

SUBGEOMETRIC ERGODICITY OF STRONG MARKOV PROCESSES

G. FORT AND G.O. ROBERTS

SUMMARY

We derive sufficient conditions for subgeometric f -ergodicity of strongly Markovian processes. We first propose a criterion based on modulated moment of some delayed return-time to a petite set. We then formulate a criterion for polynomial f -ergodicity in terms of a drift condition on the generator. Applications to specific processes are considered, including Langevin tempered diffusions on \mathbb{R}^n and storage models.

2000 Mathematics Subject Classifications *Primary:* 60J25, *Secondary:* 60J60, 60K30.

Keywords: Markov processes, Subgeometric f -ergodicity, Drift criterion, Langevin diffusions, Storage models.

Affiliations: G. Fort (Corresponding author), CNRS/LMC-IMAG, 51, rue des Mathématiques, BP 53, 38041 Grenoble Cedex 9, France, *email:* Gersende.Fort@imag.fr

G.O. Roberts, Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, United Kingdom, *email:* g.o.roberts@lancaster.ac.uk

1. INTRODUCTION

This paper is devoted to the study of sub-geometric f -ergodicity of a strong Markov semi-group $(P^t)_{t \geq 0}$. That is, for a sub-geometrically increasing rate function $r := (r(t))_{t \geq 0}$, and a Borel function $f \geq 1$, we propose sufficient conditions implying the limit

$$\lim_{t \rightarrow +\infty} r(t) \|P^t(x, \cdot) - \pi(\cdot)\|_f = 0,$$

for π -almost all (a.a.) x where π is the unique invariant probability measure. Our main condition is couched in terms of modulated moments of return-times to a “test-set”. In this form, this condition extends earlier criteria implying different notions of stability (such as Harris-recurrence, positive Harris-recurrence, ergodicity, f -ergodicity), for continuous-time Markov processes. This condition is also analogous to the criterion for sub-geometric f -ergodicity of discrete time Markov chains. We also derive a condition for polynomial ergodicity which is easy to check in many applications. This condition is expressed in terms of inequality on the semi-group *generator*, and is analogous to the so-called *drift inequality* in the discrete-time case.

We apply our results to the study of strongly Markovian processes, giving three non-trivial examples, two of which are of considerable applied probabilistic interest. We first consider a simple jump process as a toy example, demonstrating that f -ergodicity at a logarithmic (*resp.* polynomial or sub-exponential) rate is narrowly related to the existence of a logarithmic (*resp.* polynomial or sub-exponential) moment of the mean-time spent in each state, with respect to the jump distribution. We then consider Langevin tempered diffusions on \mathbb{R}^n which are relevant to Markov chain Monte Carlo (MCMC) techniques since they construct a diffusion process with given stationary distribution π (which only needs to be available up to an unknown normalization constant). When the stationary distribution is polynomial in the tails, the (simple) Langevin diffusion can not be ergodic at a geometric rate and we show that it is polynomially ergodic. We also consider Langevin tempered diffusion in which the diffusion matrix is a scalar matrix with coefficient π^{-2d} , $d > 0$, and prove that even when the target distribution is polynomial in the tails, a convenient choice of the temperature d involves geometric ergodicity of the process. Finally, we study a compound Poisson-process driven Ornstein-Uhlenbeck process which is used in storage models and more recently in financial econometrics. It is known that when the distribution of the jump F has sufficiently light tails, the process is geometrically ergodic. We investigate the case where F is heavy tailed and establish the sub-geometric ergodicity of the process under appropriate conditions in this case.

The paper is organised as follows. We first recall basic definitions on Markov process, as well as reviewing existing results on ergodicity of strongly Markovian process. The new criteria for sub-geometric ergodicity are given in Section 2, and the proofs are postponed in Section 4. Section 3 is devoted to the three examples mentioned above.

1.1. Basic definitions on Markov process. Let \mathcal{X} be a locally compact and separable metric space endowed with the Borel σ -field $\mathcal{B}(\mathcal{X})$. $X = (\Omega, \mathcal{A}, (\mathcal{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}, \mathbb{P}_x)$, is a \mathcal{X} -valued Borel right process so that it is a temporally homogeneous Markov process, strongly Markovian with right-continuous sample paths (see e.g. Sharpe (1988)). \mathbb{P}_x (resp. \mathbb{E}_x) denotes the canonical probability (resp. expectation) associated to the Markov process with initial distribution δ_x , the Dirac distribution at point x . Let $(P^t)_{t \geq 0}$ be the associated Markov semi-group.

We recall basic definitions and properties on Markov process that will be used throughout this paper. The process X is ϕ -irreducible for some σ -finite measure ϕ on $\mathcal{B}(\mathcal{X})$ if

$$\phi(A) > 0 \implies \mathbb{E}_x \left[\int_0^\infty \mathbb{1}_A(X_s) ds \right] > 0, \quad \forall x \in \mathcal{X}.$$

If the process is ϕ -irreducible, there exists a maximal irreducibility measure ψ that dominates any irreducibility measure (Nummelin (1984)). In fact, if π is an invariant measure i.e. $\pi P^t = \pi$ for all $t \geq 0$, then π is a maximal irreducibility measure. Any measurable set which is of positive ψ -measure is said to be *accessible*. A set $C \in \mathcal{B}(\mathcal{X})$ is ν_b -petite for the process (or simply petite) if there exist a probability measure b (resp. non-trivial σ -finite measure ν_b) on the Borel σ -field of \mathbb{R}_+ (resp. on $\mathcal{B}(\mathcal{X})$) such that

$$\int_0^\infty P^t(x, \cdot) b(dt) \geq \nu_b(\cdot), \quad \text{for all } x \in C.$$

A ϕ -irreducible process always possesses an accessible closed petite set (Meyn and Tweedie, 1993a, Proposition 3.2.). A process is *Harris-recurrent* if there exists a σ -finite measure ϕ such that

$$\phi(A) > 0 \implies \mathbb{P}_x \left(\int_0^\infty \mathbb{1}_A(X_s) ds = +\infty \right) = 1, \quad x \in \mathcal{X};$$

or, equivalently, if there exists a σ -finite measure μ such that $\mu(A) > 0 \implies \mathbb{P}_x(\tau_A < \infty) = 1$ for all $x \in \mathcal{X}$. Harris-recurrence trivially implies ϕ -irreducibility. A Harris recurrent right process possesses an invariant measure (Gettoor (1980)). In fact, when the invariant measure is finite, X is called *positive Harris-recurrent*. A ϕ -irreducible process is *aperiodic* if there exist an accessible ν_{δ_m} -petite set C and t_0 such that $P^t(x, C) > 0$ for all $x \in C, t \geq t_0$. (Meyn and Tweedie, 1993c, Proposition 6.1) shows that a positive Harris recurrent process is aperiodic if some skeleton

chain P^m , $m > 0$, is irreducible i.e. if there exists a σ -finite measure ϕ on $\mathcal{B}(\mathcal{X})$ such that $\phi(A) > 0 \implies \mathbb{E}_x \left[\sum_{n \geq 0} \mathbb{1}_A(X_{nm}) \right] > 0$ for all $x \in \mathcal{X}$.

For Borel functions $f \geq 1$, g , define the norm $|g|_f := \sup_x |g(x)|/f(x)$ and the Banach space $\mathcal{L}_f := \{g, |g|_f < \infty\}$. For a signed measure μ , the *total variation norm* is given by $\|\mu\|_{\text{TV}} := \sup_A \mu(A) - \inf_A \mu(A)$; and the f -norm (for some Borel function $f \geq 1$), $\|\mu\|_f := \sup_{\{g, |g|_f=1\}} |\mu(g)|$, so that the total variation norm is the $\mathbb{1}$ -norm where $\mathbb{1}$ denotes the constant function $\mathbb{1}(t) = 1$. The process is *ergodic* if

$$\forall x \in \mathcal{X}, \quad \lim_{t \rightarrow \infty} \|P^t(x, \cdot) - \pi(\cdot)\|_{\text{TV}} = 0,$$

and f -ergodic if $\pi(f) < \infty$ and

$$\forall x \in \mathcal{X}, \quad \lim_{t \rightarrow \infty} \|P^t(x, \cdot) - \pi(\cdot)\|_f = 0. \quad (1)$$

Finally, X is *geometrically* (resp. *sub-geometrically*) f -ergodic if the limit (1) holds at a rate $r(t) := \kappa^t$, for some $\kappa > 1$ (resp. $r := (r(t))_{t \geq 0}$ for some sub-geometrically increasing rate). A sub-geometric rate is defined as follows (see e.g. Thorisson (1985)). Let Λ_0 be the set of the measurable, bounded on bounded intervals and non-decreasing functions $r : \mathbb{R}_+ \rightarrow [1, \infty)$, such that $\log r(t)/t \downarrow 0$ as $t \rightarrow +\infty$. Let Λ be the set of the rates $\bar{r} := (\bar{r}(t))_{t \geq 0}$ such that for some $r \in \Lambda_0$, $0 < \liminf_t \bar{r}(t)/r(t) \leq \limsup_t \bar{r}(t)/r(t) < \infty$. Λ is by definition, the set of the sub-geometric rates. For example, Λ contains rates such as $\bar{r}(t) \sim \log^\beta(t+1)$, $\beta \geq 0$, $\bar{r}(t) \sim (1 \vee t^\alpha) \log^\beta(t+1)$, $\alpha > 0$, $\beta \in \mathbb{R}$, and sub-exponential rates $\bar{r}(t) \sim \exp(\alpha t^\beta)$, $\alpha > 0$, $0 < \beta < 1$.

Throughout this paper, we will often make comparison with (discrete time) Markov chains; the unfamiliar reader can refer to Meyn and Tweedie (1993b).

1.2. (f, r) -modulated moments and stability. Define the hitting-time on a measurable set C , delayed by $\delta > 0$,

$$\tau_C(\delta) := \inf\{t \geq \delta, X_t \in C\}, \quad (2)$$

the moment $\tau_C(0)$ is denoted by τ_C . It is proved in the literature that modulated moments of $\tau_C(\delta)$ for some closed petite set are related to Harris-recurrence, positive Harris-recurrence, f -ergodicity and geometric f -ergodicity. For a Borel function $f \geq 1$, an increasing non-negative

rate function $r = (r(t))_{t \geq 0}$, $\delta > 0$, define the (f, r) -modulated moment

$$G_C(x, f, r; \delta) := \mathbb{E}_x \left[\int_0^{\tau_C(\delta)} r(s) f(X_s) ds \right].$$

- R1** X is Harris-recurrent if and only if there exists a petite set C such that for all $x \in \mathcal{X}$ $\mathbb{P}_x(\tau_C < \infty) = 1$ (Meyn and Tweedie, 1993a, Theorem 1.1).
- R2** If X is Harris-recurrent with invariant measure π , then for $f \geq 1$, $\pi(f) < \infty$ if and only if $\sup_{x \in C} G_C(x, f, \mathbb{1}; \delta) < \infty$ for some closed petite set C (Meyn and Tweedie, 1993a, Theorem 1.2.).
- R3** A positive Harris-recurrent process is ergodic if and only if some skeleton chain P^m is irreducible (Meyn and Tweedie, 1993c, Theorem 6.1.).
- R4** A positive Harris-recurrent process is f -ergodic if (a) some skeleton chain P^m is irreducible, (b) $\sup_{x \in C} G_C(x, h, \mathbb{1}; \delta) < \infty$ where $h \geq \sup_{s \leq m} P^s f$ and C is a closed petite set, and (c) for all x , $G_C(x, h, \mathbb{1}; \delta) < \infty$ (Meyn and Tweedie, 1993a, Proposition 4.1.) and (Meyn and Tweedie, 1993c, Theorem 7.2.).
- R5** A positive Harris-recurrent process is geometrically f -ergodic if (a) some skeleton chain P^m is irreducible, (b) there exists a closed petite set C and $\eta > 0$ and $G_C(x, h, \exp(\eta t); \delta)$ is finite for all x , where $h \geq 1$ is a Borel function such that $c_1 f \leq \int_0^\infty \exp(-t) P^t h dt \leq c_2 f$ for some finite positive constants c_i , (c) $\sup_C G_C(\cdot, h, \exp(\eta t); \delta) < \infty$ (Down et al., 1995, Theorem 7.4.).

