumption $\mathbf{L}(V,\varphi,\Psi)\mathbf{(ii)}$. $\qquad\square$

Proof of Lemma 2.10. Pick $x \in \mathcal{X}$ and recall $H \in \mathcal{D}(\mathcal{A})$. Thus, by [8, Ch 1, Def (14.15)], there exists an increasing sequence $\{T_n : n \in \mathbb{N}\}$ of $(\mathcal{F}_t)$-stopping times, such that $T_n \uparrow \infty$ as $n \to \infty$

$\mathbb{P}_x$-a.s. and $M^{(n)} := H(X_{\cdot \wedge T_n}) - H(x) - \int_0^{\cdot \wedge T_n} \mathcal{A}H(X_s)\mathrm{d}s$ is an $(\mathcal{F}_t)$-martingale under $\mathbb{P}_x$ for all $n \in \mathbb{N}$. Set $S_m := \inf\{t > 0 : H(X_t) \geq m\} \wedge T_m$ for all $m \in \mathbb{N}$ and note $S_m \uparrow \infty$ as $m \to \infty$ $\mathbb{P}_x$-a.s. Since $M_t^{(m)}$ is integrable and, on the event $\{t < S_m\}$, $H(X_{t \wedge S_m})$ is bounded, we have $\mathbb{E}_x[\int_0^{t \wedge S_m} \mathcal{A}H(X_s)\mathrm{d}s] < \infty$. Moreover, since $\mathcal{A}H \leq \xi \circ H$ on $\mathcal{X}$, for any $t \in \mathbb{R}_+$ and $m \in \mathbb{N}$ we obtain

$$\mathbb{E}_x[H(X_{t \wedge S_m})] = H(x) + \mathbb{E}_x\left[\int_0^{t \wedge S_m} \mathcal{A}H(X_s)\mathrm{d}s\right] \leq H(x) + \mathbb{E}_x\left[\int_0^{t \wedge S_m} \xi \circ H(X_s)\mathrm{d}s\right]$$

This inequality, Fatou's lemma and the monotone convergence theorem, yield

$$\mathbb{E}_x[H(X_t)] = \mathbb{E}_x[\liminf_{m \to \infty} H(X_{t \wedge S_m})] \leq \liminf_{m \to \infty} \mathbb{E}_x[H(X_{t \wedge S_m})]$$

$$\leq H(x) + \liminf_{m \to \infty} \mathbb{E}_x\left[\int_0^{t \wedge S_m} \xi \circ H(X_s)\mathrm{d}s\right] = H(x) + \mathbb{E}_x\left[\int_0^t \xi \circ H(X_s)\mathrm{d}s\right].$$

Since $\xi : [1, \infty) \to [1, \infty)$ is concave, Tonelli's theorem and Jensen's inequality imply

$$\mathbb{E}_x\left[\int_0^t \xi \circ H(X_s)\mathrm{d}s\right] = \int_0^t \mathbb{E}_x\left[\xi \circ H(X_s)\right]\mathrm{d}s \leq \int_0^t \xi\left(\mathbb{E}_x[H(X_s)]\right)\mathrm{d}s \quad \text{for all } t \in \mathbb{R}_+.$$

Denote $g(t) := \mathbb{E}_x[H(X_t)] \geq 1$. Thus $g(0) = H(x)$ and

$$(37) \qquad\qquad g(t) \leq g(0) + \int_0^t \xi(g(s))\mathrm{d}s \quad \text{for all } t \in \mathbb{R}_+.$$

The increasing function $\Xi : [1, \infty) \to \mathbb{R}_+$, given by $\Xi(t) := \int_1^t \mathrm{d}s/\xi(s)$, has a differentiable inverse Ξ^{-1}. Denote $G(t) := \int_0^t \xi(g(s))\mathrm{d}s$. By (37) we have $G'(v) = \xi(g(v)) \leq \xi(g(0)+G(v))$ for all $v \in \mathbb{R}_+$, since ξ is non-decreasing. This yields

$$\Xi(g(0) + G(t)) - \Xi(g(0)) = \int_{g(0)}^{g(0)+G(t)} \frac{\mathrm{d}z}{\xi(z)} = \int_0^{G(t)} \frac{\mathrm{d}z}{\xi(g(0)+z)} = \int_0^t \frac{G'(v)}{\xi(g(0)+G(v))}\mathrm{d}v \leq t.$$

Thus $\Xi(g(0) + G(t)) \leq \Xi(g(0)) + t$ and hence $g(t) \leq g(0) + G(t) \leq \Xi^{-1}(\Xi(g(0)) + t)$. □

6. Proofs of the examples in Section 3

This section is dedicated to proving the theorems presented in Section 3. All strong Markov processes X considered in this section are ergodic with an invariant measure π and Feller continuous (see [9, Sec. 5] for more details), and thus positive Harris recurrent by [25, Thm 1.1]. Moreover, all the models in this section are irreducible, either because of uniform ellipticity and the irreducibility of the driver (a Brownian motion or, more generally, a Lévy process) or by a direct argument in the hypoelliptic case. Thus, for any Lyapunov function V for X, we will have $\mathbb{P}_x(V(X_t) \geq r_0) > 0$ for all starting points x, times $t > 0$ and levels r_0, thus satisfying the assumptions of Lemma A.1 below and implying non-confinement. Verifying the **L**-drift condition $\mathbf{L}(V, \varphi, \Psi)$ in our examples will thus reduce to finding functions V, φ, Ψ, which satisfy the conditions **(i)** and **(ii)** in $\mathbf{L}(V, \varphi, \Psi)$. Since we are working with Feller processes, we will obtain these conditions by applying the generator to the relevant functions, establishing the appropriate point-wise inequalities and using Theorem 2.7.

6.1. **Diffusions from Section 3.1.** Since the SDE in (11) possesses a unique strong solution, the process X is strong Markov. Feller continuity follows from [17, Thm 3.4.1]. By Itô's formula applied to $g(X)$, the extended generator (see Section 2.4 above for definition) of the diffusion X takes the following form for any twice continuously differentiable $g \in C^2(\mathbb{R}^n)$:

$$(38) \qquad \mathcal{A}g(x) = \langle b(x), \nabla g(x)\rangle + \frac{1}{2}\operatorname{Tr}\left(\Sigma(x)\operatorname{Hess}(g)(x)\right), \qquad \text{for all } x \in \mathbb{R}^n,$$

where $\Sigma = \sigma\sigma^\intercal$ is the instantaneous covariance of X, ∇g is the gradient of g, $\operatorname{Hess}(g)$ is the Hessian (i.e. the matrix of the second derivatives of g) and $\operatorname{Tr}(\cdot)$ denotes the trace of a matrix in $\mathbb{R}^{n\times n}$.

6.1.1. *Polynomial tails: proofs for Section 3.1.1.* For any $m \in \mathbb{R}\setminus\{0\}$, consider a function $p_m : \mathbb{R}^n \to (0,\infty)$ in $C^2(\mathbb{R}^n)$, satisfying $p_m(x) \leq 1 + |x|^m$ for all $x \in \mathbb{R}^n$ and

$$(39) \qquad p_m(x) = |x|^m \quad \text{for all } x \in \mathbb{R}^n \text{ with } |x| \text{ sufficiently large.}$$

If $m > 0$, we assume in addition that $p_m : \mathbb{R}^n \to [1,\infty)$ takes values in $[1,\infty)$ only. Recall parameters $\alpha, \beta, \gamma \in (0,\infty)$, $\ell \in [0,2)$ and m_c in Assumption A_p. The following (deterministic) proposition allows us to construct the functions in $\mathbf{L(V,\varphi,\Psi)}$.

Proposition 6.1. *Under Assumption A_p, extended generator (38) of the diffusion X in (11) satisfies the following asymptotic inequalities: if $m \in (-\infty, m_c)\setminus\{0\}$ (resp. $m \in (m_c,\infty)$), then $\mathcal{A}p_m(x) \leq C_0|x|^{m+\ell-2}$ (resp. $\mathcal{A}p_m(x) \geq C_0|x|^{m+\ell-2}$) for all $x \in \mathbb{R}^n$ with $|x|$ sufficiently large and $C_0 := \max\{-m\beta(m_c - m), -m\beta(m_c - m)/4\} \in \mathbb{R}\setminus\{0\}$ (resp. $C_0 := m\beta(m - m_c)/4 \in (0,\infty)$).*

Proof. Since $\nabla p_m(x) = m|x|^{m-2}x$ and $\operatorname{Hess}(p_m)(x) = m|x|^{m-2}((m-2)xx^\intercal/|x|^2 + I_n)$ for all $m \in \mathbb{R}\setminus\{0\}$ and $x \in \mathbb{R}^n$ with sufficiently large $|x|$ ($I_n \in \mathbb{R}^{n\times n}$ is the identity matrix), the representation of $\mathcal{A}$ in (38), the identity $\langle\Sigma(x)x/|x|, x/|x|\rangle = \operatorname{Tr}\left(\Sigma(x)xx^\intercal/|x|^2\right)$ for all points $x \in \mathbb{R}^n\setminus\{0\}$ and Assumption A_p yield

$$(40) \quad \mathcal{A}p_m(x) = -\frac{m\beta}{2}\left(\frac{2\alpha - \gamma}{\beta} - (m-2) + o(1)\right)|x|^{m+\ell-2} = -\frac{m\beta}{2}(m_c - m + o(1))|x|^{m+\ell-2},$$

for all $x \in \mathbb{R}^n$ with $|x|$ sufficiently large. Recall $m_c > 0$ by Assumption A_p. If $m \in (-\infty, m_c)\setminus\{0\}$ (resp. $m \in (m_c,\infty)$), then a constant C_0 and the inequalities follow from representation (40). □

As with all of the results in this section, Theorem 3.2 is a direct consequence of the theory developed in Section 2 applied to an appropriate class (in this case polynomial) of Lyapunov functions. Its proof is straightforward, but somewhat tedious. It consist of verifying the assumptions of Theorem 2.7 and translating them into lower bounds via Theorems 2.1 and 2.5(b).

Proof of Theorem 3.2. Pick $\varepsilon \in (0, m_c)$. Let $V_\varepsilon := p_{m_c-\varepsilon}$ be a $C^2(\mathbb{R}^n)$ function in (39). By Proposition 6.1, we have $\mathcal{A}(1/V_\varepsilon)(x) = \mathcal{A}p_{\varepsilon-m_c}(x) \leq C_0|x|^{\varepsilon-m_c+\ell-2}$ for $C_0 > 0$ and all $x \in \mathbb{R}^n$ with large norm $|x|$. For $r \in [1,\infty)$, define $\varphi_\varepsilon(1/r) := C_0 r^{(\varepsilon-m_c+\ell-2)/(m_c-\varepsilon)}$. Then $\mathcal{A}(1/V_\varepsilon) \leq \varphi(1/V_\varepsilon)$ holds outside of a large ball centered at the origin, implying condition (7) in Theorem 2.7. Moreover, since the function $r \mapsto 1/(r\varphi_\varepsilon(1/r)) = r^{(2-\ell)/(m_c-\varepsilon)}/C_0$ is increasing with infinite limit as $r \to \infty$, all assumptions in Theorem 2.7(a) concerning φ_ε are satisfied. Define differentiable, increasing and submultiplicative function $\Psi_\varepsilon(r) := r^{1+2\varepsilon/(m_c-\varepsilon)}$, $r \in [1,\infty)$, and note (by (39)) that $\Psi_\varepsilon \circ V_\varepsilon = p_{m_c+\varepsilon}$. By Proposition 6.1, we have $\mathcal{A}(\Psi_\varepsilon \circ V_\varepsilon) \geq 0$ outside of a large ball centered at the

origin, implying conditions in Theorem 2.7(b). Thus, by Theorem 2.7, L-drift condition $\mathbf{L(V,\varphi,\Psi)}$ holds with $(V,\varphi,\Psi) \coloneqq (V_\varepsilon, \varphi_\varepsilon, \Psi_\varepsilon)$ for the diffusion X in SDE (11) and every $\varepsilon \in (0, m_c)$.

