This is a repository copy of *On convergence of 1D Markov diffusions to heavy-tailed invariant density*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/138981/

Version: Accepted Version

Article:

On convergence of 1D Markov diffusions to heavy-tailed invariant density

O.A. Manita,[*] A.Yu. Veretennikov[†]

November 9, 2018

Abstract

Rate of convergence is studied for a diffusion process on the half line with a non-sticky reflection to a heavy-tailed 1D invariant distribution which density on the half line has a polynomial decay at infinity. Starting from a standard receipt which guarantees some polynomial convergence, it is shown how to construct a new non-degenerate diffusion process on the half line which converges to the same invariant measure exponentially fast uniformly with respect to the initial data.

Key words: 1D diffusion; invariant distribution; heavy tails; fast convergence
MSC codes: 60H10, 60J60.

1   Introduction

A topical area of Markov Chain Monte Carlo (MCMC) in theoretical statistics is around the following problem: given a fixed "target" density or distribution known up to a constant multiplier – a normalizing constant – how to construct a (Markov) process which would have this density as a (unique) invariant one and which would converge to this invariant one with a rate that could be theoretically evaluated? In particular, a permanent great interest in recent decades was about dealing with "heavy-tailed" densities with a polynomial decay at infinity. With this problem in mind, let us consider a polynomially decreasing probability density π on the line $\mathbb{R}^1$; in the precise setting it will be restricted to the half-line $\mathbb{R}^1_+$. The question under consideration in this paper is constructing a Markov diffusion process with invariant measure $\pi(x)dx$ such that this measure is invariant for the constructed process and,

---

[*]Moscow State University, Moscow, Russia; email: oxana.manita @ gmail.com

moreover, so that an exponential convergence in total variation to the invariant distribution holds.

This problem has certain deep relations to ergodicity and to the Perron – Frobenius theorem for Markov chains with finite state space, to spectral gap for semigroup generators, to upper and lower bounds for convergence to stationarity; yet, a spectral gap in this paper is not used. The literature in this area is huge and we only mention a few important references related to the subject of the paper more or less directly (see [3], [7], [9], [17], [19], [20], [24], [25], et al.; also, see further references theiren).

The paper consists of four sections, the first one being this Introduction. In the section 2 two known receipts of constructing an SDE with a given stationary measure are shown: one is an SDE with a unit diffusion coefficient while another one is an SDE with an affine drift. In the section 3 we state the main result of this paper, and in the section 4 its proof is provided. The construction is based on the first one of the standard receipts from the section 2 and on a random time change. The proof uses certain recurrence type hitting time moment bounds introduced earlier in [28].

2 Quick review: two standard receipts on $\mathbb{R}^1$

2.1 Receipt 1: SDE with a unit diffusion coefficient

Suppose a continuous and differentiable strictly positive probability density π on $\mathbb{R}^1$ decreases at infinity polynomially, i.e. there exist constants $c > 0$ and $m > 1$ such that for any x,

$$c \left(1 + |x|\right)^{-m} \leq \pi \left(x\right) \leq c^{-1} \left(1 + |x|\right)^{-m}. \tag{1}$$

Here $m > 1$ is required so that the function π were integrable; for further claims a bit more restrictive condition $m > 3$ will be assumed in the sequel.

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let us fix a standard Wiener process W_t with its natural filtration $\mathcal{F}_t = \mathcal{F}_t^W$ (as usual, $\mathbb{P}$ – completed). On this probability space consider a Langevin diffusion Y_t given by an SDE

$$dY_t = dW_t + b\left(Y_t\right) dt, \quad Y_0 = \xi, \tag{2}$$

with an arbitrary nonrandom initial value ξ, where

$$b(x) = \frac{1}{2}(\ln \pi(x))'. \tag{3}$$

If there is no explosion then this equation possesses a strong solution [27]. Random initial values will be mentioned briefly in the section 2.2 and in principle could have been allowed here, too. It is assumed that two derivatives π' and π'' exist and that the drift b is locally bounded; its global boundedness is not required because, as it turns out [26, 28], a no blow-up is guaranteed just by the assumption (1) on the function π only (see below the details). Emphasize that despite the assumed two derivatives, the only quantitative assumption will be just on π itself given in (1); also there is a hypothesis that the assumption about π'' could be dropped, and this is the reason why we refer to [27] instead of more standard results under a local Lipschitz condition on the drift, while talking on strong solutions in the sequel.