In Section 2 we give a criterion of the form R1-5 that implies sub-geometric f -ergodicity.

To date, little is known about general characterisations for f -ergodicity at a sub-geometric rate for Markov processes. However we note some important special cases which have been studied in the literature. The work by Ganidis et al. (1999) is restricted (a) to convergence in total variation norm and (b) to diffusion processes on \mathbb{R}^d with diffusion matrix equals to identity. Their proof is based on spectral properties of the transition semi-group seen as an operator, and differs from the probabilistic approach adopted in the present paper. We will see in Section 3.2 how to improve their conclusions. Veretennikov (1999) and Malyshev (2001) deal with diffusion processes and can be read as a special application of the present paper. The most related work to the present one is the paper by Dai and Meyn (1995) that considers f -ergodicity at a polynomial rate of a Markovian state process, in order to study the stability of multi-class

queuing networks. These results are particularly related to our work and we will describe their results in Section 2.

1.3. Drift condition and generator. For a Borel function $0 \leq V < \infty$, denote by $\mathcal{A}V$ the Borel function - when exists - such that $t \mapsto \mathcal{A}V(X_t)$ is integrable \mathbb{P}_x -almost surely (a.s.), and there exists an increasing sequence of stopping-time $\{T_n\}_n$ such that for any stopping time τ

$$\mathbb{E}_x \left[V(X_{\tau \wedge T_n}) - V(X_0) - \int_0^{\tau \wedge T_n} \mathcal{A}V(X_s) ds \right] = 0, \quad \text{for all } x \in \mathcal{X}, n \geq 0.$$

When $\mathcal{A}V$ exists, V is said to be in the domain of \mathcal{A} . If there exists h such that $t \mapsto h(X_t)$ is integrable \mathbb{P}_x -a.s. and $t \mapsto V(X_t) - V(X_0) - \int_0^t h(X_s) ds$ is a right continuous \mathbb{P}_x -local martingale (with respect to the filtration \mathcal{F}_t), then V is in the domain of \mathcal{A} and $\mathcal{A}V = h$ (Davis (1993)). If V is in the domain of the weak infinitesimal generator $\tilde{\mathcal{A}}$ then V is in the domain of \mathcal{A} and $\mathcal{A}V = \tilde{\mathcal{A}}V$ (Dynkin (1965)). If the functions V and $\mathcal{A}V$ are right-continuous, these two sufficient conditions are equivalent and $\tilde{\mathcal{A}}V = \mathcal{A}V$.

When $\mathcal{A}V$ satisfies a drift condition $\mathcal{A}V \leq -f + b\mathbb{1}_C$ for some closed set C , and a non-negative function f such that $t \mapsto f(X_t)$ is integrable \mathbb{P}_x -a.s., we have $G_C(x, f, \mathbb{1}; \delta) \leq V(x) + \delta b$; this will be the basic tool to upper bound the (f, r) -modulated moments.

Conditions on $\mathcal{A}V$ are analogous to conditions on the variation $P^m V - V$ for a discrete time Markov chain with transition kernel P^m . It is well-known that the condition $P^m V - V \leq -f$ outside a “test set” for the skeleton P^m is related (a) to the f -ergodicity of the Markov chain $(X_{km})_k$ (Meyn and Tweedie (1993b)); (b) to the geometric V -ergodicity if $f = \lambda V$ for some $0 < \lambda < 1$ (Meyn and Tweedie (1993b)); (c) to the polynomial V -ergodicity if $f \propto V^{1-\alpha}$ for some $0 < \alpha \leq 1$ (Jarner and Roberts (2002); Fort and Moulines (2003)); (d) and more generally sub-geometric f -ergodicity is $f \propto \phi(V)$ for some concave function ϕ (Douc et al. (2004)). Similar results hold for continuous Markov process. Meyn and Tweedie (1993d) prove that the condition $\mathcal{A}V \leq -f$ outside a closed petite set is related (a) to the f -ergodicity of the Markov process X ; and (b) to the geometric V -ergodicity if $f \propto V$ (see also Down et al. (1995); Roberts and Rosenthal (1996); Roberts and Tweedie (2000)). In Section 2, we establish that the case $f \propto V^{1-\alpha}$ is related to polynomial f -ergodicity.

2. STATEMENTS OF THE RESULTS

In Theorem 1, we establish that modulated moment on some delayed hitting-time on a closed petite set C provides a criterion for sub-geometric f -ergodicity. We assume that there exist $\delta > 0$, a Borel function $f_* \geq 1$ and a rate function $r_* \in \Lambda$ such that

$$\sup_C G_C(\cdot, f_*, \mathbb{I}; \delta) < \infty, \quad \sup_C G_C(\cdot, \mathbb{I}, r_*; \delta) < \infty. \quad (3)$$

We will establish that r_* is the maximal rate of convergence (that can be deduced from these assumptions) and it is associated to convergence in total variation norm *i.e.* in \mathbb{I} -norm, which is the minimal one. On the other hand, we will show that f_* is the largest norm in which convergence occurs and the associated convergence rate is the minimal one \mathbb{I} .

Using an interpolation technique, we also derive a convergence rate $1 \leq r_f \leq r_*$ in Λ associated to some f -norm, $1 \leq f \leq f_*$ (see Douc et al. (2004) for a similar approach in the discrete time case). The simplest interpolation technique is given by Hölder's inequality which yields (from (3)) $\sup_C G_C(\cdot, f_*^p, r_*^{1-p}; \delta) < \infty$. By analogy to the discrete-time case, one would expect convergence in f_*^p -norm at the rate r_*^{1-p} , and we will prove the continuous time version of this result.

More generally, if there exists a pair of non-decreasing positive functions (Ψ_1, Ψ_2) satisfying

$$\Psi_1(x) \Psi_2(y) \leq x + y, \quad x, y \geq 1, \quad (4)$$

then $\sup_C G_C(\cdot, \Psi_2(f_*), \Psi_1(r_*); \delta) < \infty$. We will establish that if $\Psi_1(r_*) \in \Lambda$, this condition yields convergence in $\Psi_2(f_*)$ -norm at the rate $\Psi_1(r_*)$. Young functions are closely related to these pairs of functions (Ψ_1, Ψ_2) , say. Specifically, if (H_1, H_2) is a pair of Young function, then (H_1^{-1}, H_2^{-1}) satisfies (4) (see e.g. (Krasnosel'skii and Rutickii, 1961, Chapter 1)). Let \mathcal{I} be the set of pairs of inverse Young functions, augmented with the pairs (Id, \mathbb{I}) and (\mathbb{I}, Id) . As commented above, \mathcal{I} contains the pairs $((x/p)^p, (y/(1-p))^{1-p})$, $0 < p < 1$, and more generally, the pairs of functions increasing at infinity as $(x^p \ln^b x, y^{1-p} \ln^{-b} y)$ for some $0 < p < 1$ and $b \in \mathbb{R}$, $p = 0$ and $b \geq 0$, $p = 1$ and $b \leq 0$.

Theorem 1. *Let $f_* \geq 1$ be a Borel function and $r_* \in \Lambda$. Assume that*

- (i) *X is Harris-recurrent with invariant measure π , and some skeleton chain, say P^m , is ψ -irreducible.*
- (ii) *there exist a closed petite set C and some $\delta > 0$, such that (3) holds.*

(iii) there exists a finite constant c such that $\sup_{t \leq m} P^t f_* \leq c f_*$.

Then π is an invariant probability measure, $\pi(f_*) < \infty$ and for any pair $\Psi := (\Psi_1, \Psi_2) \in \mathcal{I}$

$$\lim_{t \rightarrow +\infty} \{\Psi_1(r_*(t)) \vee 1\} \|P^t(x, \cdot) - \pi(\cdot)\|_{\Psi_2(f_*) \vee 1} = 0 \quad \text{for all } x \in \mathcal{S}_\Psi,$$

where \mathcal{S}_Ψ , which is of π -measure one, is defined by

$$\mathcal{S}_\Psi := \left\{ x \in \mathcal{X}, \mathbb{E}_x \left[\int_0^{\tau_C(\delta)} \Psi_1(r_*(s)) \Psi_2(f_*(X_s)) ds \right] < \infty \right\}.$$

The proof of Theorem 1 is postponed in Section 4.1. We first verify that $\Psi_1(r_*(t)) \vee 1 \in \Lambda$. Under (i-ii), C is accessible and the following lemma holds (Meyn and Tweedie, 1993c, Proposition 6.1.).

Lemma 2. *Suppose that X is positive Harris recurrent with invariant distribution π and some skeleton chain P^m is irreducible. For any accessible petite set C , there exist $t_0 > 0$ and an irreducibility measure ν for the process such that $\nu(C) > 0$ and $\inf_{x \in C} \inf_{t \geq t_0} P^t(x, \cdot) \geq \nu(\cdot)$.*

Based on this lemma and on condition (ii), the second step consists in proving that the skeleton P^m is irreducible, aperiodic and possesses a petite set A such that

$$\sup_{x \in A} G_A^{(m)}(x, \Psi_2(f_*), \Psi_1(r_*)) < \infty, \quad \text{with } G_A^{(m)}(x, f, r) = \mathbb{E}_x \left[\sum_{k=0}^{T_{m,A}} r(k) f(X_{km}) \right], \quad (5)$$

where $T_{m,A} \geq 1$ is the return-time to A for the skeleton chain P^m

$$T_{m,A} := \inf\{k \geq 1, X_{km} \in A\}. \quad (6)$$

By application of Tuominen and Tweedie (1994), Theorem 2.1, this proves that for π -a.a. x , $\lim_{k \rightarrow +\infty} \Psi_1(r_*(k)) \|P^{km}(x, \cdot) - \pi(\cdot)\|_{\Psi_2(f_*)} = 0$. Using (iii), the limit still holds replacing $r_*(k)$ (resp. P^{km}) by $r_*(t)$ (resp. P^t). We finally establish that the limit holds for all the points $x \in \mathcal{S}_\Psi$ and $\pi(\mathcal{S}_\Psi) = 1$.

Remark 3. Theorem 1 remains valid by substituting condition (i) for the condition : there exist a ψ -irreducible, aperiodic and positive recurrent transition kernel P^m .

Theorem 1 remains valid by substituting (ii-iii) for the condition : there exist a closed petite set C and some $\delta > 0$ such that $\sup_C G_C(\cdot, h, \mathbb{I}; \delta) < \infty$ and $\sup_C G_C(\cdot, \mathbb{I}, r_*; \delta) < \infty$ where $h \geq \sup_{t \leq m} P^t f_*$.