The function $L_{\varepsilon,q}$ in (3) (with $q \in (0,1)$) satisfies $L_{\varepsilon,q}(r) = r\varphi(1/r)\Psi(2r/(1-q))(\log\log r)^\varepsilon \leq r^{1+(\ell-2+2\varepsilon)/(m_c-\varepsilon)}(\log\log r)^\varepsilon/C_1$ for some constant $C_1 > 0$ and all $r \in [1,\infty)$. Since $V = V_\varepsilon$, by Theorem 2.1, there exists $c'_\pi > 0$ such that

$$\pi(\{|x| \geq r\}) = c'_\pi/L_{\varepsilon,q}(r^{m_c-\varepsilon}) \geq c'_\pi C_1/(r^{\ell-2+m_c+\varepsilon}(\log\log(r^{m_c-\varepsilon}))^\varepsilon) \geq c_\pi/r^{\ell-2+m_c+2\varepsilon}$$

for all $r \in [1,\infty)$ and a sufficiently small constant $c_\pi > 0$, implying part (a) of the theorem.

Theorem 2.5(b) provides a lower bound on the tail of the return time $\tau_D(\delta)$. The inverse function $G_1(t)$ (of the function proportional to $r \mapsto 1/(r\varphi_\varepsilon(1/r)) = r^{(2-\ell)/(m_c-\varepsilon)}/C_0$) is proportional to $t \mapsto t^{(m_c-\varepsilon)/(2-\ell)}$, implying in particular that there exist $t_0, C' > 0$ such that $\Psi(2G_1(t)/(1-q)) \leq C't^{(m_c+\varepsilon)/(2-\ell)}$ for all $t \in (t_0,\infty)$. Since $\varepsilon \in (0, m_c)$ can be chosen to be arbitrarily small, part (b) follows. By Remark 2.6, following Theorem 2.5 above, the lower bound holds for all $\delta > 0$ since the diffusion X in SDE (11) has full support at every positive time. $\qquad\square$

Proof of Theorem 3.1. Pick $\varepsilon \in (0, (2-\ell)/3)$. Let $V_\varepsilon = p_{m_c-\varepsilon}$ (where $p_{m_c-\varepsilon}$ is a $C^2(\mathbb{R}^n)$ function in (39)), $\Psi_\varepsilon(r) = r^{1+2\varepsilon/(m_c-\varepsilon)}$ and $\varphi_\varepsilon(1/r) = C_0 r^{(\varepsilon-m_c+\ell-2)/(m_c-\varepsilon)}$, for $r \in [1,\infty)$ and some constant $C_0 \in (0,\infty)$. Recall, from the proof of Theorem 3.2, that the $\mathbf{L}$-drift condition $\mathbf{L(V,\varphi,\Psi)}$ holds with $(V,\varphi,\Psi) \coloneqq (V_\varepsilon, \varphi_\varepsilon, \Psi_\varepsilon)$.

Pick $k \in [0, \ell + (2\alpha-\gamma)/\beta)$ and note $k < 2 + (2\alpha-\gamma)/\beta = m_c$. Consider the functions $h, f_\star, g : [1,\infty) \to [1,\infty)$ given by $h(r) = r$, $f_\star(r) = r^{k/(m_c-\varepsilon)}$ and $g(r) = h(r)/f_\star(r) = r^{1-k/(m_c-\varepsilon)}$. By Proposition 6.1, there exists $C'_h \in (1,\infty)$ such that $\mathcal{A}(h \circ V_\varepsilon)(x) \leq C'_h$ for all $x \in \mathbb{R}^n$. Thus, Lemma 2.10 (with $H = h \circ V_\varepsilon$ and $\xi \equiv C'_h$) yields $\mathbb{E}_x[h \circ V_\varepsilon(X_t)] \leq C_h(h \circ V_\varepsilon(x) + t)$ for all $x \in \mathbb{R}^n$, $t \in \mathbb{R}_+$ and some $C_h \in (1,\infty)$.

The function $L_{\varepsilon,q}$ in (3) (with $q \in (0,1)$) satisfies $L_{\varepsilon,q}(r) = r\varphi(1/r)\Psi(2r/(1-q))(\log\log r)^\varepsilon \leq r^{1+(\ell-2+2\varepsilon)/(m_c-\varepsilon)}(\log\log r)^\varepsilon/C \leq r^{1+(\ell-2+3\varepsilon)/(m_c-\varepsilon)}/C'$ for some constants $C, C' > 0$ and all $r \in [1,\infty)$. We define the function $a : [1,\infty) \to \mathbb{R}_+$ by

$$\frac{c_{\varepsilon,q}f_\star(g^{-1}(t))}{L_{\varepsilon,q}(g^{-1}(t))} \geq \frac{\widetilde{c}_{\varepsilon,q}t^{k/(m_c-\varepsilon-k)}}{t^{(\ell-2+m_c+2\varepsilon)/(m_c-\varepsilon-k)}} = \widetilde{c}_{\varepsilon,q}t^{-1-(\ell-2+3\varepsilon)/(m_c-k-\varepsilon)}$$

$$= \widetilde{c}_{\varepsilon,q}t^{-1+(2-\ell)/(m_c-k)-b(\varepsilon)} =: a(t),$$

where $b(\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$. Let $A(t) \coloneqq ta(t) = \widetilde{c}_{\varepsilon,q}t^{(2-\ell)/(m_c-k)-b(\varepsilon)}$ for $t \in [1,\infty)$. As $(2-\ell)/(m_c-k) > 0$, for all sufficiently small $\varepsilon > 0$, we have $\lim_{t\to\infty} A(t) = \infty$.

Applying Theorem 2.2 with functions h, $f_\star$, g, a and A defined above yields: for every $x \in \mathbb{R}^n$, there exists a constant $c \in (0,\infty)$ such that

$$\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{f_\star \circ V_\varepsilon} \geq a \circ A^{-1}(2C_h(h \circ V_\varepsilon(x) + t)) \geq c/t^{\eta_\varepsilon},$$

where $\eta_\varepsilon \coloneqq (m_c - k)/((2-\ell) - b(\varepsilon)(m_c-k)) - 1$. Since $b(\varepsilon) \downarrow 0$ as $\varepsilon \downarrow 0$ and $k < m_c$, we get $\eta_\varepsilon \downarrow \alpha_k = m_c/(2-\ell) - 1 - k/(2-\ell)$ as $\varepsilon \downarrow 0$ (recall the definition of α_k in the statement of Theorem 3.1). Moreover, since we have $V_\varepsilon(x) \leq 1 + |x|^{m_c-\varepsilon}$ for all $x \in \mathbb{R}^n$ and $f_\star(r) = r^{k/(m_c-\varepsilon)}$ is concave on $[0,\infty)$, it follows that $f_\star \circ V_\varepsilon(x) \leq f_\star(1+|x|^{m_c-\varepsilon}) \leq f_\star(1) + f_\star(|x|^{m_c-\varepsilon}) = 1 + |x|^k = f_k(x)$ for all $x \in \mathbb{R}^n$. This implies $\| \cdot \|_{f_\star \circ V_\varepsilon} \leq \| \cdot \|_{f_k}$, which concludes the proof. $\qquad\square$

Having given full details of the proofs of Theorems 3.1 and 3.2, in the proofs of the remaining results of Section 3, we will be less explicit in the applications of our theory.

In the last proof of this section, we will show that the assumptions on the upper bounds of the model parameters, used in [9, 12, 36], are not sufficient for determining the rate of convergence. We will prove that, under assumption on upper bounds from [12], one-dimensional Langevin diffusions may achieve polynomial ergodicity of *any* order.

Proof of the inequalities in (14). Fix $\alpha \in (1, \infty)$, pick arbitrary $k \in (\alpha, \infty)$ and set $b \in (0, \infty)$ such that $\alpha = kb/(1 + b)$ holds. Consider π satisfying $\pi(x) = |x|^{-k}((1 + b) + \sin(k|x|)/|x|)$ for $|x|$ sufficiently large and note that $\limsup_{|x| \to \infty} x(\log \pi)'(x) = -kb/(1 + b) = -\alpha$.

For $\eta \in (0, 1]$ define the Lyapunov function $V_\eta : \mathbb{R} \to [1, \infty)$ satisfying $V_\eta(x) = |\int_0^x 1/\pi(y)^\eta \mathrm{d}y|$ for all $|x|$ sufficiently large. Choosing $\eta \in (0, 1)$ implies the equalities $\mathcal{A}V_\eta(x) = ((\log \pi)'V_\eta' + V_\eta'')(x) = (1 - \eta)(\log \pi)'(x)/\pi(x)^\eta$ for all x outside of some compact set, where $\mathcal{A}$ is the generator of the process in (13). Since there exist $c, C \in (0, \infty)$ such that $c/|x|^k \leq \pi(x) \leq C/|x|^k$ for all $x \in \mathbb{R}$ outside a compact set, there also exist constants $c_1, c_2 \in (0, \infty)$ satisfying $V_\eta(x) \geq c_1 x^{1+\eta k}$ and $(\log \pi)'(x)/\pi(x)^\eta \leq c_2 x^{\eta k - 1}$ for all x with large $|x|$. Thus, $\mathcal{A}V_\eta(x) \leq c_3 V_\eta(x)^{(\eta k - 1)/(\eta k + 1)}$ for all $x \in \mathbb{R}$ with $|x|$ sufficiently large and some constant $c_3 \in (0, \infty)$. By [9, Thms 3.2 and 3.4], for each $x \in \mathbb{R}$ there exists $C_0 \in (0, \infty)$ such that $\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq C_0 t^{-(\eta k - 1)/2}$ for all $t \in [1, \infty)$.