The receipt (2)–(3) is, actually, a continuous time analogue of (one of) a standard MCMC receipt(s) in discrete time after a suitable limit. We do not recall it because the paper does not rely upon this limiting procedure; however, this is likely to signify a possible link to MCMC algorithms in discrete time. Obviously $\pi(x)dx$ is the (unique) invariant distribution of the process Y_t; this can be shown explicitly by checking the Kolmogorov equation for the invariant distribution. The process Y_t is ergodic and has a polynomial rate of convergence to the stationary distribution with density π [26, Theorem 1] (under a bit more restricted conditions see also [28]), at least, if m is not too small. Note that unlike in most of other works on convergence or mixing rates, Lyapunov functions are not used here, as they were not used in [26, 28]. For close results for some particular distributions and for homogeneous Markov processes under various assumptions (usually more restrictive because of explicit assumptions about the derivative π') see also [1, 2, 19]; for discrete time examples – that is, actually, about MCMC algorithms – see, e.g., [8, 9, 25] and further references therein. Emphasize that the assumptions in [26, Theorem 1] as well as in [28, Theorem 1] are all on π and not on π' except that the latter derivative exists and that b is locally bounded. It remains to be our goal to avoid any assumptions on π' beyond its existence and local boundedness of b in the sequel.

Moreover, it is known that under the *additional assumption* about π',

$$\liminf_{|x|\to\infty} xb(x) = \liminf_{|x|\to\infty} \frac{x\pi'(x)}{2\pi(x)} = -r < -3/2, \tag{4}$$

the *beta-mixing rate* of Y_t is *no faster than polynomial*, $\geq Ct^{-k}$ *with any* $k > r - 1/2$ *and some* $C > 0$ (see the definition and the details in [29]). The notion of beta-mixing – which is neither defined nor discussed here in detail – is rather close although not identical to the convergence in total variation. Hence, and also because of close results about lower bounds for convergence rates in [15, 24], it is likely that convergence of Y to the stationary distribution of Y_t under (1) is also no faster than some polynomial. The assumption (4) will not be used in the sequel but was shown just for information. Recall that our aim is a faster convergence, and that we want to avoid any conditions on the derivative π' except for its existence and local boundedness.

Note that in the case if for large $|x|$ the density equals *exaclty* $c(1+|x|)^{-m}$, it apparently follows that we need $m > 3$ in order to have the inequality $r > 3/2$ in (4). Yet, we will not use conditions in terms of π', assuming just (1). Also, emphasize that the requirement $m > 3$ is for the quick reference on some existing earlier results. We do not claim that for $m \leq 3$ a similar analysis and asymptotics are not possible, but just that we are not aware of such asymptotics for $m \leq 3$.

Note that for the density on a half-line $\mathbb{R}^1_+$ a natural analogue of (2) is the process satisfying an SDE with a non-sticky reflection at zero,

$$dY_t = dW_t + b(Y_t)\,dt + d\phi_t^Y, \quad Y_0 = \xi. \tag{5}$$

For the process satisfying (5) similar mixing and convergence bounds follow from the bounds and from the calculus quite similar to those in [28] applied to the situation of the reflected SDE, or just from a consideration of an SDE (2) with a symmetric ($b(-x) = b(x)$) drift.