Condition (iii) implies that the semi-group $(P^t)_{t \geq 0}$ and the resolvent kernel $R = \int_0^\infty \exp(-t) P^t dt$ are bounded on \mathcal{L}_{f_*} .

Remark 4. By (4), it is readily seen that $\{x; G_C(x, f_*, \mathbb{I}; \delta) + G_C(x, \mathbb{I}, r_*; \delta) < \infty\} \subset \mathcal{S}_\Psi$. It may be read from the proof that

$$\lim_{t \rightarrow +\infty} \{\Psi_1(r_*(t)) \vee 1\} \|\mu P^t(\cdot) - \pi(\cdot)\|_{\Psi_2(f_*) \vee 1} = 0, \quad (7)$$

for all probability measure μ such that $G_C(x, \Psi_2(f_*) \vee 1, \Psi_1(r_*) \vee 1; \delta)$ is μ -integrable. Applying again (4), (7) holds for all distribution μ such that $\{G_C(x, f_*, \mathbb{I}; \delta) + G_C(x, \mathbb{I}, r_*; \delta)\}$ is μ -integrable.

Remark 5. For any pair $(\Psi_1, \Psi_2) \in \mathcal{I}$, if Ψ_1 strongly increases at infinity (for example, $\Psi_1(x) \propto x^p$ for some $p < 1$ close to one), then Ψ_2 slowly increases ($\Psi_2(x) \propto x^{1-p}$ for some $1 - p$ close to zero) (Krasnosel'skii and Rutickii, 1961, Theorem 2.1. Chapter 1). Hence, the stronger the norm, the weaker the rate (and conversely). This compromise between the rate function and the norm of convergence is well-known for discrete parameter Markov chain (Tuominen and Tweedie (1994); see also Jarner and Roberts (2002), Fort and Moulines (2003), Douc et al. (2004)). As expected, this property remains valid for continuous-time Markov process.

Corollary 6 provides a condition based on \mathcal{A} , well-adapted to prove polynomial ergodicity.

Corollary 6. *Let $1 \leq V < \infty$ be a Borel function and $0 < \alpha \leq 1$. Assume that*

- (i) *some skeleton chain P^m is irreducible.*
- (ii) *there exists a closed petite set C such that $\sup_C V < \infty$ and for all $\alpha \leq \eta \leq 1$, $t \mapsto V^{\eta-\alpha}(X_t)$ is integrable \mathbb{P}_x -a.s. and*

$$\mathcal{A}V^\eta \leq -c_\eta V^{\eta-\alpha} + b\mathbb{I}_C, \quad 0 \leq b < \infty, 0 < c_\eta < \infty. \quad (8)$$

Then there exists an unique invariant distribution π , $\pi(V^{1-\alpha}) < \infty$ and for all $0 < p < 1$ and $b \in \mathbb{R}$ or $p = 1$ and $b \geq 0$ or $p = 0$ and $b \leq 0$,

$$\lim_{t \rightarrow +\infty} (1+t)^{(1-p)(1-\alpha)/\alpha} (\log t)^b \|P^t(x, \cdot) - \pi(\cdot)\|_{V^{(1-\alpha)p} (\ln V)^{-b} \vee 1} = 0 \quad x \in \mathcal{X}.$$

The proof is given in Section 4.2. From (ii), we obtain $G_C(x, V^{1-\alpha}, \mathbb{I}; \delta) + G_C(x, \mathbb{I}, (1+t)^{1/\alpha-1}; \delta) \leq cV(x)$; and then we apply Theorem 1.

By choosing $b = 0$ and $p = (1 - \kappa\alpha)/(1 - \alpha)$ for some $1 \leq \kappa \leq 1/\alpha$, Corollary 6 yields

$$\forall x \in \mathcal{X}, \quad \lim_{t \rightarrow \infty} (t+1)^{\kappa-1} \|P^t(x, \cdot) - \pi(\cdot)\|_{V^{1-\kappa\alpha}} = 0. \quad (9)$$

If (9) holds for some V function, we shall say that the Markov chain is *polynomially ergodic* with rate $(1+t)^{(1-\alpha)/\alpha}$.

Remark 7. Corollary 6 can be compared to the paper by Jarner and Roberts (2002) for the discrete parameter case. They start with proving that if there exist a Borel function $1 \leq V < \infty$, $0 < \alpha \leq 1$, a set C such that for all $\alpha \leq \eta \leq 1$

$$P^m V^\eta - V^\eta \leq -c_\eta V^{\eta-\alpha} + b\mathbb{I}_C \quad 0 \leq b < \infty, 0 < c_\eta < \infty, \quad (10)$$

there exists $c < \infty$ such that $G_C^{(m)}(\cdot, V^{1-\alpha}, \mathbb{I}) + G_C^{(m)}(\cdot, \mathbb{I}, (1+t)^{1/\alpha-1}) \leq cV$, where $G_C^{(m)}$ is given by (5). The drift condition (10) is analogous to (8) and the controls of the moments $G_C^{(m)}$ and G_C are similar. If in addition P^m is irreducible, aperiodic and C is petite for the skeleton, P^m is positive with invariant distribution π such that $\pi(V^{1-\alpha}) < \infty$ and for all $1 \leq \kappa \leq 1/\alpha$, the skeleton is $V^{1-\kappa\alpha}$ -ergodic with rate $(n+1)^{\kappa-1}$. These rates coincide with those in (9).

Remark 8. From the proof of Corollary 6, it may be read than only a finite number of nested drift conditions is required; nevertheless, in practice, it is not more restricting to verify a continuum of drift conditions than to verify a finite number of drift conditions. More precisely, assumption (ii) can be substituted for the conditions : (iii) there exist a closed petite set and functions $1 \leq V_{q-1} \leq cf_q$, such that for all integers $1 \leq q \leq p$, $\mathcal{A}V_q \leq -f_q + b\mathbb{I}_C$, $t \mapsto f_q(X_t)$ is integrable \mathbb{P}_x -a.s., and $\sup_C V_p < \infty$; (iv) there exists $\beta > 0$ such that $\mathbb{E}_x[\tau_C^\beta] \leq f_1(x)$. If such, following the same lines as in the proof of Proposition 26, it may be proved that $G_C(\cdot, \mathbb{I}, (t+1)^{p-1+\beta}; \delta) + G_C(\cdot, f_p, \mathbb{I}; \delta) \leq cV_p$ for some $\delta > 0$. Together with condition (i), this yields $f_*^{1-\eta}$ -ergodicity at a rate $(t+1)^{(p-1+\beta)\eta}$ for all $0 \leq \eta \leq 1$, where f_* is any function satisfying $\sup_{t \leq m} P^t f_* \leq f_p$. Dai and Meyn (1995) (hereafter DM) are, to our best knowledge, the first to exhibit this kind of nested drift conditions and hence, the first to address ergodicity at a polynomial rate; they proved this yields f_1 -ergodicity at a rate $(t+1)^{p-1}$ (Theorem 6.3, DM). We are able to obtain the same result : to that goal, we observe that conditions (iii-iv) are verified with functions $f_k \leq f_p^{k/p}$, $\beta = 1$ (as a consequence of Proposition 5.3 and Eq. (6.1), in DM) and $f_* \propto f_p$.

We proved that nested drift conditions on the generator \mathcal{A} provide a control of moments G_C with a polynomially increasing rate function. The converse seems to be an open question. We nevertheless make mention of Propositions 5.4 and 6.1 in Dai and Meyn (1995), that provide a (partial) converse condition : from the condition $\sup_C G_C(x, f, \mathbb{I}; \delta) < \infty$, they deduce a drift condition on \mathcal{A} (we point out that this single condition implies a continuum of conditions by using the same convexity argument as in (Jarner and Roberts, 2002, Lemma 3.5)). Unfortunately, this drift condition in turn implies only a control of the moment $G_C(x, Rf, \mathbb{I}; \delta)$ where $Rf(x)$ is a function, which is, in general difficult to compare with f .

3. EXAMPLES

In this section, $\langle \cdot, \cdot \rangle$ and $|\cdot|$ denote respectively the scalar product and the Euclidean norm in \mathbb{R}^n . If u is a twice continuously differentiable real valued function on \mathbb{R}^n , ∇u (resp. $\nabla^2 u$) denotes its gradient (resp. its Hessian matrix); and $\partial u / \partial x_i$ its partial derivative with respect to the i -th variable. For a matrix u , $\text{Tr}(u)$ stands for the matrix trace and u' the matrix transpose. For $r \in \Lambda$, define the sequence r^0 by $r^0(t) = \int_0^t r(s) ds$. Finally, we largely make use of the inequality $r(s+t) \leq r(s)r(t)$, $s, t \geq 0$, which holds for any rate $r \in \Lambda$ (Thorisson, 1985, Lemma 1).

3.1. Toy example: Jump process. Consider the jump process on \mathbb{Z}_+ such that given that $X_t = i$, the waiting-time to the next jump has an exponential distribution with expectation λ_i^{-1} and is independent of the past history. We assume that for all $i \geq 0$, $\lambda_i > 0$, and $\sup_{i \geq 0} \lambda_i < \infty$. The probability that the jump leads to state j is given by the matrix entry $Q(i, j)$. We consider the case when $Q(0, i) = p_i$ and $Q(i, 0) = 1$ for all $i \geq 1$, for some positive sequence $(p_i)_{i \geq 1}$ such that $\sum_{i \geq 1} p_i = 1$. We assume in addition that

$$\liminf_i \lambda_i = 0 \quad \text{and} \quad \sum_{i \geq 0} p_i \lambda_i^{-1} < \infty. \quad (11)$$

Since $\sup_{i \geq 0} \lambda_i < \infty$, there exists a \mathbb{Z}_+ -valued right-continuous strong Markov process satisfying the heuristic description above and such that for all $(i, j) \in \mathbb{Z}_+^2$, the limit exists

$$\lim_{t \rightarrow 0} \frac{P^t(i, j) - \delta_i(j)}{t} =: A(i, j) < \infty \quad (12)$$

where δ_i is the Dirac-mass at point i , and for all $i \geq 1$

$$A(0, 0) = -\lambda_0, \quad A(0, i) = \lambda_0 p_i, \quad A(i, 0) = -A(i, i) = \lambda_i, \quad (13)$$

and $A(i, j) = 0$ otherwise (see e.g. (Feller, 1971, p.330)).

Lemma 9. *The process is Harris-recurrent, reversible with invariant distribution π given by $\pi(0) = \{1 + \sum_{j \geq 1} p_j \lambda_j^{-1}\}^{-1}$ and $\pi(i) = p_i \lambda_i^{-1} \pi(0)$, $i \geq 1$. Any skeleton chain P^m is irreducible.*

Proof. We have $\mathbb{E}_i[\tau_0] = (1 - \mathbb{1}_0(i)) \lambda_i^{-1}$ and for all $i \geq 0$, $j \neq i$, $\delta > 0$, $P^\delta(i, j) \leq p_j$. Then, $\mathbb{E}_i[\tau_0(\delta)] = \delta + \sum_{j \geq 1} P^\delta(i, j) \mathbb{E}_j[\tau_0] \leq \delta + 2 \sum_{j \geq 1} p_j \lambda_j^{-1}$. Hence, for all $i \in \mathbb{Z}_+$, $\mathbb{P}_i(\tau_0(\delta) < \infty) = 1$ and as $\{0\}$ is a closed petite set, the process is Harris-recurrent. π is the unique invariant

probability measure (as unique solution to $\pi \mathbf{A} = 0$), and since X obeys the detailed balance *i.e.* $\pi(i)\mathbf{A}(i, j) = \pi(j)\mathbf{A}(j, i)$ for all i, j the process is reversible. Finally, for all $m > 0$, and $i, j > 0$

$$P^m(i, j) \geq p_j \lambda_0 \lambda_i \lambda_j \int_0^m ds \exp(-\lambda_i s) \int_0^{m-s} dt \exp(-\lambda_0 t) \int_0^{m-(t+s)} du \exp(-\lambda_j u) > 0,$$

where the inequality says that $P^m(i, j)$ is greater than the probability of a single visit to 0 before a jump to j . Similarly, it is easy to prove that $P^m(0, j) > 0$ and $P^m(j, 0) > 0$ for any $j \in \mathbb{Z}_+$. This proves the irreducibility of any skeleton. \square

We deduce from R3 that the process is ergodic. Nevertheless, this convergence fails to occur at a geometric rate as shown in lemma 10, the proof of which relies on the notion of conductance.