To obtain the matching lower bound, note that the inequality $\mathcal{A}V_1 = ((\log \pi)'V_1' + V_1'') \leq C'$ holds on $\mathbb{R}$ for some constant $C' \in (0, \infty)$. Thus, Lemma 2.10 (with $H = V_1$ and $\xi \equiv C'$) yields a constant $C'' \in (1, \infty)$, such that $\mathbb{E}_x[V_1(X_t)] \leq C''(V_1(x) + t)$ for all $x \in \mathbb{R}$ and $t \in \mathbb{R}_+$. Since $\pi(\{V \geq r\}) \geq c_\pi r^{(1-k)/(1+k)}$ for some $c_\pi > 0$ and all large $r \in [1, \infty)$, by Lemma 5.1 with $f \equiv 1$, for every $x \in \mathbb{R}$ there exist $c, C \in (0, \infty)$ such that

$$ct^{(1-k)/2} \leq \|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq Ct^{(1-\eta k)/2} \quad \text{for all } t \in [1, \infty).$$

The proof of (14) is complete since $\eta \in (0, 1)$ was chosen arbitrarily. □

Proof of the exponential ergodicity in (16). In [20, Example 2.5], the authors show that X is polynomially ergodic of *any* order. Let $V : \mathbb{R} \to [1, \infty)$ satisfy $V(x) = \int_0^{|x|} \exp(cy + 2 \sin y)\mathrm{d}y$ for $x \in \mathbb{R}$ with large $|x|$. Since $|\sin y| \leq 1$ for $y \in \mathbb{R}$, we have $0 < \frac{c}{\exp(4)}V(x) \leq \frac{x}{|x|}V'(x)$ for all $x \in \mathbb{R}$ with $|x|$ sufficiently large. As $V''(x) = \frac{x}{|x|}V'(x)(c + 2 \cos x)$ for x in the complement of a compact, we have

$$\mathcal{A}V(x) = -(c + \cos x)\frac{x}{|x|}V'(x) + \frac{1}{2}V''(x) = -\frac{c}{2}\frac{x}{|x|}V'(x) \leq -c^2 V(x)/(2 \exp(4)),$$

where $\mathcal{A}$ is the generator of the diffusion X in (15). Thus $V(X)$ satisfies the drift condition in (10) above with linear $\phi(r) = c^2 r/(2 \exp(4))$, yielding exponential ergodicity by [9, Thm 3.2]. □

6.1.2. *Stretched exponential tails: proofs for Section 3.1.2.* Recall from Assumption A_{se} parameters $p \in (0, 1)$, $\ell \in [0, 2p]$. For any $u \in \mathbb{R}\setminus\{0\}$, consider a function $g_u : \mathbb{R}^n \to (0, \infty)$ in $C^2(\mathbb{R}^n)$, satisfying

$$(41) \qquad g_u(x) = \exp(u|x|^{1-p}) \quad \text{for all } x \in \mathbb{R}^n \text{ with } |x| \text{ sufficiently large.}$$

We may assume that g_u is a function of $|x|$ for all $x \in \mathbb{R}^n$ and, if $u > 0$, then $g_u \geq 1$ on $\mathbb{R}^n$.

The following (deterministic) proposition allows us to construct the functions in $\mathbf{L}(V, \varphi, \Psi)$.

Proposition 6.2. *Under Assumption A_{se}, the extended generator $\mathcal{A}$ in (38) of the diffusion X in (11) satisfies the following.*

(i) For some $u_c \in (0, \infty)$ and all $u \in (u_c, \infty)$, there exists a constant $C_u \in (0, \infty)$ such that

$$0 \leq \mathcal{A}g_u \leq C_u g_u / (\log g_u)^{(2p-\ell)/(1-p)} \qquad \text{outside of some compact set.}$$

(ii) For the function p_1, defined in (39) above, we obtain $\mathcal{A}(1/p_1)(x) \leq C_2/p_1(x)^{2+p-\ell}$ for some constant $C_2 \in (0, \infty)$ and all $x \in \mathbb{R}^n$ with $|x|$ large.

Proof. Denote $g(|x|) = g_u(x)$ for $x \in \mathbb{R}^n$. Note that $\nabla g_u(x) = g'(|x|)x$ and

$$\text{Hess}(g_u)(x) = \text{Hess}(g(|x|)) = g'(|x|)|x|^{-1}I_n + (g''(|x|)/|x|^2 - g'(|x|)/|x|^3)xx^{\mathsf{T}},$$

where $I_n \in \mathbb{R}^{n \times n}$ is the identity matrix. By Assumption A_{se}, $\limsup_{|x| \to \infty} \text{Tr}(\Sigma(x))/|x|^\ell < \infty$ and there exist $0 < \alpha_U < \alpha_L$ and $0 < \beta_L < \beta_U$ such that

$$-\alpha_L \leq \langle b(x), x/|x| \rangle / |x|^{\ell-p} \leq -\alpha_U \quad \text{and} \quad \beta_L \leq \langle \Sigma(x)x/|x|, x/|x| \rangle / |x|^\ell \leq \beta_U$$

for all x outside of some compact set. By the formula for $\mathcal{A}$ in (38), we get

$$(42) \quad (\beta_L u(1-p)/2 - \alpha_L + o(1))/|x|^{2p-\ell} \leq \frac{\mathcal{A}g_u(x)}{u(1-p)g_u(x)} \leq (\beta_U u(1-p)/2 - \alpha_U + o(1))/|x|^{2p-\ell},$$

as $|x| \to \infty$. Thus, for $u \in (2\alpha_L/(\beta_L(1-p)), \infty)$ we have $\mathcal{A}(g_u)(x) \geq 0$ for all x with $|x|$ large. Since $|x| = (u^{-1} \log g_u(x))^{1/(1-p)}$ for $x \in \mathbb{R}^n$ with large $|x|$, there exists a constant $C_u > 0$, such that the upper bound on $\mathcal{A}g_u$ in (i) holds.

For $x \in \mathbb{R}^n$ outside of a compact set, $p_1(x) = |x|$, $\nabla(1/p_1)(x) = -x/|x|^3$ and $\text{Hess}(1/p_1)(x) = -(-3xx^{\mathsf{T}}/|x|^2 + I_n)/|x|^3$. Thus the representation of $\mathcal{A}$ in (38) and the Assumption A_{se} imply

$$\mathcal{A}(1/p_1)(x) \leq (\alpha_L + o(1))/p_1(x)^{2+p-\ell},$$

for all $x \in \mathbb{R}^n$ with sufficiently large $|x|$. This concludes the proof of (ii). $\qquad \square$

Proof of Theorem 3.7. Let $V := p_1$ be a $C^2(\mathbb{R}^n)$ Lyapunov function in (39). By Proposition 6.2(ii) we have $\mathcal{A}(1/V)(x) \leq C/V(x)^{2+p-\ell} = \varphi(1/V(x))$ for some constant $C \in (0, \infty)$ and all $x \in \mathbb{R}^n$ outside of some compact set, where the function $\varphi : (0, 1] \to \mathbb{R}$ is given by $\varphi(1/r) := C/r^{2+p-\ell}$. Moreover, by Proposition 6.2(i), there exists $u_c \in (0, \infty)$, such that for all $u \in (u_c, \infty)$ the function $\Psi_u(r) = \exp(ur^{1-p})$ satisfies $\mathcal{A}(\Psi_u \circ V)(x) = \mathcal{A}g_u \geq 0$ for all $x \in \mathbb{R}^n$ with $|x|$ large (g_u is defined in (41)). The functions (V, φ, Ψ_u) (with any $u > u_c$) defined above, satisfy the **L**-drift condition $\mathbf{L}(V, \varphi, \Psi)$ by Theorem 2.7. For any $u_0 \in (2u, \infty)$, Theorem 2.1 yields

$$(43) \qquad \pi(\{V \geq r\}) \geq c \exp(-u_0 r^{1-p}) \quad \text{for some constant } c \in (0, \infty) \text{ and all } r \in [1, \infty),$$

implying Theorem 3.7(a) (polynomial and logarithmic terms in (4) of Theorem 2.1 are negligible, compared to the stretched exponential decay of $1/\Psi$, implying inequality (43)). Moreover, an application of Theorem 2.5(b) yields the lower bound on the tail of the return time in Theorem 3.7(b).

Pick $u \in (u_0, \infty)$ (for u_0 in (43)) and define $h(r) := \exp(ur^{1-p})$. By the inequality in (43) we have $\pi(\{h \circ V \geq r\}) \geq c/r^{u_0/u}$ for some $c \in (0, 1)$ and all $r \in [1, \infty)$. Since $h \circ V = g_u$ outside of a compact set, by Proposition 6.2(i) there exists a constant $C_h' \in (0, \infty)$, such that

$\mathcal{A}(h \circ V) \leq C_h' h \circ V/(\log h \circ V)^{(2p-\ell)/(1-p)}$ outside of a compact set. Lemma 2.10 (with $H = h \circ V$, $\xi(r) = C_h' r/(\log r)^{(2p-\ell)/(1-p)}$ and hence $\Xi(u) = \int_1^u \mathrm{d}s/\xi(s) = (C_h'^{-1} \log u)^{(1+p-\ell)/(1-p)})$ implies

$$\mathbb{E}_x[h \circ V(X_t)] \leq \Xi^{-1}(\Xi(h \circ V(x)) + t) \leq \exp(C_h(h \circ V(x) + t)^{(1-p)/(1+p-\ell)}) =: v(x, t)$$

for all $x \in \mathbb{R}^n$, $t \in \mathbb{R}_+$ and some $C_h \in (1, \infty)$. Thus, by Corollary 2.3, applied with $a(r) = c/r^{u_0/u}$ and $v(x, t)$, we obtain the claimed lower bound in part (c) of Theorem 3.7. □

6.1.3. Exponential tails: proofs for Section 3.1.3.

Proposition 6.3. *Under Assumption A_e, the extended generator $\mathcal{A}$ in (38) of the diffusion X in (11) satisfies the following.*

(i) For some $u_c \in (0, \infty)$ and all $u \in (u_c, \infty)$, there exists a constant $C_u \in (0, \infty)$ such that

$$0 \leq \mathcal{A}\exp(u|x|) \leq C_u \exp(u|x|) \qquad \text{outside of some compact set.}$$

(ii) For the function p_1, defined in (39) above, we obtain $\mathcal{A}(1/p_1)(x) \leq C'/p_1(x)^2$ for some constant $C' \in (0, \infty)$ and all $x \in \mathbb{R}^n$ with $|x|$ large.

Proof. For any $u \in (0, \infty)$ it holds that $\nabla \exp(u|x|) = u\exp(u|x|)x/|x|$ and $\mathrm{Hess}(\exp(u|x|)) = u\exp(u|x|)(I_n/|x| + (u/|x|^2 + 1/|x|^3)xx^{\mathsf{T}})$. By Assumption A_e, $\limsup_{|x| \to \infty} \mathrm{Tr}(\Sigma(x)) < \infty$ and there exist $0 < \alpha_U < \alpha_L$ and $0 < \beta_L < \beta_U$ such that

$$-\alpha_L \leq \langle b(x), x/|x| \rangle \leq -\alpha_U \quad \text{and} \quad \beta_L \leq \langle \Sigma(x)x/|x|, x/|x| \rangle / \leq \beta_U$$

for all x outside of some compact set. By the formula for $\mathcal{A}$ in (38), we get

$$(\beta_L u/2 - \alpha_L + o(1)) \leq \frac{\mathcal{A}\exp(u|x|)}{u\exp(u|x|)} \leq (\beta_U u/2 - \alpha_U + o(1)) \qquad \text{as } |x| \to \infty.$$

Thus, for $u \in (2\alpha_L/\beta_L, \infty)$ we have $\mathcal{A}(\exp(u|x|)) \geq 0$ for all x with $|x|$ large, implying (i).