2.2 Receipt 2: SDE with an affine "mean-reverted" drift

There is another receipt different from (2) offered in [4] (see also the references therein concerning some other earlier constructions): in a slightly simplified form it suggests to consider an SDE on the line

$$dZ_t = -(Z_t - \mu) \, dt + \sqrt{v(Z_t)} \, dW_t, \tag{6}$$

with an appropriate initial distribution (e.g., stationary π as in the reference paper), with

$$v(z) = \pi(z)^{-1} \int_{-\infty}^{z} (\mu - s) \, \pi(s) \, ds, \quad \mu = \int s \, \pi(s) \, ds. \tag{7}$$

It is, of course, assumed that μ is finite, and then it is easily proved that $v \geq 0$, so that the SDE (6) is well-defined. (Indeed, $\int_{-\infty}^{\mu} (\mu - s) \, \pi(s) \, ds \geq 0$ since $\mu - s \geq 0$ for $s \leq \mu$; and for $s > \mu$ the values $\mu - s$ are negative but the whole integral $\int_{-\infty}^{+\infty} (\mu - s) \, \pi(s) \, ds = 0$, so that for any $z < +\infty$, $v(z) \geq 0$ as required.) However, of course, "good" ergodic properties of the solution of this equation depend on some features of the density π. The solution locally exists due to local Lipschitz property of $\sqrt{v}$ combined with the affine drift assumption, but no-explosion should be derived from other conditions. Some related papers are, for example, [1, 2, 19] which tackle particular parametric families of target densities π – Student, reciprocal Gamma, and Fisher-Snedekor diffusions. In all three papers a quadratic Lyapunov function allows to show an exponential convergence in total variation which is non-uniform in the initial state Z_0. It is interesting that in [4] an exponential character of the stationary correlation function is established; yet, convergence of a *non*-stationary process to a stationary regime was not studied. In fact, the process under investigation in [4] is stated to be "ergodic", which ergodicity is understood in the sense of being stationary without any convergence statements. At the same time, the assumptions on the density in [4] involve the stationary density π (in our notations) but not on its derivative (possibly with some additional non-restrictive requirements in some theorems like continuity of the target density). Recall that in the present paper C^2-differentiability of π is assumed, but convergence rate bounds only depend on the asymptotic assumptions at infinity on the density π itself. It looks plausible that, in principle, it may be possible to work with "weak" definitions of the process via Dirichlet forms theory ([10, 21]), but we prefer to have a well-defined solution trajectory; in particular, we will be working with strong solutions due to [27]. The receipt 2 naturally rises the question whether it is possible to arrange even a faster convergence, let theoretically.

3 The setting & main result

Our primary goal is a density π on $\mathbb{R}_+^1 = [0, \infty)$ satisfying

$$c \, (1 + x)^{-m} \leq \pi(x) \leq c^{-1} \, (1 + x)^{-m}, \quad x \geq 0. \tag{8}$$

Receipts I & II in the previous section can be applied to this setting if we just extend the density in a symmetric way to the whole line, with a natural normalisation. As was said in

the Introduction, we aim at constructing a diffusion process on $\mathbb{R}_+^1$ which converges towards π with an exponential rate uniformly with respect to the initial data. It is likely that a similar result holds true for a *symmetric* density π on the whole line $\mathbb{R}^1$ satisfying (1), which we mention as a remark.

In order to achieve yet a better convergence than typically guaranteed by the receipts (2)–(3) or even by (6)–(7), and for yet a more general class of densities than in [1, 4, 19] and in quite a few other works, let us consider two diffusion processes Y_t and X_t on $\mathbb{R}_+$ satisfying, respectively, SDEs with a non-sticky reflection at zero (5) and

$$dX_t = f(X_t)\,dW_t + f^2(X_t)\,b(X_t)\,dt + d\phi_t^X, \quad X_0 = \xi, \tag{9}$$

with a local time ϕ_t^X at zero and with a special auxiliary function f,

$$f(z) := \left(1 + \int_0^z \frac{dy}{\pi(y)}\right)^{1/2}, \quad z \geq 0. \tag{10}$$

The generator of this process is given by

$$L = f^2 L_0,$$

where L_0 is the generator of the reflected diffusion (5):

$$L_0 v(x) = \frac{1}{2}v''(x) + b(x)v'(x), \ \forall\, x > 0, \quad \& \quad L_0 v(0) = v'(0+).$$