Lemma 10. *X fails to be geometrically ergodic.*

Proof. As X is reversible, any Markov kernel P^m is reversible. It is proved in Lawler and Sokal (1988) that for a reversible Markov kernel P^m , the conductance κ_m given by $\kappa_m := \inf_A c_m(A)$ where $c_m(A) := \{\pi(A)\pi(A^c)\}^{-1} \int_A P^m(x, A^c) \pi(dx)$, is positive if and only if P^m is geometrically ergodic. We verify that for any skeleton P^m , the conductance is zero which will involve that the skeleton fails to be geometrically ergodic. Consider the set of states i such that $\pi(i) \leq 1/2$. Then $c_m(i) \leq 2(1 - P^m(i, i)) \leq 2(1 - \exp(-\lambda_i m))$ upon noting that $P^m(i, i)$ is lower bounded by the probability that the waiting-time in state i is greater than m . Since $\liminf_{i \rightarrow +\infty} \lambda_i = 0$, for all $\epsilon > 0$, there exists a state i such that $c_m(i) \leq \epsilon$, which involves $\kappa_m = 0$. \square

We now identify functions V that are in the domain of \mathcal{A} .

Lemma 11. *Let $0 \leq V < \infty$ be a Borel function such that $\sum_{i \geq 1} p_i V(i) < \infty$. Then V is in the domain of \mathcal{A} and $\mathcal{A}V = AV$.*

Proof. For a function $f \geq 0$ such that $\sum_j p_j f(j) < \infty$, the monotone convergence theorem yields

$$\sum_{j \geq 1} \lim_{t \downarrow 0} \frac{P^t(i, j) - \delta_i(j)}{t} f(j) = \lim_{J \uparrow +\infty} \sum_{j=0}^J \lim_{t \downarrow 0} \frac{P^t(i, j) - \delta_i(j)}{t} f(j) = \mathbf{A}f(i);$$

in addition, $\sup_{\mathbb{Z}_+} f^{-1} |\mathbf{A}f| < \infty$. This proves that V is in the domain of the weak infinitesimal generator \mathbf{A} , and thus in the domain of \mathcal{A} . \square

The expression of the generator suggests that function V on the form $\lambda_i^{-\rho}$ is a candidate to solve the drift inequality (8). This yields f -ergodicity at a log-polynomial rate.

Proposition 12. *Assume that there exists $\beta \geq 1$ such that $\sum_{i \geq 1} p_i \lambda_i^{-\beta} < \infty$. Then for all $i \in \mathbb{Z}_+$, $0 < \kappa < \beta - 1$ and $b \in \mathbb{R}$ or $\kappa = 0$ and $b \leq 0$, or $\kappa = \beta - 1$ and $b \geq 0$,*

$$\lim_{t \rightarrow \infty} (1+t)^{\beta-1-\kappa} [\ln(1+t)]^b \|P^t(i, \cdot) - \pi(\cdot)\|_{1+\lambda_x^{-\kappa} [\ln(1+\lambda_x^{-1})]^{-b}} = 0.$$

Proof. We apply Corollary 6 : we choose $V \geq 1$ such that for all $i \geq 1$, $V(i) = c^{-1}V(0)\lambda_i^{-\beta}$ for some $c > V(0)$. Then (8) is verified with $\alpha = \beta^{-1}$ and the closed petite set $C = \{0\}$. \square

When $\beta = 1$ (i.e. with nothing more than the condition (11)), this establishes the convergence in total variation norm at the rate \mathbb{I} , which corroborates the ergodicity of the process proved above. Nevertheless, if for some $\beta > 0$, the sum $\sum_{i \geq 1} p_i (1 \vee \lambda_i^{-1}) [\log(1 \vee \lambda_i^{-1})]^\beta$ exists, Corollary 6 does not yield a stronger convergence result than the ergodic one. We prove by application of Theorem 1, that covers more general rates than the polynomial ones, that convergence in total variation norm occurs at the rate $r_*(t) \sim [\log(t)]^\beta$, and convergence in norm $f_*(x) = [\log(1 \vee \lambda_x^{-1}) + 1]^\beta$ occurs at rate \mathbb{I} . We also derive sufficient conditions for sub-exponential ergodicity.

Lemma 13. *Let $f_* : \mathbb{Z}_+ \rightarrow [1, \infty)$ and $r_* \in \Lambda$ such that*

$$\sum_{j \geq 1} p_j (1 \vee \lambda_j^{-1}) f_*(j) < \infty \quad \text{and} \quad \sum_{j \geq 1} p_j \lambda_j^{-1} \int_0^{+\infty} r_*(s) \lambda_j \exp(-\lambda_j s) ds < \infty. \quad (14)$$

Then there exists a finite constant c such that for all $m > 0$, $\sup_{t \leq m} P^t f_ \leq c f_*$. For all $\delta > 0$, there exists a finite constant c such that*

$$G_0(x, f_*, \mathbb{I}; \delta) \leq c (1 \vee \lambda_x^{-1}) f_*(x) \quad G_0(x, \mathbb{I}, r_*; \delta) \leq c \int_0^{+\infty} r_*(s) \exp(-\lambda_x s) ds.$$

Proof. Since $P^t(x, j) \leq p_j$ for all $x \neq j$, it is trivial to prove that $\sup_{t > 0} \sup_{i \in \mathbb{Z}_+} f_*^{-1} P^t f_* < \infty$. For $f \geq 1$ and $r \in \Lambda$,

$$G_0(x, f, r; \delta) \leq \int_0^\delta r(s) P^s f(x) ds + r(\delta) \sum_{j \geq 1} P^\delta(x, j) f(j) \mathbb{E}_j[r^0(\tau_0)].$$

To conclude the proof, observe $\mathbb{E}_j[r^0(\tau_0)] = \lambda_j \int r^0(t) \exp(-\lambda_j s) ds = \int r(t) \exp(-\lambda_j s) ds$. \square

Proposition 14. (i) *Assume that $\sum_{i \geq 1} p_i (1 \vee \lambda_i^{-1}) [\log(1 \vee \lambda_i^{-1})]^\beta < \infty$, for some $\beta \geq 0$.*

For all $0 \leq \kappa \leq \beta$, $i \in \mathbb{Z}_+$, $\lim_{t \rightarrow +\infty} [\log(t+1)]^{\beta-\kappa} \|P^t(i, \cdot) - \pi(\cdot)\|_{[1+\log(1 \vee \lambda_i^{-1})]^\kappa} = 0$.

(ii) *Assume that $\sum_{i \geq 1} p_i (1 \vee \lambda_i^{-1}) \lambda_i^{-1/2} \exp(z^2 \lambda_i^{-1}) < \infty$, for some $z > 0$. For all $0 \leq p \leq 1$,*

$$i \in \mathbb{Z}_+, \lim_{t \rightarrow +\infty} \exp(2z(1-p)t^{1/2}) \|P^t(i, \cdot) - \pi(\cdot)\|_{[1+\lambda_i^{-1/2} \exp(z^2 \lambda_i^{-1})]^p} = 0.$$

Proof. In both cases, apply Theorem 1; for case (i), set $r_*(t) = \{\log(\exp(\beta - 1) + t)\}^\beta$ and $f_*(i) = 1 + \log(1 \vee \lambda_i^{-1})^\beta$; and for case (ii), set $r_*(t) = \exp(2zt^{1/2})$, $f_*(i) = 1 + \lambda_i^{-1/2} \exp(z^2 \lambda_i^{-1})$ and observe that $\int \exp(2zs^{1/2}) \lambda \exp(-\lambda s) ds \leq 1 + 2\sqrt{\pi} z \lambda^{-1/2} \exp(z^2 \lambda^{-1})$. \square

3.2. Langevin Tempered diffusions on \mathbb{R}^n . Let us consider a stochastic integral equation

$$X_t = X_0 + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s, \quad (15)$$

where W_t is a n -dimensional Brownian motion, the drift coefficient $b = (b_1, \dots, b_n)'$ is on the form, $1 \leq i \leq n$,

$$b_i(x) = \frac{1}{2} \sum_{j=1}^n a_{i,j}(x) \frac{\partial}{\partial x_j} \log \pi(x) + \frac{1}{2} \sum_{j=1}^n \frac{\partial}{\partial x_j} a_{i,j}(x),$$

where $a = \sigma \sigma'$ is the $n \times n$ symmetric positive definite matrix. Such a diffusion is the so-called Langevin diffusion and is defined in such a way that π is, up to a multiplicative constant, the density of the unique invariant probability distribution (with respect to Lebesgue measure on \mathbb{R}^n). This property motivates recent interests in Langevin diffusion for their use as MCMC methods, where the scope of these techniques is to draw samples from a Markov chain with given stationary density π . The efficiency of these algorithms is linked to the rate at which f -moments $\mathbb{E}_x[f(X_t)]$ converge to the constant $\pi(f)$. This motivates the study of the f -ergodicity. In practice, discretizations of the continuous-time process are used to solve the MCMC simulation problem and recent works proved that it is possible to find methods of discretizing which inherit the convergence rates of the continuous-time diffusion (see Roberts and Tweedie (1996); Stramer and Tweedie (1999a,b); Roberts and Stramer (2003); Roberts and Tweedie (2002) for methods of discretizing and their use in MCMC techniques). Roberts and Tweedie proved that, on the real line, when the target density π is heavy tailed, the Langevin diffusion with $a := 1$ can not be geometrically ergodic. We complement this assertion when π is polynomial in the tails, and prove that the Langevin diffusion in the one-dimensional case as well as in the multi-dimensional one is f -ergodic at a polynomial rate. For such polynomial target density on the real line, it was observed in Jarner and Roberts (2001) that the polynomial rate of convergence of the Metropolis-Hastings algorithm could be improved by choosing a heavy-tailed proposal distribution. This idea, when adapted to the diffusion on the real line, suggests the choice of a speed measure *i.e.* of the coefficient σ such that σ is small when the process is close to the modes of π and big when far from the modes (Stramer and Tweedie (1999a)). In the multi-dimensional case, this suggests $a(x)$ on the form $\pi^{-2d}(x) \mathbb{I}_n$, where \mathbb{I}_n is the identity matrix on \mathbb{R}^n , $d > 0$. In that case ($d > 0$) we call these processes Langevin tempered diffusion (see Roberts and Stramer

(2003) for the justification of these heated diffusions). It was observed in the literature that by choosing d large enough, a diffusion on the real line with target density polynomial in the tails is geometrically ergodic. We investigate the behavior of this Langevin tempered diffusions in the multi-dimensional case, contrarily to most of the mentioned contributions that cover the one-dimensional case. In Theorem 16, it is proved that up to some critical temperature d_* the diffusion is polynomially ergodic and the larger d , the better the rate. When $d \geq d_*$, the diffusion is geometrically ergodic. We henceforth consider a diffusion matrix $a(x) = \sigma^2(x)\mathbb{I}_n$ where $\sigma(x) := \pi^{-d}(x)$ for some $d \geq 0$. Assume that

A1 π is, up to a multiplicative constant, a positive and twice continuously differentiable density on \mathbb{R}^n (with respect to Lebesgue measure).