For all $x \in \mathbb{R}^n$ with sufficiently large $|x|$, we have $p_1(x) = |x|$, $\nabla(1/p_1)(x) = -|x|^{-3}x$ and $\mathrm{Hess}(1/p_1)(x) = -|x|^{-3}(-3xx^{\mathsf{T}}/|x|^2 + I_n)$. Thus,

$$\mathcal{A}(1/p_1)(x) = (\alpha_L + o(1))/p_1(x)^2 \quad \text{as } |x| \to \infty. \quad \square$$

Proof of Theorem 3.10. Let $V = p_1$, where p_1 is a $C^2(\mathbb{R}^n)$ function in (39). Then, by Proposition 6.3 we have $\mathcal{A}(1/V)(x) = \mathcal{A}(1/p_1)(x) \leq \varphi(1/p_1(x))$, for all $x \in \mathbb{R}^n$ outside of some compact set centered at the origin, where $\varphi(1/r) = C/r^2$ for all $r \in [1, \infty)$ and some constant $C \in (\alpha, \infty)$. Moreover, by Proposition 6.3, there exists $u_c \in (0, \infty)$ such that for any $u_0 \in (u_c, \infty)$ the function $\Psi(r) := \exp(u_0 r)$ satisfies $\mathcal{A}(\Psi \circ p_1) = \mathcal{A}\exp(u_0|x|) \geq 0$ for all $x \in \mathbb{R}^n$ outside of some compact set. Thus, by Theorem 2.7, the conditions of $\mathbf{L(V, \varphi, \Psi)}$ are satisfied with the aforementioned functions (V, φ, Ψ). By Theorem 2.1 we obtain a lower bound on the tail of the invariant measure, implying Theorem 3.10(a). Moreover, an application of Theorem 2.5(b) yields the lower bound on the tail of the return time in Theorem 3.10(b).

The function $L_{\varepsilon,q}$ in (3) (with $q, \varepsilon \in (0, 1)$) satisfies $L_{\varepsilon,q}(r) = r\varphi(1/r)\Psi(2r/(1-q))(\log \log r)^\varepsilon \leq \exp(2u_0 r/(1-q))/C_1$ for some constant $C_1 > 0$ and all $r \in [1, \infty)$. Pick $u \in (2u_0/(1-q), \infty)$ and denote $h(r) := \exp(ur^{1-p})$. Define the function $a : [1, \infty) \to \mathbb{R}_+$ by $a(r) := c_{\varepsilon,q}'c/r^{2u_0/(1-q)/u} \leq c_{\varepsilon,q}/L_{\varepsilon,q}(h^{-1}(r))$. Since $h \circ V = \exp(u|x|)$ outside of a compact set, by Proposition 6.2(i) there

exists a constant $C'_h \in (0, \infty)$, such that $\mathcal{A}(h \circ V) \leq C'_h h \circ V$ outside of a compact set. Lemma 2.10 (with $H = h \circ V$, $\xi(r) = C'_h r$ and hence $\Xi(u) = \int_1^u \mathrm{d}s/\xi(s) = (C'^{-1}_h \log u))$ implies

$$\mathbb{E}_x[h \circ V(X_t)] \leq \Xi^{-1}(\Xi(h \circ V(x)) + t) \leq \exp(C_h(h \circ V(x) + t)) =: v(x, t)$$

for all $x \in \mathbb{R}^n$, $t \in \mathbb{R}_+$ and some $C_h \in (1, \infty)$. Thus, by Corollary 2.3, applied with a and $v(x, t)$, we obtain the claimed lower bound in part (c) of Theorem 3.10. $\qquad\square$

6.2. **Lévy-driven SDE: proofs for Section 3.2.** Let X be the solution of SDE (18) with a bounded Lipschitz dispersion coefficient σ, driven by a pure-jump Lévy process L with Lévy measure ν. By Itô's formula [29, Thm II.36] applied to $g(X)$, where $g \in C^1(\mathbb{R})$ and $\int_{\mathbb{R}\setminus(-1,1)} |g(y)|\nu(\mathrm{d}y) < \infty$, the extended generator (see Section 2.4 above for definition) of the process X takes the form

$$(44) \quad \mathcal{A}g(x) = -\mu x g'(x) + \int_{\mathbb{R}} (g(x + \sigma(x)y) - g(x) - g'(x)\sigma(x)y \mathbb{1}\{y \in [-1, 1]\})\nu(\mathrm{d}y), \qquad x \in \mathbb{R}.$$

For any $m \in [1, \infty)$, consider a non-decreasing function $g_m : \mathbb{R} \to [1, \infty)$ in $C^2(\mathbb{R})$ satisfying

$$(45) \qquad g_m \equiv 1 \text{ on } (-\infty, 1] \quad \& \quad g_m(x) = (\log x)^m \quad \text{for all } x \in \mathbb{R}_+ \text{ sufficiently large.}$$

Note that, in particular, the derivatives g', g'' are globally bounded.

Proposition 6.4. *Under Assumption A_L, extended generator (44) of the process X in (18) satisfies the following inequalities.*

(a) *For any $m \in [1, m_c)$ there exists a constant $C_1 \in (0, \infty)$ such that*

$$\mathcal{A}g_m(x) \leq C_1 \quad \text{for all } x \in \mathbb{R}.$$

(b) *There exist constants $C_2, b, \ell_0 \in (0, \infty)$, such that the following holds*

$$\mathcal{A}(1/g_1)(x) \leq C_2/g_1(x)^2 + b\mathbb{1}\{g_1(x) \leq \ell_0\} \quad \text{for all } x \in \mathbb{R}.$$

Proof. For any $g \in C^2(\mathbb{R})$ and $C'_1 := \sup_{x \in \mathbb{R}} \sigma(x)^2 \int_{[-1,1]} y^2 \nu(\mathrm{d}y) \in [0, \infty)$, Lagrange's theorem yields

$$(46) \quad \int_{[-1,1]} (g(x + \sigma(x)y) - g(x) - g'(x)\sigma(x)y\nu(\mathrm{d}y) \leq C'_1 \sup_{u \in [x-\sigma(x), x+\sigma(x)]} |g''(u)| \quad \text{for all } x \in \mathbb{R}.$$

Part (a). Let $m \in [1, m_c)$. For all large $x \in \mathbb{R}_+$ we have $g'_m(x) = m(\log x)^{m-1}/x$, implying $g'_m(x + u) \leq g'_m(1 + u)$ for all $u > 0$. Thus, by Tonelli's theorem, for all large $x \in \mathbb{R}_+$ we have

$$\int_{[1,\infty)} (g_m(x + \sigma(x)y) - g_m(x))\nu(\mathrm{d}y) = \int_{[1,\infty)} \nu(\mathrm{d}y) \int_0^{\sigma(x)y} g'_m(x + u)\mathrm{d}u$$

$$= \int_0^\infty \nu([\max\{1, u/\sigma(x)\}, \infty))g'_m(x + u)\mathrm{d}u$$

$$(47) \qquad\qquad \leq \int_0^\infty \nu([\max\{1, u/\sup_{y \in \mathbb{R}} \sigma(y)\}, \infty))g'_m(1 + u)\mathrm{d}u < \infty,$$

where the final inequality holds by Assumption A_L since $m < m_c$ and σ is bounded on $\mathbb{R}$.

Since g_m is non-decreasing and $\sup_{u \in \mathbb{R}} |g_m''(u)| < \infty$ (by definition (45)), inequalities (46) and (47) imply there exists $x_0 \in (0, \infty)$, such that, for all $x \in [x_0, \infty)$, the representation of $\mathcal{A}$ in (44) yields

$$(48) \qquad \mathcal{A}g_m(x) \leq -\mu m (\log x)^{m-1} + C_1' \sup_{u \in \mathbb{R}} |g_m''(u)| + \int_{[1,\infty)} (g_m(x + \sigma(x)y) - g_m(x))\nu(\mathrm{d}y)$$

$$\leq C_1' \sup_{u \in \mathbb{R}} |g_m''(u)| + \int_0^\infty \nu([\max\{1, u/\sup_{y \in \mathbb{R}} \sigma(y)\}, \infty)) g_m'(1+u) \mathrm{d}u < \infty.$$

Let $x \in (-\infty, x_0)$. Then $\int_{[1,\infty)} g_m(x + \sigma(x)y)\nu(\mathrm{d}y) \leq \int_{[1,\infty)} g_m(x_0 + \sup_{u \in \mathbb{R}} \sigma(u)y)\nu(\mathrm{d}y) < \infty$ since g_m is non-decreasing and integrable with respect to ν by Assumption $\mathrm{A_L}$. Hence the inequality in (46) and the global boundedness of g_m'' imply

$$\mathcal{A}g_m(x) \leq C_1' \sup_{u \in \mathbb{R}} |g_m''(u)| + \int_{[1,\infty)} g_m(x_0 + \sup_{y \in \mathbb{R}} \sigma(y)u)\nu(\mathrm{d}u) < \infty \quad \text{for all } x \in (-\infty, x_0).$$

Part (b). By definition of g_1 in (45), there exists large $x_0 \in \mathbb{R}_+$, such that

$$\int_{[-\frac{x}{2\sigma(x)}, -1]} ((1/g_1)(x + \sigma(x)y) - (1/g_1)(x))\nu(\mathrm{d}y) = \int_{[-\frac{x}{2\sigma(x)}, -1]} \frac{\log x - \log(x + \sigma(x)y)}{\log(x + \sigma(x)y) \log x} \nu(\mathrm{d}y)$$

$$\leq \nu((-\infty, -1]) \log 2/(\log(x/2) \log(x))$$

$$(49) \qquad\qquad\qquad \leq \nu((-\infty, -1]) \log 4/g_1(x)^2 \qquad \text{for all } x \geq x_0.$$

By enlarging x_0 if necessary, we may assume $-x(1/g_1)'(x) = 1/g_1(x)^2$ and $g_1''(x) \leq 1/g_1(x)^2$ for all $x \geq x_0$. By (49), the bound on small jumps in (46) and the representation of $\mathcal{A}$ in (44), we get

$$\mathcal{A}(1/g_1)(x) \leq C_1' \sup_{u \in [x - \sigma(x), x + \sigma(x)]} |(1/g_1)''(u)| - \mu x(1/g_1)'(x) + \int_{(-\infty, -1]} \frac{g_1(x) - g_1(x + \sigma(x)y)}{g_1(x)g_1(x + \sigma(x)y)} \nu(\mathrm{d}y)$$

$$\leq C_2'/g_1(x)^2 + \nu((-\infty, -\frac{x}{2\sigma(x)})) + \int_{[-\frac{x}{2\sigma(x)}, -1]} ((1/g_1)(x + \sigma(x)y) - (1/g_1)(x))\nu(\mathrm{d}y)$$

$$\leq C_2/g_1(x)^2 \qquad \text{for all } x \geq x_0 \text{ and some constants } C_2', C_2 \in (0, \infty).$$

Jumps in $[1, \infty)$ can be disregarded since $1/g_1$ is non-increasing and the third inequality follows from Assumption $\mathrm{A_L}$ on the negative tail of ν, bound (49) and the fact that σ is bounded.