Recall the requirements on the non-sticky solution and on its local time: ϕ^X is a monotonically non-decreasing function; for any $t > 0$,

$$\phi_t^X = \int_0^t 1(X_s = 0)d\phi_s^X; \quad \int_0^t 1(X_s = 0)ds = 0 \text{ a.s.}$$

Of course, a question about existence of solution of this equation (9) on the whole half-line $t \geq 0$ arises here, and a positive answer to this question for the first sight may look doubtful given fast increasing coefficients. However, it will be justified with the help of a random time change and of the law of large numbers that such a (strong) solution exists on the whole line and does not explode. The main result is the following Theorem.

Theorem 1. *Assume that for a strictly positive probability density $\pi \in C^2$ with two locally bounded derivatives the bounds (8) hold with some $m > 3$. Then the SDE (9) has a strong solution X_t for all $t \geq 0$ which is strongly (pathwise) unique and which possesses an exponential rate of convergence to the stationary distribution $\pi(x)\,dx$,*

$$\|\mu_t^\xi - \mu\|_{TV} \leq C \exp(-\lambda t), \quad t \geq 0, \tag{11}$$

uniformly with respect to ξ, with some constants λ and C which both admit certain evaluation, where μ_t^ξ is a marginal measure of the process X_t that starts from ξ at $t = 0$, and $\mu(dx) = \pi(x)dx$ is the (unique) invariant measure of the process.

The right hand side in (11) does not depend on the initial value ξ. Theoretical evaluations of both constants in the bound (11) is likely to be not very efficient, yet possible which is clearly better than pure existence of such constants.

5

4 Proof of Theorem 1

The proof will be split into several steps.

Step 1. Random change of time. Define the function $f(z)$ on $\mathbb{R}^1_+$ by (10). Obviously there exists $0 < a \leq 1$ (namely, any $a \in [0, c^2]$, with c from (8)) such that

$$a\,(1+z)^{m+1} \leq f^2(z) \leq a^{-1}\,(1+z)^{m+1}, \quad \forall z \in \mathbb{R}^1_+. \tag{12}$$

Let us define a random time change (cf. [11, 23]) by

$$\chi_t := \int_0^t f^{-2}\,(Y_s)\,ds, \quad \& \quad \beta_t := \chi_t^{-1} \quad \text{(the inverse function)}. \tag{13}$$

In other words,

$$\beta_t' = f^2(Y_{\beta_t}),$$

and

$$t = \int_0^{\beta_t} f^{-2}\,(Y_s)\,ds.$$

This time change $t \mapsto \beta_t$ is non-degenerate, that is, the following two conditions hold:
(i) there is no blow up at finite time:

$$\mathbb{P}(\chi_t|_{t \to T-0} \to +\infty) = 0 \quad \forall T \in (0, +\infty). \tag{14}$$

(ii) χ_t is unbounded as $t \to +\infty$ (i.e. when "real" time goes to infinity):

$$\chi_t \geq 0, \quad \chi_t|_{t \to \infty} \to +\infty \quad \mathbb{P} - \text{a.s.} \tag{15}$$

To prove (14), it suffices to notice that, due to (12), for any $s < t$,

$$0 \leq \chi_t - \chi_s \leq a^{-1} \int_s^t (1 + |Y_r|)^{-m-1}\,dr,$$

hence

$$\chi_t' \leq a^{-1} \cdot \sup_{r \in \mathbb{R}^1} (1 + |r|)^{-m-1} = a^{-1}, \quad \mathbb{P} - \text{a.s.}$$

Then (14) immediately follows.

From here we find,

$$\inf_{t \geq 0} \beta_t' \geq a > 0, \quad \mathbb{P} - \text{a.s.},$$

and

$$\mathbb{P}(\limsup_{t \to \infty} \beta_t < \infty) = 0.$$

The assertion (15) follows from the following Lemma.