Define the drift vector

$$b(x) := \frac{1}{2}\sigma^2(x) (\nabla \log\{\pi(x)\sigma^2(x)\}) = \frac{1-2d}{2} \pi^{-2d}(x) \nabla \log \pi(x). \quad (16)$$

Under A1, the coefficients b and σ are locally Lipschitz-continuous, which implies that for any compact set \mathcal{K} , $\sup_{x \in \mathcal{K}}\{|b(x)| + |\sigma(x)|\}(1 + |x|)^{-1} < \infty$. These local conditions allow the construction of a continuous process satisfying the stochastic integral equation (15) up to the explosion time $\zeta := \lim_{n \rightarrow \infty} \zeta_n$, where $\zeta_n := \inf\{t \geq 0, |X_t| \geq n\}$. We thus formulate the following assumption

A2 The process is regular *i.e.* $\zeta = +\infty$ a.s.

Under A1, a sufficient condition for regularity is the existence of a twice continuously differentiable non-negative function V and a constant $c \geq 0$ such that $\mathsf{L}V \leq cV$ on \mathbb{R}^n and $\lim_{n \rightarrow \infty} \inf_{|x| \geq n} V(x) = +\infty$ (Has'minskii, 1980, Theorem 3.4.1.) where L is the elliptic operator

$$\begin{aligned} \mathsf{L}V(x) &= \langle b(x), \nabla V(x) \rangle + \frac{\text{Tr}(\nabla^2 V(x) a(x))}{2} \\ &= \frac{\pi^{-2d}(x)}{2} \left((1-2d) \langle \nabla \log \pi(x), \nabla V(x) \rangle + \sum_{i=1}^n \frac{\partial^2 V(x)}{\partial x_i^2} \right). \end{aligned}$$

In the one dimensional case, (Has'minskii, 1980, Remark 2, p.105) establishes that the process is regular if d is chosen such that

$$\int_{\mathbb{R}} \pi^{2d-1}(x) dx = +\infty, \quad (17)$$

since the function $V(x) := \text{sign}(x) \int_0^x Q(y)dy$ where $\ln Q(x) = -2 \int_0^x b(t)\sigma^{-2}(t)dt = (2d - 1)(\ln \pi(x) - \ln \pi(0))$ is finite and satisfies $LV = 0$ on \mathbb{R} . To cover the multi-dimensional case, we adapt this condition and claim that the process is regular if d is chosen such that

$$\int_r^\infty t^{1-n} \exp \left(-(1-2d) \int_r^t s^{-1} \sup_{\{x, |x|=s\}} \langle \nabla \log \pi(x), x \rangle ds \right) dt = +\infty. \quad (18)$$

Indeed, the function $V(x) := U(|x|)$ where for all $u \geq 0$,

$$U(u) := \int_r^u \exp \left(- \int_r^t \sup_{\{x, |x|=s\}} \left\{ \left\langle \frac{2b(x)}{\sigma^2(x)}, \frac{x}{|x|} \right\rangle + \frac{n-1}{|x|} \right\} ds \right) dt,$$

is finite and satisfies $LV = 0$ on \mathbb{R}^n .

In the one-dimensional case, condition (17) is necessary for the existence of an invariant probability measure (Has'minskii, 1980, Remark 2, p.105); thus, for the objective of this paper, d has to be chosen in the set \mathcal{D}_1 of the positive real numbers such that (17) hold. Observe that \mathcal{D}_1 is non-empty and contains $\{0, 1/2\}$. In the multi-dimensional case, a necessary condition for (positive) recurrence is that d checks a condition on the form (18) where the supremum is replaced by the infimum (Has'minskii, 1960, Theorem II p.194). This involves the definition of an interval \mathcal{D}_n limiting the range of the possible temperature d .

Under A1-2, there exists a solution $(\Omega, \mathcal{F}, (\mathcal{F}_t), (W_t), (X_t), \mathbb{P})$ where $(\Omega, \mathcal{F}, (\mathcal{F}_t), (W_t), \mathbb{P})$ is n -dimensional Brownian motion, $(X_t)_t$ is a \mathcal{F}_t -adapted homogeneous and continuous Markov process with Feller transition probability, satisfying (15) \mathbb{P} -a.s. and such that both the integral exist *i.e.* for all $t > 0$,

$$\mathbb{P} \left(\int_0^t b(X_s)ds + \int_0^t \sigma^2(X_s)ds < \infty \right) = 1. \quad (19)$$

A transition semi-group $(P^t)_{t \geq 0}$ has the Feller property if for any continuous bounded real-valued function f , $x \mapsto P^t f(x)$ is continuous. (X_t) is thus a strongly Markovian process as a (right)-continuous process with Feller transition probability (Dynkin (1965)).

Let $0 \leq V < \infty$ be a twice continuously differentiable function such that there exists a non-negative Borel function ϕ , bounded on compact sets, a constant $b < \infty$ and a compact set C such that $LV \leq -\phi \mathbb{I}_{C^c} + b \mathbb{I}_C$. From (19) and the continuity of $t \mapsto \nabla V(X_t)$, the process $t \mapsto \int_0^t \sigma(X_s) \{ \nabla V(X_s) \}' dW_s$ is a local martingale. Application of the Itô's rule yields $LV = \mathcal{A}V$.

A3 For all $1 \leq i, j \leq n$, $\partial^2 \sigma^2(x) / \partial x_i \partial x_j$ and $\partial^2 \log \pi(x) / \partial x_i \partial x_j$ are locally uniformly Hölder continuous.

Proposition 15. *Under A1-3, the process is reversible and π is, up to a multiplicative constant, the density of an invariant probability measure. Any skeleton chain is irreducible, and compact sets are closed petite sets.*

Proof. There exists a continuous function $p : (t, x, y) \mapsto p(t, x, y)$ such that $P^t(x, dy) = p(t, x, y)dy$ (Kent, 1978, Theorem 1.1.). Since the process is regular (or conservative, in the terminology of Kent) and π is Lebesgue integrable, (Kent, 1978, Theorems 4.1. and 6.2.) imply that the process is time-reversible and

$$\lim_{t \rightarrow \infty} \int_A p(t, x, y) dy = \left(\int \pi(x) dx \right)^{-1} \int_A \pi(x) dx. \quad (20)$$

Hence, $\pi(dx)$ is invariant. Irreducibility of skeletons results from the (20), and petiteness of compact sets from the continuity of $p(t, \cdot, \cdot)$. \square

Finally, we restrict our attention to densities π that are polynomially decreasing in the tails.

A4 π satisfies A1 and A3 and there exists some $0 < \beta < 1/n$,

$$\begin{aligned} 0 < \liminf_{|x| \rightarrow +\infty} \frac{|\nabla \log \pi(x)|}{\pi^\beta(x)} &\leq \limsup_{|x| \rightarrow +\infty} \frac{|\nabla \log \pi(x)|}{\pi^\beta(x)} < \infty, \\ 2\beta - 1 < \liminf_{|x| \rightarrow +\infty} \frac{\text{Tr}(\nabla^2 \log \pi(x))}{|\nabla \log \pi(x)|^2} &\leq \limsup_{|x| \rightarrow +\infty} \frac{\text{Tr}(\nabla^2 \log \pi(x))}{|\nabla \log \pi(x)|^2} < \infty. \end{aligned}$$

Set $\gamma = \liminf_{|x| \rightarrow +\infty} \text{Tr}(\nabla^2 \log \pi(x)) |\nabla \log \pi(x)|^{-2}$.

This class is non empty and contains the densities that are polynomially decreasing in the tails $\pi(x) = c|x|^{-1/\beta}$ for large $|x|$, where $0 < \beta < 1/n$; in that case, $\gamma = \beta(2 - n) > 2\beta - 1$. For this family, the regularity criterion (17) or (18) says that the temperature d has to be chosen in $\mathcal{D}_n = [0, (1 + \beta(2 - n))/2]$. For any density in the class A4, $0 < \liminf_{|x|} |x|^\beta \pi(x) \leq \limsup_{|x|} |x|^\beta \pi(x) < \infty$. Hence $\mathcal{D}_1 = [0, (1 + \beta)/2]$ and for $n \geq 3$, $1/2 \notin \mathcal{D}_n$. If $\sup_{s \geq r} \sup_{\{x, |x|=s\}} \langle \nabla \log \pi(x), x \rangle =: -\varrho^{-1} < 0$ exists, then $[0, 1/2 + \varrho(1 - n/2)] \subseteq \mathcal{D}_n$.

It is readily seen that setting, $V = 1 + \text{sign}(\rho)\pi^{-\rho}$ outside a compact set, and $V = 1$ otherwise,

$$\mathbb{L}V = -\frac{|\rho|}{2} V \frac{\pi^{-\rho}}{1 + \pi^{-\rho}} \pi^{2(\beta-d)} \left(\frac{|\nabla \log \pi|}{\pi^\beta} \right)^2 \left(1 - \rho - 2d + \frac{\text{Tr}(\nabla^2 \log \pi)}{|\nabla \log \pi|^2} \right), \quad (21)$$

for large $|x|$. As established in (Stramer and Tweedie, 1999a, Theorem 3.1.), the diffusion can not be geometrically ergodic when $0 \leq d < \beta$: by choosing $f := \pi^{d-\beta}$ and applying Itô's formula,

$df(X_t) \approx c_1 \pi^{\beta-d}(X_t)dt + c_2 dW_t$ for some constants c_i ; and the drift coefficient tends to zero for large value of the process. The process $(f(X_t))_t$ fails to be geometrically ergodic, and henceforth, $(X_t)_t$ itself.

From (21), for large $|x|$,

$$LV \leq -cV^{1-\alpha} \quad \text{where } \alpha := 2\rho^{-1}(\beta - d), \quad \text{and } c > 0 \iff 1 + \gamma - \rho - 2d > 0.$$

In any cases, one has to choose ρ such that $c > 0$. If $\alpha \leq 0$ and $\rho > 0$, then the process is geometrically V -ergodic (Meyn and Tweedie, 1993d, Theorem 6.1.). If $0 < \alpha \leq 1$ and $\rho > 0$, the diffusion is polynomially ergodic as discussed in Section 2. If $\alpha \leq 1$ and ρ can be set negative, the process is uniformly ergodic *i.e.* there exist $\kappa > 1$ and a constant c such that for all x , $\lim_{t \rightarrow \infty} \kappa^t \|P^t(x, \cdot) - \pi(\cdot)\|_{TV} \leq c$ and the convergence does not depend on the starting point. This yields Theorem 16 : the first assertion results from Roberts and Tweedie (1996) and Corollary 6 of the present paper. The second and third assertions result from (Meyn and Tweedie, 1993d, Theorem 6.1.).