By (45), $1/g_1$ is non-increasing, $x(1/g_1)'(x)$ is bounded from below for $x \in \mathbb{R}$ and $1/g_1 \leq 1$. Thus, since $\mu \in (0, \infty)$, by (46) we have

$$\mathcal{A}(1/g_1)(x) \leq C_1' \sup_{u \in \mathbb{R}} (1/g_1)''(u) - \inf_{u \in \mathbb{R}} \mu u(1/g_1)'(u) + \nu((-\infty, -1]) =: b < \infty \quad \text{for all } x \in \mathbb{R},$$

implying part (b) of the proposition with $\ell_0 := g_1(x_0)$. $\qquad\qquad\qquad\qquad\qquad\square$

In this section we will work with the Lyapunov function $V := g_1$. Recall the definitions $S_{(\ell)} = \inf\{t \geq 0 : V(X_t) < \ell\}$ and $T^{(r)} = \inf\{t \geq 0 : V(X_t) > r\}$, for $r, \ell \in (1, \infty)$.

Proposition 6.5. *Let Assumption A_L hold. For some $\ell_0 \in (1, \infty)$ and each $\ell \in (\ell_0, \infty)$ there exists $C_\ell \in (0, \infty)$ such that*

$$\mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq C_\ell/r^{m_c} \quad \text{for all } r \in (\ell + 1, \infty) \text{ and } x \in \{\ell + 1 \leq V < r\}.$$

The proof of this proposition is based on a simple idea, which we first explain informally. The process X satisfies SDE (18),

$$(50) \quad X_t = x - \mu \int_0^t X_s \mathrm{d}s + \int_0^t \sigma(X_{s-})\mathrm{d}L_s^{(-)} + \int_0^t \sigma(X_{s-})\mathrm{d}L_s^{(M)} + \int_0^t \sigma(X_{s-})\mathrm{d}L_s^{(+)}, \quad \text{for all } t \in \mathbb{R}_+,$$

where the driving pure-jump Lévy process $L = L^{(-)} + L^{(M)} + L^{(+)}$ is decomposed into a sum of independent pure-jump Lévy processes $L_t^{(-)}$, $L_t^{(M)}$ and $L_t^{(+)}$ with Lévy measures $\nu^{(-)}(\cdot) = \nu(\cdot \cap (-\infty, -1])$, $\nu^{(M)}(\cdot) = \nu(\cdot \cap (-1, 1))$ and $\nu^{(+)}(\cdot) = \nu(\cdot \cap [1, \infty))$, respectively. For any $t \in \mathbb{R}_+$, the process X has no negative jumps in $(-\infty, -1]$ on the event

$$(51) \quad A_t := \{L_t^{(-)} = 0\} \cap \left\{ \sup_{0 \leq s \leq t} \left| \int_0^s \sigma(X_{u-})\mathrm{d}L_u^{(M)} \right| \leq 1/8 \right\}.$$

On A_t, it is thus necessary for X, started at $x \in \{V \geq \ell + 1\}$, to accumulate sufficient negative drift in order to return to $\{V \leq \ell\}$ before time t. Since the drift of the process X is bounded on $\{\ell \leq V \leq \ell + 1\}$, we will prove $\mathbb{P}_x(\{S_{(\ell)} \leq t\} \cap A_t) = 0$ for all sufficiently small $t > 0$, implying

$$\{T^{(r)} < S_{(\ell)}\} \supset \{T^{(r)} < S_{(\ell)}\} \cap A_t \supset \{T^{(r)} < t\} \cap A_t \supset \left\{ \sigma_- \sup_{0 \leq s \leq t} (L_s - L_{s-}) \geq V^{-1}(r) \right\} \cap A_t$$

$\mathbb{P}_x$-a.s., where $\sigma_- := \inf_{u \in \mathbb{R}} \sigma(u) > 0$. Evaluating the probability of the smallest event in the last display will complete the proof.

Proof of Proposition 6.5. Recall that X follows (50) and $V = g_1$, where g_1 is a $C^2(\mathbb{R})$ function in (45). Choose $\ell_0 \in (1, \infty)$ such that $V^{-1}(y) = \exp(y)$ for all $y \in [\ell_0, \infty)$. Note that for all $\ell, r \in [\ell_0, \infty)$, we have $S_{(\ell)} = \inf\{t > 0 : X_t < \exp(\ell)\}$ and $T^{(r)} = \inf\{t > 0 : X_t > \exp(r)\}$. By possibly increasing ℓ_0, we may (and do) assume that $\nu(\exp(\ell_0)/\sigma_-, \infty) < 2\mu \exp(\ell_0)$ holds.

Fix $\ell \geq \ell_0$ and define the upcrossing and downcrossing times of the interval $[\exp(\ell+1), \exp(\ell+2)]$ as follows: $\bar{\theta}_1 := 0$ and for $k \in \mathbb{N}$,

$$\underline{\theta}_k := \inf\{t > \bar{\theta}_k : X_t < \exp(\ell+1)\} \quad \text{and} \quad \bar{\theta}_{k+1} := \{t > \underline{\theta}_k : X_t > \exp(\ell+2)\}.$$

Thus we have $0 = \bar{\theta}_1 \leq \underline{\theta}_1 \leq \cdots \leq \bar{\theta}_k \leq \underline{\theta}_k \leq \bar{\theta}_{k+1} \leq \ldots$. The equality $\bar{\theta}_k = \underline{\theta}_k$ may occur if the process X jumps downwards over the entire interval $[\exp(\ell+1), \exp(\ell+2)]$.

Claim 1. For all $x \in \{V \geq \ell + 1\}$ and $t \in (0, 1/(2\mu \exp(\ell+2)))$ we have $\mathbb{P}_x(\{S_{(\ell)} \leq t\} \cap A_t) = 0$.

Proof of Claim 1. Pick $x \in \{V \geq \ell + 1\}$. Since $\{\bar{\theta}_k \leq S_{(\ell)} < \underline{\theta}_k\} = \emptyset$ for all $k \in \mathbb{N}$, the following holds

$$(52) \quad \{S_{(\ell)} \leq t\} \cap A_t = \cup_{k=1}^\infty \{\underline{\theta}_k \leq S_{(\ell)} \leq t \wedge \bar{\theta}_{k+1}\} \cap A_t \qquad \mathbb{P}_x\text{-a.s.}$$

It is thus sufficient to show that $\mathbb{P}_x(\{\underline{\theta}_k \leq S_{(\ell)} \leq t \wedge \bar{\theta}_{k+1}\} \cap A_t) = 0$ for every $k \in \mathbb{N}$. By (51), on A_t, the negative jumps of X can only come from the Lévy process $L^{(M)}$. The modulus of the negative jumps of $\int_0^{\cdot} \sigma(X_{s-})\mathrm{d}L_s^{(M)}$ on the event A_t is by definition (51) bounded above by $1/4$, implying $X_{\underline{\theta}_k} \geq \exp(\ell+1) - 1/4$. Clearly, for $u \in [\underline{\theta}_k, \bar{\theta}_{k+1})$, we have $X_u \leq \exp(\ell+2)$. Hence, by (50), on the event A_t we obtain

$$\inf_{\underline{\theta}_k \leq s \leq t \wedge \bar{\theta}_{k+1}} X_s = X_{\underline{\theta}_k} + \inf_{s \in [\underline{\theta}_k, t \wedge \bar{\theta}_{k+1}]} \left\{ -\mu \int_{\underline{\theta}_k}^s X_u \mathrm{d}u + \int_{\underline{\theta}_k}^s \sigma(X_{u-})\mathrm{d}L_u^{(M)} + \int_{\underline{\theta}_k}^s \sigma(X_{u-})\mathrm{d}L_u^{(+)} \right\}$$

$$\geq \exp(\ell+1) - 1/4 - \mu\exp(\ell+2)t - 1/4 > \exp(\ell+1) - 1 > \exp(\ell),$$

where the second inequality follows from $t < 1/(2\mu \exp(\ell + 2))$. Hence, on the event A_t, the process X cannot go below $\exp(\ell)$ during the time interval $[\underline{\theta}_k, t \wedge \overline{\theta}_{k+1}]$ for any $k \in \mathbb{N}$, implying $\mathbb{P}_x(\{\underline{\theta}_k \leq S_{(\ell)} \leq t \wedge \overline{\theta}_{k+1}\} \cap A_t) = 0$ for all $k \in \mathbb{N}$. By (52), the claim follows.

To conclude the proof of the proposition, pick $r \in (\ell + 1, \infty)$ and $x \in \{\ell + 1 \leq V < r\}$. Note that by Claim 1 above we have $\{S_{(\ell)} > T^{(r)}\} \supset \{S_{(\ell)} > t\} \cap \{T^{(r)} < t\} \cap A_t = \{T^{(r)} < t\} \cap A_t$ $\mathbb{P}_x$-a.s. for any $t \in (0, 1/(2\mu \exp(\ell + 2)))$. Since the positive jumps of L greater than one can only come from $L^{(+)}$ in (50), we have

$$\{T^{(r)} < t\} \cap A_t \supset \left\{\sup_{0 \leq s \leq t}(L_s - L_{s-}) \geq \frac{V^{-1}(r)}{\sigma_-}\right\} \cap A_t = \left\{\sigma_- \sup_{0 \leq s \leq t}(L_s^{(+)} - L_{s-}^{(+)}) \geq V^{-1}(r)\right\} \cap A_t$$

(recall $V^{-1}(r) = \exp(r) > \exp(\ell_0) > 1$). For some $t, c \in (0, \infty)$, Claim 2 below yields

$$\mathbb{P}_x(T^{(r)} < S_{(\ell)}) \geq \mathbb{P}_x(\{T^{(r)} < t\} \cap A_t) \geq \mathbb{P}_x\left(\left\{\sup_{0 \leq s \leq t}(L_s^{(+)} - L_{s-}^{(+)}) \geq \exp(r)/\sigma_-\right\} \cap A_t\right)$$

$$\geq c\,(1 - \exp(-t\nu(\exp(r)/\sigma_-, \infty))) > c\nu(\exp(r)/\sigma_-, \infty)t/2,$$

where the constant $c \in (0, \infty)$ is such that the third inequality holds uniformly in $r \in (\ell + 1, \infty)$ and $x \in \{V \geq \ell + 1\}$. Recall

$$t\nu(\exp(r)/\sigma_-, \infty)) < t\nu(\exp(\ell_0)/\sigma_-, \infty)) < \nu(\exp(\ell_0)/\sigma_-, \infty))/(2\mu \exp(\ell_0 + 2)) < 1$$

for the last inequality. Proposition 6.5 now follows from the lower bound in Assumption A_L.

Claim 2. There exists $c \in (0, \infty)$ such that

$$\mathbb{P}_x(\{\sup_{0 \leq s \leq t}(L_s^{(+)} - L_{s-}^{(+)}) \geq \exp(r)/\sigma_-\} \cap A_t) > c(1 - \exp(-t\nu(\exp(r)/\sigma_-, \infty)))$$

for all $x \in \{V \geq \ell + 1\}$, $r \in (\ell + 1, \infty)$ and all sufficiently small $t \in (0, \infty)$.