Lemma 1. *Let $m > 3$, and let g be a bounded continuous function on $\mathbb{R}^1$. Assume that the diffusion process Y_t satisfies (2), and let μ_{inv} be its unique invariant measure. Then for any $\delta > 0$ and $\varepsilon > 0$ there exists $T_0 > 0$ such that*

$$\mathbb{P}\left(\left| \frac{1}{t} \int_0^t g\left(Y_s\right) ds - \int g(x) d\mu_{inv}(x) \right| > \varepsilon \right) < \delta \qquad \text{for any } t \geq T_0.$$

The Lemma with $g(r) = (1 + |r|)^{-m-1}$ yields the assertion (15). Indeed, let us fix any $\delta \in (0,1)$ and $\varepsilon = a_g/2$. Naturally, $a_g > 0$. Then with $\mathbb{P}$-probability at least $1 - \delta$ one has $\chi_t \geq (a_g - \varepsilon) t = a_g t/2$ for **any** t large enough. This means that with probability at least $1 - \delta$ the change of time mapping does not stop up to at least $a_g t/2$. Since $\delta \in (0,1)$ is arbitrary, (15) holds.

Proof of Lemma 1. First of all, we will refer to the mixing results for SDEs on the whole line; however, in the case of symmetric coefficients ($b(-x) = b(x)$ and similarly for the diffusion if it is not a constant) such results straightforward imply similar bounds and convergence rates for (non-sticky) reflected at zero diffusions, too. In other words, The process (Y_t) is Markov ergodic with a finite variance (and, in fact, with any moment $m' < m - 2$; g is bounded) with a polynomial beta-mixing rate as well as convergence in total variation $\beta^\xi(t) + |\mu_t^\xi - \mu_\infty|_{TV} \leq C_k(\xi)(1 + t)^{-k}$ with some $C(\xi)$ for any $k < m - 1$, to the stationary regime μ_∞, see [28, 26]. Indeed, the assumptions of [28] are met with $p = m - 1$ where p is the standing parameter in [28]. The assumption $m > 3$ implies $k > 2$. Moreover, the function g is bounded; hence, the process $\int_0^t g(Y_s) ds$ possesses all moments (including exponential with any constant, although, this is far too much for our goal). The beta-mixing coefficient dominates the alpha-mixing, while certain convergence rate to zero of the alpha coefficient is the standing assumption in the Theorem 18.5.4 of [13]. Hence, for the stationary regime, the assertion of the Lemma – LLN – follows from the Central Limit Theorem [13, Theorem 18.5.4]. Indeed, splitting the integral from zero to t into a sum $\sum_1^{[t]}$ plus $\int_{[t]}^t$, the claim follows. For a *nonstationary* regime the desired LLN follows again from the CLT for the *stationary* case, from the Markov property, and from the polynomial convergence of Law (Y_s) to μ_{inv} in total variation, similarly to the proof of " non-stationary CLT" in [31, Theorem 4] with the help of the results from [28, Theorem 1] with $m > 3$. After mixing bounds have been found, see also [32] for LLN (formally, in [32] mixing is exponential, but obviously any polynomial would do such that the related sums or integrals converge). This finishes the proof of the Lemma 1. $\qquad\square$

See also [20] for close results under slightly different assumptions. As was already mentioned, the statement of the Lemma will be used straight away for our reflected diffusion (5).

Step 2. Constructing the process X_t.