Theorem 16. *Consider the Langevin tempered diffusion on \mathbb{R}^n where the target density π is from the class A_4 and $\sigma := \pi^{-d}$ for some d satisfying (17) if $n = 1$ or (18) if $n \geq 2$.*

(i) *If $0 \leq d < \beta$, the process fails to be geometrically ergodic. For all $0 \leq \kappa < 1 + \gamma - 2\beta$,*

$$\lim_{t \rightarrow +\infty} (t+1)^\tau \|P^t(x, \cdot) - \pi(\cdot)\|_{1+\pi^{-\kappa}} = 0 \quad \tau < \frac{1 + \gamma - 2\beta - \kappa}{2(\beta - d)}. \quad (22)$$

(ii) *If $\beta \leq d < (1 + \gamma)/2$, then for all $0 < \kappa < 1 + \gamma - 2d$, the diffusion is geometrically V -ergodic with $V := 1 + \pi^{-\kappa}$.*

(iii) *If $\beta < d$, the diffusion is uniformly ergodic.*

Theorem 16 extends earlier results to the multi-dimensional case and provides polynomial rates of convergence of the “cold” Langevin tempered diffusions for a wide family of norms. In the one-dimensional case, when $d = 0$, (Ganidis et al., 1999, Result R3, p.245) only claim that the convergence in total variation norm is polynomial, with no explicit value of the rate of convergence. We establish that for a given $\pi^{-\kappa}$ -norm, the minimal rate of convergence is achieved with $d = 0$ and in that case coincides with the rate of convergence of the symmetric random walk Hastings-Metropolis algorithm with light proposal distribution (Jarner and Roberts (2001)). By choosing a diffusion matrix which is heavy where the target distribution is light, and conversely, improves the rate of convergence as evidenced by (22). The critic temperature is $d = \beta$. For $d \geq \beta$, the diffusion is no more polynomially ergodic and geometric rates can be reached. This

critic temperature coincides with the critic one given in (Stramer and Tweedie, 1999a, Theorem 3.1.) for the real-valued diffusion.

Remark: General diffusions on \mathbb{R}^n . The techniques above can be adapted for the analyzes of diffusions satisfying (15). Under conditions implying (a) the existence of a solution, (b) the condition (i) of Corollary 6 and (c) the petiteness property of the compact sets, (see e.g. Has'minskii (1980); Veretennikov (1999); Malyskin (2001)), we are able to prove that when there exist $M, \beta, \gamma > 0$ and $l < 2$ such that

$$\begin{aligned} \sup_{\{x, |x| \geq M\}} |x|^{-(2+l)} \langle x, a(x)x \rangle &=: \beta, & \sup_{\{x, |x| \geq M\}} |x|^{-l} \text{Tr}(a(x)) &=: \gamma \\ \sup_{\{x, |x| \geq M\}} |x|^{-l} \langle b(x), x \rangle &=: -r & \text{for some } r > (\gamma - \beta l)/2, \end{aligned}$$

then the diffusion is polynomially ergodic and for all x , for all $0 \leq \kappa < l + \beta^{-1}(2r - \gamma)$

$$\lim_{t \rightarrow \infty} (1+t)^\tau \|P^t(x, \cdot) - \pi(\cdot)\|_{1+|x|^\kappa} = 0, \quad \tau < \frac{2(r + \beta) - \gamma}{\beta(2-l)} - 1 - \frac{\kappa}{2-l}.$$

3.3. Compound Poisson-Process driven Ornstein-Uhlenbeck processes. Let X be an Ornstein-Uhlenbeck process driven by a finite rate subordinator:

$$dX_t = -\mu X_t dt + dZ_t,$$

where $Z_t := \sum_{i=1}^{N_t} W_i$, $\{W_i\}_{i \geq 1}$ is an independent and identically distributed collection of random variables from probability measure F and $\{N_t\}_{t \geq 0}$ is a Poisson-Process of finite rate λ , independent of the collection $\{W_i\}_{i \geq 1}$. Such processes are used as storage models (see for example Lund et al. (1996)) and have recently been used in financial econometrics as models for stochastic volatility (see Barndorff-Nielsen and Shephard (2001)).

The exponential decay of X except at jump points leads to geometric ergodicity of X when the tails of $F(\cdot)$ are sufficiently light. Here we shall explore the case where $F(\cdot)$ is extremely heavy-tailed. First we make this concept precise: we say a probability measure is *heavy-tailed* if, under that probability measure, for all $\kappa > 0$, $\mathbf{E}[e^{\kappa X}] = \infty$. Now let G denote the law of the log jump sizes, that is $G(A) = F(e^A)$. We have the following negative result showing that for sufficiently heavy-tailed jumps, geometric ergodicity, and even ergodicity can fail. As usual, we let π denote the invariant probability measure (should it exist).

Lemma 17. (i) Suppose $\int x G(dx) = \infty$, then X fails to be positive recurrent.

(ii) Suppose G is heavy-tailed, then X fails to be geometrically ergodic.

Proof. Suppose $X_0 = 2$ and consider the petite set $C = [0, 1]$. Then

$$\mathbf{P}(\tau_C > t) \geq \mathbf{P}[\text{jump of size } \geq e^{\mu t} \text{ occurs before time } \log 2/\mu] = (1 - 2^{-\lambda/\mu}) \int_{\mu t}^{\infty} G(x) dx. \quad (23)$$

For positive recurrence, we require that $\mathbf{E}(\tau_C)$ to be finite, that is that $\mathbf{P}[\tau_C > t]$ be integrable. However the integrated right hand side of (23) is just

$$\int_0^{\infty} dt \int_{\mu t}^{\infty} G(x) dx = \int_0^{\infty} \mu^{-1} x G(x) dx = \infty$$

by hypothesis, so that $\mathbf{E}(\tau_C) = \infty$ too, so that positive recurrence must fail, proving (i). For (ii), we recall that for geometric ergodicity, we require that for some $\kappa > 0$, $\mathbf{E}[e^{\kappa \tau_C}] < \infty$. (Although not necessary, we shall again assume that $X_0 = 2$ and $C = [0, 1]$.) Thus from (23) we require that

$$\int_0^{\infty} e^{\kappa t} dt \int_{\mu t}^{\infty} G(x) dx = \mu \kappa^{-1} \int_0^{\infty} (e^{\kappa x} - 1) G(\mu x) dx < \infty. \quad (24)$$

However this is precluded by the heavy-tailed nature of G , thus proving (ii).

□

Examples of jump distributions for which geometric ergodicity fails (case 2 above), though we will see that X is positive Harris recurrent, include the following:

$$F(dx) = \frac{dx}{x(\log x)^k} \text{ at least for } k > 1; \quad F(dx) = \frac{e^{-(\log x)^\beta} dx}{x} \text{ for some } \beta \leq 1.$$

Lemma 18. *Suppose that for some $r > 1$, $m_r := \int_0^{\infty} [\log(1+u)]^r F(du) < \infty$. Then, X is polynomially ergodic with rate $(1+t)^{(r-1)}$.*

Proof. For differentiable functions V in the domain of \mathcal{A} ,

$$\mathcal{A}V = \int_0^{\infty} (V(x+u) - V(x)) \lambda F(du) - \mu x V'(x).$$

Now set $V(x) = (\log x)^r$, then by direct calculation,

$$\mathcal{A}V^\eta = \int_0^{\infty} ((\log(x+u))^{r\eta} - (\log x)^{r\eta}) \lambda F(du) - \frac{\mu x r \eta (\log x)^{r\eta-1}}{x}. \quad (25)$$

Now the finiteness of m_r merely ensures the finiteness of the first term on the right hand side of (25). So, noting that $(\log x)^{r\eta}$ is concave beyond $x = e^{r-1}$ for all $0 < \eta \leq 1$, we find that in fact the first term on the right hand side of (25) is bounded as a function of x , so that for some positive constant c ,

$$\mathcal{A}V^\eta \leq \int_0^{\infty} \frac{ur(\log x)^{r\eta-1}}{x} \lambda F(du) c - r\mu(\log x)^{r\eta-1}.$$

It is easy to check that all bounded sets are petite in this example, and therefore the conditions for the application of Corollary 6 with $\alpha = r^{-1}$.

4. PROOFS OF SECTION 2

When not explicitly defined, c denotes a generic finite positive constant. θ is the usual shift operator on the canonical probability space of the strong Markov process.

Lemma 19. *If Ψ^{-1} is a Young function and $r \in \Lambda_0$ (resp. Λ), $[\Psi(r) \vee 1] \in \Lambda_0$ (resp. Λ).*

Proof. Let $r \in \Lambda_0$. Ψ^{-1} is a continuous, increasing and convex function, so Ψ is measurable and bounded on bounded sets ((Krasnosel'skii and Rutickii, 1961, Chapter 1)). Furthermore, there exists a right-continuous non decreasing function ϕ such that $\ln \Psi(r(t)) = \ln r(t) + \ln \{r(t)^{-1} \int_0^{r(t)} \phi(s) ds\}$; thus proving that $\ln \Psi(r(t))/t \downarrow 0$ as $t \rightarrow \infty$. This yields $\Psi \in \Lambda_0$. The second assertion deduces easily from the definition of Λ and the upper bound $\sup_{t \geq 1} \Psi(at)/\Psi(t) < \infty$ for all $a > 0$ (Krasnosel'skii and Rutickii, 1961, Chapter 1, p.7-8). \square

While Theorem 1 and Corollary 6 are claimed for a rate function $r \in \Lambda$, Lemma 19 shows that they can be established for a rate $r \in \Lambda_0$, and we will do so.

4.1. Proof of Theorem 1. Without loss of generality, we assume $\Psi_1(r_*) \geq \mathbb{I}$ and $\Psi_2(f_*) \geq 1$.

Lemma 20. *Let $r \in \Lambda_0$ and $f \geq 1$ be a Borel function. For any closed set C such that $\sup_C G_C(\cdot, f, r; \delta) < \infty$, there exists a constant $M < \infty$ such that for all $x \in \mathcal{X}$ and $t \geq \delta$, $G_C(x, f, r; t) \leq M^{\lfloor t/\delta \rfloor} G_C(x, f, r; \delta)$.*

Proof. The proof is on the same lines than the proof of Lemma 4.1. Meyn and Tweedie (1993a) that addresses the case $r = \mathbb{I}$, and the details are omitted. Using the property $r(s+t) \leq r(s)r(t)$ (Thorisson, 1985, Lemma 1(d)), we obtain $M = 1 + \sup_{t \geq \delta} [r(t)/r^0(t)] \sup_C G_C(\cdot, f, r; \delta)$ which is finite since $\lim_t r(t)/r^0(t) = 0$ (this is a consequence of (Thorisson, 1985, Lemma 1)). \square

Proposition 21. *Let $r \in \Lambda_0$ and $f \geq 1$ be a Borel function. Assume that X is ϕ -irreducible and $\sup_C G_C(\cdot, f, r; \delta) < \infty$ for some closed petite set C and $\delta > 0$. $x \mapsto G_C(x, f, r; \delta)$ is finite ψ -almost surely for some (and then any) maximal irreducibility measure ψ , and C is accessible.*