Proof of Claim 2. The processes $L^{(-)}$ and $L^{(+)}$ are independent with $\mathbb{P}_x(L_t^{(-)} = 0) = e^{-t\nu((-\infty, -1])}$ and $\mathbb{P}_x(\sup_{0 \leq s \leq t}(L_s^{(+)} - L_{s-}^{(+)}) \geq \exp(r)/\sigma_-) = 1 - \exp(-t\nu(\exp(r)/\sigma_-, \infty))$ for all $t \in \mathbb{R}_+$. For any $t > 0$, let $\mathcal{B}_t$ be σ-algebra generated by $(L_s^{(+)}, L_s^{(-)})_{s \in [0,t]}$. We have to show that there exists a constant $\widetilde{c} \in (0, \infty)$ such that $\mathbb{P}_x(\sup_{0 \leq s \leq t}|\int_0^s \sigma(X_{u-})dL_u^{(M)}| \leq 1/8|\mathcal{B}_t) > \widetilde{c}$, for all small $t \in (0, \infty)$ and $x \in \{V \geq \ell + 1\}$. By Markov's inequality, it is sufficient to identify a constant $C \in (0, \infty)$, such that $\mathbb{E}_x[\sup_{0 \leq s \leq t}|\int_0^s \sigma(X_{u-})dL_u^{(M)}|^2|\mathcal{B}_t] \leq Ct$ holds for all $t \in (0, \infty)$ and $x \in \{V \geq \ell + 1\}$.

For any $y \in \mathbb{R}$, let the process $(Y_t^y)_{t \in \mathbb{R}_+}$ be the unique strong solution [29, Thm V.6] of the SDE

$$(53) \qquad Y_t^y = y - \mu \int_0^t Y_s^y ds + \int_0^t \sigma(Y_{s-}^y)dL_s^{(M)}, \quad \text{for all } t \in \mathbb{R}_+.$$

For any pair of deterministic sequences $S = (s_k)_{k \in \mathbb{N}}$ and $J = (j_k)_{k \in \mathbb{N}}$, such that $0 < s_k \uparrow \infty$ and $j_k \in \mathbb{R} \setminus \{0\}$, define the process Z recursively as follows:

$$Z_s^{S,V} := Y_s^x \mathbb{1}\{s \in [0, s_1)\} + \sum_{k \in \mathbb{N}} \mathbb{1}\{s \in [s_k, s_{k+1})\}Y_{s-s_k}^{Z_{s_k}^{S,V}+j_k}.$$

Note that the conditional law $(\int_0^s \sigma(X_{u-})dL_u^{(M)})_{s \in [0,t]}$, given the σ-algebra $\mathcal{B}_t$, equals the law of $(\int_0^s \sigma(Z_{u-}^{S,V})dL_u^{(M)})_{s \in [0,t]}$ for the sequences S and V equal to the jump times and sizes of the components of the Lévy process $(L_s^{(+)}, L_s^{(-)})_{s \in [0,t]}$. This is because the Lévy-driven SDE in (53) has a unique strong solution. Since σ is bounded, [29, Thm V.66] yields a constant $C \in (0, \infty)$ such

that for all $x \in \mathbb{R}$ and any pair of deterministic sequences S and J satisfying the above conditions, we have $\mathbb{E}[\sup_{0 \le s \le t} | \int_0^s \sigma(Z_{u-}^{S,V}) dL_u^{(M)}|^2] \le Ct$ for all $t \in \mathbb{R}_+$. This implies that

$$\mathbb{E}_x \left[\sup_{0 \le s \le t} \left| \int_0^s \sigma(X_{u-}) dL_u^{(M)} \right|^2 \middle| \mathcal{B}_t \right] \le Ct \quad \text{for all } t \in \mathbb{R}_+ \text{ and } x \in \mathbb{R},$$

which concludes the proof of Claim 2.   □

Remark 6.6. Under Assumption A_L, the process $\Psi \circ V(X) = (\log X)^{m_c}$ is not a submartingale as its marginals are not integrable. As discussed in Remark 2.9 above, this makes a submartingale argument, similar to the one in the proof of Lemma 2.8, to deduce a lower bound on the probability $\mathbb{P}_x(T^{(r)} < S_{(\ell)})$ infeasible. A slight perturbation $(\log X)^{m_c - \varepsilon}$ (for a small $\varepsilon > 0$) makes the process integrable under Assumption A_L. However, $\Psi(r) = r^{m_c - \varepsilon}$ cannot be used because the process $(\log X)^{m_c - \varepsilon}$ is a supermartingale for all small $\varepsilon > 0$ by (47)–(48) in the proof of Proposition 6.4(a).

Proof of Theorem 3.12. The process X is a Feller process by [19, Thm 1.1], since we can represent it as a solution of the SDE $dX_t = x + \int_0^t \widetilde{\sigma}(X_{s-}) d\widetilde{L}_s$ for the Lévy process $\widetilde{L}_t = (t, L_t)$ and a covariance matrix $\widetilde{\sigma}(x) = (-\mu x, \sigma(x))^{\mathsf{T}}$ (note that the condition on the Lévy measure of [19, Thm 1.1] holds since σ is bounded).

The irreducibility of X follows from the structure of the SDE in (18) since the Lévy driver L has unbounded positive jumps, the function σ is uniformly bounded from below and the drift pushes X linearly towards the origin. Indeed, for any starting point of X, with positive probability a large positive jump of L takes X beyond any given finite interval I the process X needs to spend time in. By the strong Markov property, at this jump time we consider an event (such as A_t in (51) above, which has positive probability for every $t \in (0, \infty)$) without future large negative or positive jumps of L and with a bounded martingale part of L. On such an event, the drift of the process X will take it down to the interval I and, once inside I, will make X stay there for a positive amount of time (recall that X is right-continuous so the martingale part can be controlled). As the details of this argument are tedious and uninformative, they are omitted for brevity. We conclude (via an argument analogous to that in [12, Sec. 3.1]) that X is Harris recurrent and ergodic by [9, Thm 3.2] with a Lyapunov function g_2 defined in (45) above.

Recall that we are working with the Lyapunov function $V = g_1$, where g_1 is a $C^2(\mathbb{R})$ function in (45). By Proposition 6.4(b), there exists a constant $C_2 \in (0, \infty)$ such that the function $\varphi(1/r) := C_2/r^2$, $r \in [1, \infty)$, satisfies the assumptions of Theorem 2.7(a), which in turn implies condition $\mathbf{L(V, \varphi, \Psi)(i)}$. Moreover, by Proposition 6.5 the function $\Psi(r) := r^{m_c}$, $r \in [1, \infty)$, satisfies condition $\mathbf{L(V, \varphi, \Psi)(ii)}$. Note that under $\mathbf{L(V, \varphi, \Psi)}$ the function $L_{\varepsilon, q}$ in (3) (with arbitrary $q, \varepsilon \in (0, 1)$) is bounded above by

$$L_{\varepsilon, q}(r) = r\varphi(1/r)\Psi(2r/(1-q))(\log \log r)^{\varepsilon} \le C_1 r^{m_c - 1}(\log \log r)^{\varepsilon} \quad \text{for all } r \in [1, \infty)$$

and some constant $C_1 \in (0, \infty)$. By Theorem 2.1, for any $\varepsilon > 0$, we obtain the lower bound $\pi(\{V \ge r\}) \ge c_{\varepsilon, q}/r^{m_c - 1 + \varepsilon}$, and hence $\pi([r, \infty)) \ge c'_{\varepsilon, q}/(\log r)^{m_c - 1 + \varepsilon}$, for all $r \in [1, \infty)$ and some constants $c_{\varepsilon, q}, c'_{\varepsilon, q} \in (0, 1)$, implying part (a) of the theorem. Moreover, applying Theorem 2.5(b) yields part (b) of the theorem.

Pick $\varepsilon \in (0, 1/2)$ and consider the function $h(r) := r^{m_c - \varepsilon}$. Proposition 6.4(a) implies the inequality $\mathcal{A}(h \circ V) \le C'_h$ on $\mathbb{R}$ for some $C'_h \in (1, \infty)$, and Lemma 2.10 (with $H = h \circ V$ and $\xi \equiv C'_h$) yields a

constant $C_h \in (1, \infty)$ such that $\mathbb{E}_x[h \circ V(X_t)] \leq C_h(h(V(x)) + t)$ holds for all $x \in \mathbb{R}$ and $t \in \mathbb{R}_+$. Thus we may apply Corollary 2.3 with functions h and $a(r) := C_a r^{(1-m_c-\varepsilon)/(m_c-\varepsilon)} \leq c_{\varepsilon,q}/L_{\varepsilon,q}(h^{-1}(r))$, $r \in [1, \infty)$, and some $C_a \in (0,1)$. For each $x \in \mathbb{R}$ we obtain a constant $c_{\mathrm{TV}} \in (0, \infty)$, such that the lower bound $\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \geq c_{\mathrm{TV}} t^{(1-m_c-\varepsilon)/(1-2\varepsilon)}$ holds for all $t \in [1, \infty)$. $\qquad\square$

6.3. Stochastic damping Hamiltonian system: proofs for Section 3.3.

Let $X = (Z, Y)$ be the hypoelliptic diffusion satisfying the stochastic damping Hamiltonian system in (19). Recall that by Assumption A_H we have $\inf_{z \in \mathbb{R}} U(z) =: -b > -\infty$ and constants $\sigma, c \in (0, \infty)$ in (19) satisfy $ac/\sigma^2 > 1/2$. Following [39], for $\varepsilon \in (0,1)$ and $k \in (0, \infty)$, define a twice differentiable function

$$(54) \qquad g_{k,\varepsilon}(z, y) := (y^2/2 + U(z) + c(1-\varepsilon)(zy + cz^2/2) + b + 1)^k \geq 1 \qquad \text{for all } (z,y) \in \mathbb{R}^2.$$

By Itô's formula applied to $g(Z, Y)$, the extended generator (see Section 2.4 above for definition) of the Hamiltonian system (Z, Y) takes the following form:

$$(55) \qquad \mathcal{A}g(z,y) = \frac{1}{2}\sigma^2 \partial_y^2 g(z,y) + y \partial_z g(z,y) - (cy + U'(z)) \partial_y g(z,y) \qquad \text{for any } g \in C^2(\mathbb{R}^2).$$

We now apply the generator $\mathcal{A}$ to the function $g_{k,\varepsilon}$.