On the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with a solution Y_t to the equation (5), let us introduce stochastic processes

$$X_t := Y_{\beta_t}, \quad \phi_t^X = \phi_{\beta_t}^Y.$$

Then due to the time change [11, Theorem 3.15.5] it follows that the process X_t satisfies an SDE

$$dX_t = f\left(X_t\right) dW_t + f^2\left(X_t\right) b\left(X_t\right) dt + d\phi_t^X, \quad X_0 = \xi, \qquad (16)$$

with a new Wiener process $\tilde{W}_t = \int_0^{\beta_t} f^{-1}(X_s) \, dW_s$, and with the local time at zero ϕ_t^X; recall that $f(0) = 1$. Indeed, outside zero the "main part" here

$$1(X_t > 0)dX_t = 1(X_t > 0)\left[f(X_t) \, d\tilde{W}_t + f^2(X_t) b(X_t) \, dt \right]$$

follows straightforward from [11, Theorem 3.15.5], and

$$1(X_t = 0)dX_t = 1(X_t = 0)d\phi_t^X$$

is a direct consequence of the equation

$$1(Y_t = 0)dY_t = 1(Y_t = 0)d\phi_t^Y.$$

Also, we have,

$$\int_0^t 1(X_s = 0)ds = 0, \quad \int_0^t 1(X_s = 0)d\phi_s^X = \phi_t^X.$$

Finally,

$$X_t - \xi - \int_0^t f(X_s) \, dW_s + \int_0^t f^2(X_s) b(X_s) \, ds - \phi_t^X = 0, \quad \text{a.s.}$$

Thus, X is the solution of the equation (16) with a non-sticky reflection, as required.

The equation (16) can be also derived from the time change for the SDE (2) on the whole line with a symmetric drift and symmetrically extended f after the application of Itô–Tanaka's formula to the modulus,

$$d\bar{X}_t = \bar{f}\left(\bar{X}_t\right) \, d\tilde{W}_t + \bar{f}^2\left(\bar{X}_t\right) b\left(\bar{X}_t\right) dt, \quad \bar{X}_0 = \xi, \tag{17}$$

with

$$\bar{b}(x) = \text{sign}(x)b(|x|), \quad \bar{f}(x) = f(|x|), \quad \forall\, x \in \mathbb{R}^1.$$

By construction, the processes X_t and ϕ_t^X are regular, i.e. are defined for all $t \geq 0$, and adapted to the filtration $\tilde{\mathcal{F}}_t \equiv \mathcal{F}_{\beta_t}$, see [11]. Recall the well-known fact that the new filtration $\mathcal{F}_{\beta_t}$ is well-defined because of the fact that for any t, the random variable β_t is a stopping time.

Emphasize that the process X_t is well defined on the whole half-line $t \geq 0$, it does not explode, and it neither reaches infinity from zero, nor vice versa (zero from infinity) over a finite time, all of these because of the construction via the time change.

Step 3. The solution X_t is strong. Indeed, it is well-defined on $t \geq 0$, and the diffusion coefficient is locally continuously differentiable, and locally bounded, and locally non-degenerate, while the drift coefficient is also locally bounded. Due to the results in [27], this suffices for strong uniqueness via the stopping time arguments with the help of the strong Markov property – see [18]. This will be used in the sequel in the coupling procedure (although, probably could be done with weak solutions, too).

8

Step 4. Stationary distribution for X_t.

Let us prove that the process X_t has a unique invariant distribution $\pi(x)dx$. The stationary distribution μ satisfies the stationary Kolmogorov equation $L^*(\mu) = 0$ on $\mathbb{R}_+$ – or, equivalently, $L^*\pi = 0$ – where

$$L = \frac{f^2}{2}D_x^2 + (f^2 b)D_x \tag{18}$$

is the generator of X_t and * is the adjoint with respect to the Lebesgue measure. First of all, the Kolmogorov equation $L^*(\mu) = 0$ has at most one probability solution due to [6, Example 4.1.1]. Next, the measure $\mu(dx) = \pi(x)dx$ satisfies this equation. Indeed, since

$$\frac{1}{2}\pi' - (b\pi) = 0, \tag{19}$$

we have

$$\frac{1}{2}\left((f^2)\pi\right)'' - \left((f^2)b\pi\right)' = \frac{1}{2}\left((f^2)''\pi + 2(f^2)'\pi' + (f^2)\pi''\right) - (f^2)'(b\pi) - (f^2)(b\pi)' =$$

$$(f^2)\left(\frac{1}{2}\pi'' - (b\pi)'\right) + (f^2)'\pi' + \frac{1}{2}(f^2)''\pi - (f^2)'(b\pi) \overset{(19)}{=} \frac{1}{2}(f^2)''\pi + \frac{1}{2}(f^2)'\pi' = \frac{1}{2}\left((f^2)'\pi\right)'. \tag{20}$$