Proof. By (Meyn and Tweedie, 1993a, Proposition 3.2.(ii)), for all $\lambda > 0$, there exist a positive integer m and a maximal irreducibility measure ψ such that $\psi(\cdot) \leq \inf_{x \in C} R_\lambda^m(x, \cdot)$, where R_λ is the resolvent kernel $R_\lambda(x, \cdot) = \int \lambda \exp(-\lambda t) P^t(x, \cdot) dt$. By Lemma 20, $R_\lambda G_C(\cdot, f, r; \delta)(x) \leq c G_C(x, f, r; \delta)$ where c is finite for some convenient λ . Hence, $\psi G_C(\cdot, f, r; \delta) < \infty$, proving the first statement. This implies that there exists an accessible set B such that $\sup_{x \in B} \mathbb{E}_x[\tau_C(\delta)] \leq \sup_{x \in B} G_C(x, f, r; \delta) < \infty$. Then for q large enough, $\inf_{x \in B} \mathbb{P}_x(\tau_C(\delta) \leq q) > 0$ and, for any x , $\mathbb{E}_x[\eta_C] \geq P^n(x, B) \inf_{x \in B} \mathbb{P}_x(\tau_C(\delta) \leq q) > 0$ for some n depending upon (x, B) . Hence C is accessible. \square

Proposition 22. *Suppose Assumptions (i-ii) of Theorem 1. Then*

- (i) *There exist t_0 and a measure ν such that $\inf_{t \geq t_0} \inf_{x \in C} P^t(x, \cdot) \geq \nu(\cdot)$, and $\nu(C) > 0$.*
- (ii) *For any set B such that $\nu(B) > 0$, $\mathbb{E}_x[r^0(T_{m,B})] \leq R_B \mathbb{E}_x[r^0(\tau_C(\delta))]$ for some finite constant $R_{t,B}$.*
- (iii) *For any $t \geq 0$ and any accessible set B , $\mathbb{E}_x[r^0(\tau_B(t))] \leq R_{t,B} \mathbb{E}_x[r^0(\tau_C(\delta))]$ for some finite constant $R_{t,B}$.*

Proof. (i) results from R2, Proposition 21 and Lemma 2. (ii) Let t_0 and ν be given by (i). Set $\tau = \tau_C(t_0 + m)$; and define the sequence of iterates $\tau^1 = \tau$ and for $n \geq 2$, $\tau^n = \tau^{n-1} + \tau \circ \theta^{\tau^{n-1}}$. Finally, let $(u_n)_{n \geq 2}$, be a $\{0, 1\}$ -valued process given by $u_n = 1$ if $X_{[(\tau^{n-1} + t_0)/m]m} \in B$ and 0 otherwise. $[t]$ denotes the upper integer part of t . Then $u_n \in \mathcal{H}_n$ with $\mathcal{H}_n = \sigma(X_t, t \leq \tau^n)$, and by the strong Markov property $\mathbb{P}_x(u_n = 1 | \mathcal{H}_{n-1}) \geq \nu(C) > 0$ for $n \geq 2$. Finally, set $\eta = \inf\{n \geq 2, u_n = 1\}$, so that $\mathbb{E}_x[r^0(T_{m,B})] \leq \mathbb{E}_x[r^0(\tau^\eta)]$. Using again the strong Markov property and the inequality $r^0(t_1 + t_2) \leq r^0(t_1) + r(t_1)r^0(t_2)$ (Thorisson (1985)),

$$\mathbb{E}_x[r^0(\tau^\eta)] \leq \sum_{n \geq 2} \mathbb{E}_x[r^0(\tau^n) \mathbb{1}_{\eta \geq n}] = \sum_{n \geq 2} \{a_x(n) + \sup_{x \in C} \mathbb{E}_x[r^0(\tau)] b_x(n)\}, \quad (26)$$

for all $n \geq 2$, where $a_x(n) = \mathbb{E}_x[r^0(\tau^{n-1}) \mathbb{1}_{\eta \geq n}]$ and $b_x(n) = \mathbb{E}_x[r(\tau^{n-1}) \mathbb{1}_{\eta \geq n}]$. Since, by Lemma 20, $\sup_C \mathbb{E}_x[r^0(\tau)] < \infty$, there exists $0 < \rho < 1$ and a finite constant c such that

$$b_x(n) \leq \rho b_x(n-1) + c(1 - \nu(C))^{n-1}, \quad a_x(n) \leq (1 - \nu(C))a_x(n-1) + b_x(n-1) \sup_{x \in C} \mathbb{E}_x[r^0(\tau)];$$

and $b_x(2) = \mathbb{E}_x[r(\tau)]$, $a_x(2) = \mathbb{E}_x[r^0(\tau)]$. The proof is on the same lines than the proof of (Nummelin and Tuominen, 1983, Lemma 3.1.) and is omitted for brevity. Hence, $\mathbb{E}_x[r^0(T_{m,B})] \leq c(\mathbb{E}_x[r^0(\tau)] + \mathbb{E}_x[r(\tau)])$ for some $c < \infty$. The proof is concluded, applying again Lemma 20 and the bound $\sup_{t \geq a} r(t)/r^0(t) < \infty$ for all $a > 0$ (see the proof of Lemma 19).

(iii) B is accessible and C petite so there exist $t_0 \geq 0$ and $\gamma > 0$ such that $\inf_{x \in C} \mathbb{P}_x(\tau_B \leq t_0 + t) \geq \inf_{x \in C} \mathbb{P}_x(\tau_B \leq t_0) \geq \gamma$. Set $\tau = \tau_C(t + t_0)$ and $u_n = 1$ if for some $\tau^{n-1} \leq s \leq \tau^{n-1} + t + t_0$, $X_s \in B$; and $u_n = 0$ otherwise. Following the same lines as in the proof of (ii), it may be proved that there exists $c < \infty$ such that $\mathbb{E}_x[r^0(\tau_B(t))] \leq c \mathbb{E}_x[r^0(\tau_C(t + t_0))]$. The proof is concluded by applying Lemma 20. \square

Proposition 23. *Suppose Assumptions (i-ii) of Theorem 1. For any $(\Psi_1, \Psi_2) \in \mathcal{I}$, C is a $(\Psi_2(f_*), \Psi_1(r_*))$ -regular set for the process i.e. $\sup_C G_B(\cdot, \Psi_2(f_*), \Psi_1(r_*); t) < \infty$ for any $t > 0$ and any accessible set B . $G_B(x, \Psi_2(f_*), \Psi_1(r_*); t) < \infty$ for all $x \in \mathcal{S}_\Psi$ and $\pi(\mathcal{S}_\Psi) = 1$.*

Proof. $(\Psi_2(f_*), \Psi_1(r_*))$ -regularity is a consequence of the Young's inequality (4), the (f_*, \mathbb{I}) -regularity of C (Meyn and Tweedie, 1993a, Proposition 4.1.) and Proposition 22(iii). For the second statement, write

$$G_B(x, \Psi_2(f_*), \Psi_1(r_*); t) \leq G_C(x, \Psi_2(f_*), \Psi_1(r_*); t) + \mathbb{E}_x \left[\int_{\tau_C(t)}^{\tau_B \circ \theta^{\tau_C(t)}} \Psi_1(r_*(s)) \Psi_2(f_*(X_s)) ds \right].$$

The result now follows from the strong Markov property, Lemma 20 and the inequality $\Psi_1(r_*(s+t)) \leq \Psi_1(r_*(s))\Psi_1(r_*(t))$ which holds since $\Psi_1 \circ r_* \in \Lambda_0$. Finally, $\pi(\mathcal{S}_\Psi) = 1$ by Proposition 21. \square

Proposition 24. *Suppose Assumptions (i-ii) of Theorem 1. The skeleton chain P^m is ψ -irreducible and aperiodic and possesses an accessible petite set A such that for all $(\Psi_1, \Psi_2) \in \mathcal{I}$,*

$$\sup_{x \in A} \mathbb{E}_x \left[\sum_{k=0}^{T_{m,A}-1} \Psi_1(r_*(k)) \Psi_2(f_*(X_{km})) \right] < \infty. \quad (27)$$

Proof. For the definitions of accessibility, smallness, petiteness, aperiodicity of a discrete-time Markov chains, see Meyn and Tweedie (1993b) (hereafter MT). From Proposition 22(i), C is small for the skeleton P^m and the skeleton is aperiodic (Theorem 5.4.4 MT). In addition, by R2, the skeleton is positive and $\pi(f_*) < \infty$. Let C_n be a petite set (for the skeleton P^m) such that $A = C \cap C_n$ is of positive ν -measure

$$\sup_{x \in C_n} \mathbb{E}_x \left[\sum_{k=0}^{T_{m,B}-1} f_*(X_{km}) \right] < \infty, \quad (28)$$

for any accessible set B (for the skeleton); the existence of such a set is a consequence of (Theorems 14.2.3 and 14.2.11, MT) and Proposition 22(ii). The set A is accessible and petite for the skeleton. (27) now results from the Young's inequality (4), (28) and Proposition 22(ii).

□

[Proof of Theorem 1]. By Proposition 24 and (Tuominen and Tweedie, 1994, Theorem 2.1 and Proposition 3.2), $\lim_{n \rightarrow \infty} \Psi_1(r_*(n)) \|P^{nm}(x, \cdot) - \pi(\cdot)\|_{\Psi_2(f_*)} = 0$ for π a.a. x . By Jensen's inequality, the upper bound $\sup_{t \geq 1} \Psi_2(at)/\Psi_2(t) < \infty$ for all $a > 0$, and assumption (iii), we have for all $t \leq m$, $P^t \Psi_2(f_*) \leq c \Psi_2(f_*)$. In addition, since $\Psi_1(r_*) \in \Lambda$, $\Psi_1(r_*(n+t)) \leq c \Psi_1(r_*(n))$ for all $t \leq m$ (Thorisson, 1985, Lemma 1). Hence,

$$\lim_{t \rightarrow \infty} \Psi_1(r_*(t)) \|P^t(x, \cdot) - \pi(\cdot)\|_{\Psi_2(f_*)} = 0, \quad \pi \text{ a.a. } x. \quad (29)$$

We now prove that this convergence occurs for all $x \in \mathcal{S}_\Psi$ which is of π -measure one, by Proposition 23. To that goal, we mimic the proof of (Meyn and Tweedie, 1993c, Theorem 7.2.). By Egorov's Theorem, there exists a set A , $\pi(A) > 0$, such that (29) holds uniformly for all $x \in A$. For all Borel function $g \in \mathcal{L}_{\Psi_2(f_*)}$, set $\bar{g} := g - \pi(g)$. Since $\Psi_1(r_*) \in \Lambda_0$,

$$\begin{aligned} \Psi_1(r_*(t)) |\mathbb{E}_x [\bar{g}(X_t) \mathbb{1}_{\tau_A \leq t}]| &\leq \Psi_1(r_*(t)) \int_0^t \sup_{y \in A} |P^{t-s} \bar{g}|(y) \mathbb{P}_x(\tau_A \in ds) \\ &\leq M \{ \mathbb{E}_x [\Psi_1(r_*(\tau_A))] + \mathbb{E}_x [\Psi_1(r_*(\tau_A)) \mathbb{1}_{\tau_A \geq t/2}] \}, \end{aligned}$$

where $M = \sup_{y \in A} \sup_{s \geq 0} r_*(s) |P^s \bar{g}|(y)$. Let $x \in \mathcal{S}_\Psi$; from Proposition 23, $\mathbb{E}_x [\Psi_1(r_*(\tau_A))] < \infty$ and $\lim_{t \rightarrow \infty} \mathbb{E}_x [\Psi_1(r_*(\tau_A)) \mathbb{1}_{\tau_A \geq t/2}] = 0$. Since the limit (29) holds uniformly for all $x \in A$, M is finite. Hence, $\lim_{t \rightarrow \infty} \Psi_1(r_*(t)) |\mathbb{E}_x [\bar{g}(X_t) \mathbb{1}_{\tau_A \leq t}]| = 0$ uniformly for all $g \in \mathcal{L}_{\Psi_2(f_*)}$.