Proposition 6.7. *Under Assumption A_H, extended generator $\mathcal{A}$ of the Hamiltonian system in (19) satisfies the following: if $\varepsilon \in (0, 1/2)$ and $k := 1/2 + ac(1-2\varepsilon)/\sigma^2$, the constant $C := k\varepsilon ac/2 > 0$ satisfies $\mathcal{A}g_{k,\varepsilon}(z,y) \leq -C(g_{k,\varepsilon}(z,y))^{(k-1)/k}$ for all (z,y) outside of some large compact set. Moreover, $\mathcal{A}g_{k,\varepsilon} \leq C'$ on $\mathbb{R}^2$ for some positive constant $C' > 0$.*

Proof. Fix $\varepsilon \in (0, 1/2)$. Applying the generator $\mathcal{A}$ in (55) to the function $g_{1,\varepsilon}$ defined in (54) yields

$$\mathcal{A}g_{1,\varepsilon}(z,y) = \sigma^2/2 - \varepsilon cy^2 - (1-\varepsilon)cU'(z)z.$$

Since $\mathcal{A}g_{k,\varepsilon} = kg_{k-1,\varepsilon}\mathcal{A}g_{1,\varepsilon} + (\sigma^2/2)k(k-1)g_{k-2,\varepsilon}(\partial_y g_{1,\varepsilon})^2$ and $g_{k-1,\varepsilon} = g_{1,\varepsilon}g_{k-2,\varepsilon}$ we obtain

$$\mathcal{A}g_{k,\varepsilon}(z,y) = kg_{k-1,\varepsilon}(z,y)\left(\mathcal{A}g_{1,\varepsilon}(z,y) + \sigma^2(k-1)(\partial_y g_{1,\varepsilon}(x,y))^2/(2g_{1,\varepsilon}(z,y))\right)$$
$$\leq kg_{k-1,\varepsilon}(z,y)(\sigma^2/2 - \varepsilon cy^2 - c(1-\varepsilon)U'(z)z + \sigma^2(k-1))$$
$$= kg_{k-1,\varepsilon}(z,y)((k-1/2)\sigma^2 - \varepsilon cy^2 - c(1-\varepsilon)U'(z)z)$$
$$(56) \qquad \leq -k\varepsilon ac g_{k-1,\varepsilon}(z,y)/2, \qquad \text{as } z^2 + y^2 \to \infty,$$

where the first inequality follows since

$$\frac{(\partial_y g_{1,\varepsilon}(z,y))^2}{2g_{1,\varepsilon}(z,y)} \leq (y + c(1-\varepsilon)z)^2/((y + c(1-\varepsilon)z)^2 + c^2(1-\varepsilon)\varepsilon z^2 + 1) \leq 1 \text{ for all } (z,y) \in \mathbb{R}^2.$$

The inequality in (56) follows from the fact that, for $z^2 + y^2$ sufficiently large, the following holds: either $|z|$ is large and hence by Assumption A_H, $c(1-\varepsilon)U'(z)z \geq ac(1 - 3/2\varepsilon)$ or $|z|$ is small and $|y|$ is large and then $\varepsilon cy^2 \geq (k - 1/2)\sigma^2 + zU'(z) + k\varepsilon ac$, since $zU'(z)$ is bounded for $z \in \mathbb{R}$ by Assumption A_H. Moreover, since $\mathcal{A}g_{k,\varepsilon}$ is continuous, it is also globally bounded from above. $\qquad\square$

Recall that the solution X of (19) is a strong Markov process, all the skeletons are irreducible with respect to the Lebesgue measure and compact sets are petite [39, Lem. 1.1, Prop. 1.2].

Proof of Theorem 3.15. By Assumption A_H, we have $1 < 1/2 + ac/\sigma^2$. Thus, we may choose $\varepsilon \in (0, 1/2)$ such that $k := 1/2 + ac(1 - 2\varepsilon)/\sigma^2 > 1$. Let $V := g_{k,\varepsilon}$ be the function given in (54). By Proposition 6.7 there exist positive constants $C, b > 0$, such that the following inequality $\mathcal{A}V \leq -\phi \circ V + b\mathbb{1}_D$ holds for a sufficiently large compact (hence petite) set D and an increasing function $\phi(r) := Cr^{(k-1)/k}$. Then, by [9, Thm 3.4], the assumptions of [9, Thm 3.2] (i.e. the drift condition in (10)) are satisfied with (V, ϕ, D, b).

Recall that for any $\eta \in [0, 1)$, we have $(r_1/\eta)^\eta (r_2/(1 - \eta))^{1-\eta} \leq r_1 + r_2$ for all $r_1, r_2 \in [1, \infty)$, where for $\eta = 0$ we take $(1/\eta)^\eta = 1$. Note that $g_{\eta(k-1),\varepsilon} = (\phi \circ V)^\eta$. Thus, by [9, Thm 3.2], for every $\eta \in [0, 1)$ and $x \in \mathbb{R}^2$, there exists $C_\eta \in (0, \infty)$ such that $\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{g_{\eta(k-1),\varepsilon}} \leq C_\eta / t^{(k-1)(1-\eta)}$ holds for all $t \in [1, \infty)$. Moreover, by the definition of $g_{k,\varepsilon}$ in (54), it holds that

$$(57) \qquad (\phi \circ V)^\eta(z, y) = g_{\eta(k-1),\varepsilon}(z, y) \geq (1 + c^2(1 - \varepsilon)\varepsilon z^2/2)^{\eta(k-1)}$$

$$\geq \widetilde{C}(1 + |z|^{2(k-1)\eta}) = f_{2(k-1)\eta}(x)$$

for some $\widetilde{C} \in (0, 1)$ and all $x = (z, y) \in \mathbb{R}^2$, where $f_m(z, y) = 1 + |z|^m$ was defined in Theorem 3.15 for any $m \in [0, 2(k - 1))$. Hence the upper bound follows.

As in (57), we have $H(z, y) := C_H(1 + c^2(1 - \varepsilon)\varepsilon z^2/2)^k \leq C_H g_{k,\varepsilon}(z, y)$ for all $(z, y) \in \mathbb{R}^2$. The constant $C_H \in (1, \infty)$ is chosen so that $G_m := H/f_m \geq 1$ on $\mathbb{R}^2$ for any $m \in [0, 2(k - 1))$. By Proposition 6.7 we have $\mathcal{A}g_{k,\varepsilon}(z, y) \leq C'$ on $\mathbb{R}^2$ for some $C' \in (0, \infty)$. By Lemma 2.10 there exists $C'' \in (0, \infty)$, such that $\mathbb{E}_x[H(X_t)] \leq \mathbb{E}_x[C_H g_{k,\varepsilon}(X_t)] \leq C''(g_{k,\varepsilon}(x) + t) =: v(x, t)$ holds for all $x \in \mathbb{R}^2$ and $t \in [1, \infty)$. By [39], $X = (Z, Y)$ admits an invariant measure π with density proportional to $(z, y) \mapsto \exp(-2c/\sigma^2(y^2/2 + U(z)))$. Moreover, by Assumption A_H, we have $U(z) \leq a(1 + \varepsilon) \log |z|$ for all $z \in \mathbb{R}$ outside of some compact set. Thus, $\int_{\mathbb{R}} \pi(z, \mathrm{d}y) \geq c'|z|^{-2ca(1+\varepsilon)/\sigma^2}$ for some constant $c' \in (0, 1)$ and all $|z|$ sufficiently large. We have $\liminf_{|z| \to \infty} G_m(z, y)/|z|^{2k-m} > 0$ since $G_m = H/f_m$. Thus, for all large $r \in \mathbb{R}_+$, there exists $c_\pi > 0$ such that (recall $2 - 2k = 1 - 2ac(1 - 2\varepsilon)/\sigma^2$)

$$\int_{\{G_m \geq r\}} f_m(z, y)\pi(\mathrm{d}z, \mathrm{d}y) \geq \int_{\{|z| \geq r^{1/(2k-m)}\}} c'|z|^{m-2ca(1+\varepsilon)/\sigma^2}\mathrm{d}z \geq c_\pi r^{(m+2-2k-6ca\varepsilon/\sigma^2)/(2k-m)}.$$

Applying Lemma 5.1 with functions f_m, H, G_m, v and $a(r) := c_\pi r^{(m+2-2k-6ca\varepsilon/\sigma^2)/(2k-m)}$, for every $x \in \mathbb{R}^2$ yields a positive constant c_m satisfying

$$\|\mathbb{P}_x(X_t \in \cdot) - \pi(\cdot)\|_{f_m} \geq c_m t^{(m+2-2k-6ca\varepsilon/\sigma^2)/(2(1-3ca\varepsilon/\sigma^2))} \quad \text{for all } t \in [1, \infty).$$

Since $\varepsilon > 0$ is arbitrary, the theorem follows. $\square$

7. CONCLUSION

This paper develops a general theory for establishing lower bounds in f-variation for Markov processes in continuous time. The applications discussed in Section 3 demonstrate the wide applicability of our results. Nevertheless, many questions remain open. We now discuss briefly some interesting possible further directions of research.

Lower bounds for the convergence in Wasserstein metrics. Lyapunov drift conditions have also been developed for establishing upper bounds on the rates of convergence in Wasserstein distance. Applications include convergence of solutions of certain stochastic delay differential equations on infinite-dimensional state spaces [14]. Subgeometric upper bounds in this context have also

received attention [7]. It is feasible that our **L**-drift conditions, suitably adapted to this setting, could yield lower bounds on the return times to sets bounded in an appropriate metric. Due to the lack of local compactness in such state spaces, the notion of petite sets from [26] has been replaced by a weaker notion of metric-dependent small sets, suitable for applications in infinite-dimensional settings. A natural interesting question is whether our lower bounds on modulated moments, derived via suitably adapted **L**-drift conditions, could be used to characterise the decay of the tail of the invariant measure and the rate of convergence in the Wasserstein distance. Put differently, can the lower bound Lyapunov drift conditions, developed in this paper for locally compact state spaces, be adapted to the infinite-dimensional setting?

Ergodic averages. The quantification of the asymptotic behavior of an ergodic average $\frac{1}{t} \int_0^t X_s \mathrm{d}s$, as $t \to \infty$, of a Markov process X represents a fundamental problem in probability and beyond. Results on the asymptotic behaviour of ergodic averages, such as the central limit theorem, as well as moderate and large deviations, are commonly derived using Lyapunov drift conditions and associated upper bounds on the modulated moments [9, 10]. For example, the growth rate of the speed function in moderate deviations of additive functionals of Markov processes is bounded above by the tails of the modulated moments [9, Thm 3.3]. In Theorem 2.5 of the present paper we establish asymptotically matching lower bounds on the modulated moments for a wide range of models (Section 3 above). It is thus reasonable to expect that our lower bounds have a role to play in establishing the optimal growth rate of the speed function in moderate deviations.

Discrete-time Markov chains. To the best of our knowledge, there are no general results for establishing lower bounds on the rates of convergence to the invariant measure for discrete-time Markov chains in uncountable state spaces. It is natural to expect that the results presented in this paper can be readily adapted to establish convergence in the discrete-time setting. The assumption $\mathbf{L}(V, \varphi, \Psi)$ includes a supermartingale condition, exit probability, ergodicity, positive Harris recurrence and non-confinement, all of which have natural counterparts in discrete time. Moreover, since all the tools used in our proofs in Section 5 are also available in discrete time (note that the results of Section 4 hold for discrete-time chains already, if viewed as piecewise constant continuous-time Markov chains), the conclusions of our main results (Theorems 2.1, 2.2 and 2.5) are expected to hold in this setting.

Application to Markov chain Monte Carlo. Markov chain Monte Carlo (MCMC) algorithms represent a major area of application for ergodic Markov processes [11, 26]. In particular, upper bounds on the convergence rates of MCMC algorithms are often obtained using Lyapunov drift conditions. Knowing whether a convergence rate of one algorithm is better than that of another can however not be deduced from the corresponding pair of Lyapunov functions used to obtain the two upper bounds on the rates. Our results on lower bounds for the convergence rate open the door for such comparisons of MCMC algorithms.

APPENDIX A. NON-CONFINEMENT OF IRREDUCIBLE FELLER CONTINUOUS PROCESSES

A strong Markov process $X = (X_t)_{t \in \mathbb{R}_+}$ on a locally compact metric state space $\mathcal{X}$ is *Feller continuous* if the function $x \mapsto \mathbb{E}_x[f(X_t)]$ is continuous for all $t \in (0, \infty)$ and continuous bounded functions $f : \mathcal{X} \to \mathbb{R}_+$.

Lemma A.1. *Let $V : \mathcal{X} \to [1, \infty)$ be continuous and such that $\{V \leq r_0\}$ is compact for all $r_0 \in (1, \infty)$. Assume that the process X is Feller continuous and that for every $r_0 \in (1, \infty)$ there exists $k_0 \in (0, \infty)$ such that $\mathbb{P}_x(V(X_{k_0}) > r_0) > 0$ for all $x \in \mathcal{X}$. Then, we have $\mathbb{P}_x(\limsup_{t \to \infty} V(X_t) = \infty) = 1$.*

In all our examples the processes will be Feller continuous and irreducible with respect to the d-dimensional Lebesgue measure, making the conditions of Lemma A.1 easy to check.

Proof of Lemma A.1. Pick an arbitrary $r_0 \in (1, \infty)$. We will show that $\mathbb{P}_x(\limsup_{t \to \infty} V(X_t) \leq r_0) = 0$ holds for every $x \in \mathcal{X}$. Pick $x \in \mathcal{X}$ and estimate

$$\mathbb{P}_x(\limsup_{t \to \infty} V(X_t) \leq r_0) \leq \mathbb{P}_x(\limsup_{n \to \infty} V(X_{nk_0}) \leq r_0) = \mathbb{P}_x(\cup_{j=1}^{\infty} \cap_{n=j}^{\infty} \{V(X_{k_0 n}) \leq r_0\})$$

$$(58) \qquad\qquad\qquad\qquad \leq \lim_{j \to \infty} \mathbb{P}_x(\cap_{n=j}^{\infty} \{V(X_{k_0 n}) \leq r_0\})$$

where we have used the continuity from below property of the measure $\mathbb{P}_x$. Since $\mathbb{P}_{x'}(V(X_{k_0}) > r_0) > 0$ for every $x' \in \mathcal{X}$, the set $\{V \leq r_0\}$ is compact, and the process X is Feller continuous it follow that $\inf_{x' \in \{V \leq r_0\}} \mathbb{P}_{x'}(V(X_{k_0}) > r_0) =: \varepsilon > 0$. Thus, by recurrent application Markov property at times nk_0 for $j_1, j_2 \in \mathbb{N}$ satisfying $j_2 > j_2$, we have $\mathbb{P}_x(\cap_{n=j_1}^{j_2} \{V(X_{nk_0}) \leq r_0\}) \leq (1 - \varepsilon)^{j_2 - j_1}$. The continuity from above of the measure $\mathbb{P}_x$, implies that for any $j \in \mathbb{N}$

$$\mathbb{P}_x(\cap_{k=j}^{\infty} \{V(X_{km}) \leq n_0\}) = \lim_{j' \to \infty} \mathbb{P}_x(\cap_{k=j}^{j'} \{V(X_{km}) \leq n_0\}) \leq \lim_{j' \to \infty} (1 - \varepsilon)^{j' - j} = 0.$$

Combining the inequality in the above display with (58) concludes the proof. $\qquad\square$

REFERENCES

[1] Ari Arapostathis, Guodong Pang, and Nikola Sandrić, *Ergodicity of a Lévy-driven SDE arising from multiclass many-server queues*, Ann. Appl. Probab. **29** (2019), no. 2, 1070–1126. MR 3910024

[2] Dominique Bakry, Patrick Cattiaux, and Arnaud Guillin, *Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré*, J. Funct. Anal. **254** (2008), no. 3, 727–759. MR 2381160

[3] Miha Brešar and Aleksandar Mijatović, *Subexponential lower bounds for f-ergodic Markov processes*, YouTube presentations: Results and Proofs, 2024, Published on Prob-AM YouTube channel.

[4] Miha Brešar, Aleksandar Mijatović, and Andrew Wade, *Brownian motion with asymptotically normal reflection in unbounded domains: from transience to stability*, to appear in Annals of Probability (2024), 55 pages, arXiv:2303.06916.

[5] Austin Brown, *Some Convergence Results for Metropolis-Hastings Algorithms*, ProQuest LLC, Ann Arbor, MI, 2022, Thesis (Ph.D.)–University of Minnesota. MR 4511936

[6] Austin Brown and Galin L. Jones, *Exact convergence analysis for Metropolis–Hastings independence samplers in Wasserstein distances*, J. Appl. Probab. **61** (2024), no. 1, 33–54. MR 4706676

[7] Oleg Butkovsky, *Subgeometric rates of convergence of Markov processes in the Wasserstein metric*, Ann. Appl. Probab. **24** (2014), no. 2, 526–552. MR 3178490

[8] M. H. A. Davis, *Markov models and optimization*, Monographs on Statistics and Applied Probability, vol. 49, Chapman & Hall, London, 1993. MR 1283589

[9] Randal Douc, Gersende Fort, and Arnaud Guillin, *Subgeometric rates of convergence of f-ergodic strong Markov processes*, Stochastic Process. Appl. **119** (2009), no. 3, 897–923. MR 2499863

[10] Randal Douc, Arnaud Guillin, and Eric Moulines, *Bounds on regeneration times and limit theorems for subgeometric Markov chains*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 2, 239–257. MR 2446322

[11] Randal Douc, Eric Moulines, Pierre Priouret, and Philippe Soulier, *Markov chains*, Springer Series in Operations Research and Financial Engineering, Springer, Cham, 2018. MR 3889011

[12] G. Fort and G. O. Roberts, *Subgeometric ergodicity of strong Markov processes*, Ann. Appl. Probab. **15** (2005), no. 2, 1565–1589. MR 2134115

[13] M. Hairer, *Lecture notes on convergence of Markov processes*, May 2021, available at `https://www.hairer.org/notes/Convergence.pdf`.

[14] M. Hairer, J. C. Mattingly, and M. Scheutzow, *Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations*, Probab. Theory Related Fields **149** (2011), no. 1-2, 223–259. MR 2773030

[15] Martin Hairer, *How hot can a heat bath get?*, Comm. Math. Phys. **292** (2009), no. 1, 131–177. MR 2540073

[16] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR 1943877

[17] R. Z. Khasminskiĭ, *Stochastic stability of differential equations*, Monographs and Textbooks on Mechanics of Solids and Fluids, Mechanics and Analysis, vol. 7, Sijthoff & Noordhoff, Alphen aan den Rijn—Germantown, Md., 1980, Translated from the Russian by D. Louvish. MR 600653

[18] S. A. Klokov, *On lower bounds for the mixing rate for a class of Markov processes*, Teor. Veroyatn. Primen. **51** (2006), no. 3, 600–607. MR 2325548

[19] Franziska Kühn, *Solutions of Lévy-driven SDEs with unbounded coefficients as Feller processes*, Proc. Amer. Math. Soc. **146** (2018), no. 8, 3591–3604. MR 3803683

[20] Petra Lazić and Nikola Sandrić, *On sub-geometric ergodicity of diffusion processes*, Bernoulli **27** (2021), no. 1, 348–380. MR 4177373

[21] J. C. Mattingly, A. M. Stuart, and D. J. Higham, *Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl. **101** (2002), no. 2, 185–232. MR 1931266

[22] Ian Melbourne and Dalia Terhesiu, *Operator renewal theory and mixing rates for dynamical systems with infinite measure*, Invent. Math. **189** (2012), no. 1, 61–110. MR 2929083

[23] M. Menshikov and R. J. Williams, *Passage-time moments for continuous non-negative stochastic processes and applications*, Adv. in Appl. Probab. **28** (1996), no. 3, 747–762. MR 1404308

[24] M. V. Menshikov and S. Yu. Popov, *Exact power estimates for countable Markov chains*, Markov Process. Related Fields **1** (1995), no. 1, 57–78. MR 1403077

[25] S. P. Meyn and R. L. Tweedie, *Generalized resolvents and Harris recurrence of Markov processes*, Doeblin and modern probability (Blaubeuren, 1991), Contemp. Math., vol. 149, Amer. Math. Soc., Providence, RI, 1993, pp. 227–250. MR 1229967

[26] ———, *Markov chains and stochastic stability*, Communications and Control Engineering Series, Springer-Verlag London, Ltd., London, 1993. MR 1287609

[27] Sean P. Meyn and R. L. Tweedie, *Stability of Markovian processes. II. Continuous-time processes and sampled chains*, Adv. in Appl. Probab. **25** (1993), no. 3, 487–517. MR 1234294

[28] ———, *Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes*, Adv. in Appl. Probab. **25** (1993), no. 3, 518–548. MR 1234295

[29] Philip E. Protter, *Stochastic integration and differential equations*, Stochastic Modelling and Applied Probability, vol. 21, Springer-Verlag, Berlin, 2005, Second edition. Version 2.1, Corrected third printing. MR 2273672

[30] Michael Röckner and Feng-Yu Wang, *Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups*, J. Funct. Anal. **185** (2001), no. 2, 564–603. MR 1856277

[31] B. A. Rogozin, *An estimate of the remainder term in limit theorems of renewal theory*, Teor. Verojatnost. i Primenen. **18** (1973), 703–717. MR 334339

[32] Nikola Sandrić, Ari Arapostathis, and Guodong Pang, *Subexponential upper and lower bounds in Wasserstein distance for Markov processes*, Appl. Math. Optim. **85** (2022), no. 3, Paper No. 24, 45. MR 4429313

[33] Omri Sarig, *Subexponential decay of correlations*, Invent. Math. **150** (2002), no. 3, 629–653. MR 1946554

[34] Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 2013, Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. MR 3185174

[35] Pekka Tuominen and Richard L. Tweedie, *Subgeometric rates of convergence of f-ergodic Markov chains*, Adv. in Appl. Probab. **26** (1994), no. 3, 775–798. MR 1285459

[36] A. Yu. Veretennikov, *On polynomial mixing bounds for stochastic differential equations*, Stochastic Process. Appl. **70** (1997), no. 1, 115–127. MR 1472961

[37] ⎯⎯⎯, *On lower bounds for mixing coefficients of Markov diffusions*, From stochastic calculus to mathematical finance, Springer, Berlin, 2006, pp. 623–633. MR 2234294

[38] W. M. Wonham, *A Liapunov method for the estimation of statistical averages*, J. Differential Equations **2** (1966), 365–377. MR 206438

[39] Liming Wu, *Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems*, Stochastic Process. Appl. **91** (2001), no. 2, 205–238. MR 1807683