Since $\pi(f_*) < \infty$, $|\mathbb{E}_x [\bar{g}(X_t) \mathbb{1}_{\tau_A \geq t}]| \leq c \mathbb{E}_x [\Psi_2(f_*(X_t)) \mathbb{1}_{\tau_A \geq t}]$. Following the same lines as in the proof of (Meyn and Tweedie, 1993c, Theorem 7.2.), using again $\sup_{u \leq m} P^u f_* \leq c f_*$, we obtain,

$$\Psi_1(r_*(t)) \mathbb{E}_x [\Psi_2(f_*(X_t)) \mathbb{1}_{\tau_A \geq t}] \leq c \Psi_1(r_*(m)) \inf_{0 \leq u \leq m} \Psi_1(r_*(t-u)) \mathbb{E}_x [\Psi_2(f_*(X_{t-u})) \mathbb{1}_{\tau_A \geq t-u}].$$

By Proposition 23, $G_A(x, \Psi_2(f_*), \Psi_1(r_*); 0) < \infty$, which implies that the upper limit in the right-hand side is zero, proving that $\lim_{t \rightarrow \infty} \Psi_1(r_*(t)) \mathbb{E}_x [\Psi_2(f_*(X_t)) \mathbb{1}_{\tau_A \geq t}] = 0$. Hence, uniformly for $g \in \mathcal{L}_{\Psi_2(f_*)}$, $\lim_{t \rightarrow \infty} \Psi_1(r_*(t)) |\mathbb{E}_x [\bar{g}(X_t) \mathbb{1}_{\tau_A \geq t}]| = 0$. This concludes the proof.

4.2. Proof of Corollary 6. Set $f_* := V^{1-\alpha}$ and $r_*(t) := (t+1)^{1/\alpha-1}$.

Lemma 25. *Suppose Assumption (ii) of Corollary 6. For any $\alpha \leq \eta \leq 1$, $t \geq 0$, and any \mathcal{F}_t -stopping-time τ ,*

$$c_\eta \mathbb{E}_x \left[\int_0^{\tau \wedge t} V^{\eta-\alpha}(X_s) ds \right] + \mathbb{E}_x [V^\eta(X_{\tau \wedge t})] \leq V^\eta(x) + b \mathbb{E}_x \left[\int_0^{\tau \wedge t} \mathbb{1}_C(X_s) ds \right].$$

Proof. By definition of \mathcal{AV} ,

$$c_\eta \mathbb{E}_x \left[\int_0^{\tau \wedge t \wedge T_n} V^{\eta-\alpha}(X_s) ds \right] + \mathbb{E}_x [V^\eta(X_{\tau \wedge t \wedge T_n})] \leq V^\eta(x) + b \mathbb{E}_x \left[\int_0^{\tau \wedge t \wedge T_n} \mathbb{1}_C(X_s) ds \right].$$

The right-hand side is upper bounded by $V(x) + bt$ and by the monotone convergence theorem, it converges to $V(x) + b \mathbb{E}_x \left[\int_0^{\tau \wedge t} \mathbb{1}_C(X_s) ds \right]$ as $n \rightarrow \infty$. The Lemma now results from Fatou's Lemma. \square

Proposition 26. *Suppose Assumption (ii) of Corollary 6. For all $\delta > 0$, there exists $c < \infty$ such that for all $x \in \mathcal{X}$, $G_C(x, \mathbb{I}, r_*; \delta) \leq cV(x)$.*

Proof. Set $q := \lfloor 1/\alpha \rfloor$ where $\lfloor \cdot \rfloor$ denotes the lower integer part. By Lemma 25, we have $\mathbb{E}_x[\tau_C] \leq cV^\alpha(x)$ and by Jensen's inequality, we obtain $\mathbb{E}_x[\tau_C^{\alpha^{-1}-q}] \leq cV^{1-q\alpha}(x)$. We prove by induction that for all integer $1 \leq l \leq q$, $\mathbb{E}_x[\tau_C^{\alpha^{-1}-l}] \leq cV^{1-l\alpha}(x)$. The case $l = q$ holds; assume it is verified for some $2 \leq l \leq q$. The induction hypothesis and Lemma 25 yield

$$\mathbb{E}_x \left[\tau_C^{\alpha^{-1}-l+1} \right] \leq c \mathbb{E}_x \left[\int_0^{\tau_C} \mathbb{E}_{X_s} \left[\tau_C^{\alpha^{-1}-l} \right] ds \right] \leq c \mathbb{E}_x \left[\int_0^{\tau_C} V^{1-l\alpha}(X_s) ds \right] \leq cV^{1-l\alpha+\alpha}(x),$$

which concludes the induction. For $l = 1$, this yields $G_C(x, \mathbb{I}, r_*; 0) \leq cV(x)$. Finally, by standard manipulations and Lemma 25, we have $G_C(x, \mathbb{I}, r_*; \delta) \leq c(1 + P^\delta V(x)) \leq cV(x)$. \square

[Proof of Corollary 6]. We check the conditions for the application of Theorem 1. Lemma 25 and Proposition 26 imply $G_C(x, f_*, \mathbb{I}; \delta) \leq cV(x)$ and $G_C(x, \mathbb{I}, r_*; \delta) \leq cV(x)$, from which we deduce the condition (ii) of Theorem 1, and by R1, condition (i) of Theorem 1. Condition (iii) follows from Lemma 25. Finally, $\mathcal{S}_\Psi = \mathcal{X}$.

ACKNOWLEDGEMENTS

We are grateful to the referee for his comments, and in particular for bringing Dai and Meyn (1995) to our attention.

REFERENCES

BARNDORFF-NIELSEN, O. and SHEPHARD, N. (2001). Non-gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. Roy. Stat. Soc., B* **63** 167–241.

- DAI, J. and MEYN, S. (1995). Stability and convergence of moments for multiclass queuing networks via fluid limit models. *IEEE Trans. Automat. Control* **40** 1889–1904.
- DAVIS, M. (1993). *Markov Models and Optimization*. London: Chapman and Hall.
- DOUC, R., FORT, G., MOULINES, E. and SOULIER, P. (2004). Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.* **14** 1353–1377.
- DOWN, N., MEYN, S. and TWEEDIE, R. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691.
- DYMKIN, E. (1965). *Markov Processes*. Springer ; Academic Press.
- FELLER, W. (1971). *An introduction to probability theory and its applications. Vol. II*. 2nd ed. John Wiley & Sons Inc., New York.
- FORT, G. and MOULINES, E. (2003). Polynomial ergodicity of markov transition kernels. *Stoc. Proc. Appl.* **103** 57–99.
- GANIDIS, H., ROYNETTE, B. and SIMONOT, F. (1999). Convergence rate of semi-groups to their invariant probability. *Stoc. Proc. Appl.* **79** 243–263.
- GETTOOR, R. (ed.) (1980). *Transience and recurrence of Markov processes*. 784, Lect. Notes Math. 397–409.
- HAS'MINSKII, R. (1960). Ergodic properties of recurrent diffusion processes and stabilization of the solution to the Cauchy problem for parabolic equations. *Theory Probab. Appl.* **5** 179–196.
- HAS'MINSKII, R. (1980). *Stochastic stability of differential equations*. Sijthoff and Noordhoff.
- JARNER, S. and ROBERTS, G. (2001). Convergence of heavy tailed MCMC algorithms. Tech. rep., Lancaster University. Available at [//www.statslab.cam.ac.uk/mcmc](http://www.statslab.cam.ac.uk/mcmc).
- JARNER, S. and ROBERTS, G. (2002). Polynomial convergence rates of Markov Chains. *Ann. Appl. Probab.* **12** 224–247.
- KENT, J. (1978). Time-reversible diffusions. *Adv. Appl. Probab.* **10** 819–835.
- KRASNOSEL'SKII, M. and RUTICKII, Y. (1961). *Convex functions and Orlicz spaces*. Noordhoff, Groningen.
- LAWLER, G. and SOKAL, A. (1988). Bounds on L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.* **309** 557–580.
- LUND, R., MEYN, S. and TWEEDIE, R. (1996). Computable exponential convergence rates for stochastically ordered Markov processes. *Ann. Appl. Probab.* **6** 218–237.
- MALYSHKIN, M. (2001). Subexponential estimates of the rate of convergence to the invariant measure for stochastic differential equations. *Theory Probab. Appl.* **45** 466–479.

- MEYN, S. and TWEEDIE, R. (1993a). *Generalized resolvents and Harris Recurrence of Markov processes*. American Mathematical Society, Providence, RI, 227–250.
- MEYN, S. P. and TWEEDIE, R. L. (1993b). *Markov chains and stochastic stability*. Springer-Verlag London Ltd., London.
- MEYN, S. P. and TWEEDIE, R. L. (1993c). Stability of markovian processes II: continuous-time processes and sampled chains. *Adv. Appl. Prob.* **25** 487–517.
- MEYN, S. P. and TWEEDIE, R. L. (1993d). Stability of markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Prob.* **25** 518–548.
- NUMMELIN, E. (1984). *General irreducible Markov chains and nonnegative operators*. Cambridge University Press, Cambridge.
- NUMMELIN, E. and TUOMINEN, P. (1983). The rate of convergence in Orey’s theorem for Harris recurrent Markov chains with applications to renewal theory. *Stoc. Proc. Appl.* **15** 295–311.
- ROBERTS, G. and ROSENTHAL, J. (1996). Quantitative bounds for convergence rates of continuous time Markov processes. *Electronic J. Probab.* **1** 1–21.
- ROBERTS, G. and STRAMER, O. (2003). Langevin diffusions and Metropolis-Hastings algorithms. To appear in *Methodology and Computing in Applied Probability*.
- ROBERTS, G. and TWEEDIE, R. (1996). Exponential convergence of Langevin diffusions and their discrete approximations. *Bernoulli* 341–364.
- ROBERTS, G. and TWEEDIE, R. (2000). Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Probab.* **37** 359–373.
- ROBERTS, G. and TWEEDIE, R. (2002). Understanding MCMC. To be published.
- SHARPE, M. (1988). *General theory of Markov Processes*. Academic Press.
- STRAMER, O. and TWEEDIE, R. (1999a). Langevin-type models I: Diffusions with given stationary distributions, and their discretizations. *Methodol. Comput. Appl. Probab.* **1** 283–306.
- STRAMER, O. and TWEEDIE, R. (1999b). Langevin-type models II: Self-targeting candidates for MCMC algorithms. *Methodol. Comput. Appl. Probab.* **1** 307–328.
- THORISSON, H. (1985). The queue GI/G/1: finite moments of the cycle variables and uniform rates of convergence. *Stoc. Proc. Appl.* **19** 83–99.
- TUOMINEN, P. and TWEEDIE, R. (1994). Subgeometric rates of convergence of f -ergodic Markov Chains. *Adv. in Appl. Probab.* **26** 775–798.
- VERETENNIKOV, A. (1999). On polynomial mixing and convergence rate for stochastic difference and differential equations. *Theory Probab. Appl.* **44** 361–374.