But

$$((f^2)'\pi)(x) = \left(1 + \int_0^x \frac{dy}{\pi(y)}\right)' \cdot \pi(x) = \frac{\pi(x)}{\pi(x)} = 1, \quad x \geq 0, \tag{21}$$

where at zero derivative is understood as right one. Hence the expression in the right hand side of (20) equals zero, i.e., π is a stationary measure for the new process X. Note that the same calculus with f replaced by 1 shows the invariance of π for Y_t

As may be expected, (21) implies the equality

$$\mathbb{E}_\pi h(X_t) = \int h(y)\pi(y)\, dy, \tag{22}$$

for any $t > 0$ and $h \in C_b(\mathbb{R}_+^1)$. By virtue of the Lebesgue dominated convergence theorem we can take $h \in C_0^\infty(\mathbb{R}_+^1)$ (continuous with a compact support), but more than that, it suffices to consider functions $h \in C_0^\infty(\mathbb{R}_+^1)$ with $h'(0+) = 0$: indeed, the latter class – denoted in the sequel as $C_{00}^\infty(\mathbb{R}_+^1)$ – is clearly dense in $C_0^\infty(\mathbb{R}_+^1)$. We have, due to $h'(0+) = 0$,

$$dh(X_t) = Lh(X_t)\, dt + h'(X_t)f(X_t)\, dW_t + 1(X_t = 0)h'(0+)d\phi_t^X$$

$$= Lh(X_t)\, dt + h'(X_t)f(X_t)\, dW_t.$$

Moreover, since h has a compact support, Lh and $h'f$ are bounded. So, by rewriting in integral form and taking expectations we get,

$$\mathbb{E}_\pi h(X_t) - \mathbb{E}_\pi h(X_0) = \mathbb{E}_\pi \int_0^t Lh(X_s)\, ds \equiv \int_0^t \mathbb{E}_\pi Lh(X_s)\, ds,$$

9

the last equality due to Fubini's theorem. Denote by $p_s(y, z)$ the transition density of the Markov process X; its existence follows, e.g., from [5, Corollary 2.9 & Remark 2.17]. In a "good case" with all appropriate derivatives, the standard *formal* calculus runs as follows:

$$\mathbb{E}_\pi L h(X_s) = \iint (L_z h(z)) \pi(y) p_s(y, z) \, dz dy$$

$$= \iint \pi(y) h(z) L_z^* p_s(y, z) dz dy = \iint \pi(y) h(z) \partial_s p_s(y, z) dz dy$$

$$= - \iint \pi(y) h(z) L_y p_s(y, z) dz dy = - \int h(z) \left(\int p_s(y, z) L_y^* \pi(y) dy \right) dz = 0,$$

due to forward and backward Kolmogorov's equations, and Fubini's theorem. Therefore, we conclude that

$$\mathbb{E}_\pi h(X_t) = \mathbb{E}_\pi h(X_0), \tag{23}$$

which is equivalent to (22). A rigorous justification without additional assumptions follows from [5].

Note that uniqueness of the invariant measure will follow from the convergence bound (11) once it is established.

Step 5. Uniform exponential moment bound.

Let us take *any* $K > 0$, and define

$$\gamma_X^\xi \equiv \gamma_X^{\xi, K} \equiv \gamma := \inf\left(t \geq 0 : X_t \leq K, X_0 = \xi \geq 0\right). \tag{24}$$

Let us show that $\mathbb{E} \exp(\alpha \gamma_X^\xi) < +\infty$ for $\alpha > 0$ small enough, uniformly with respect to the initial state of the process. Denote $v_q(\xi) := \mathbb{E}_\xi \gamma^q$, with a convention $v_0 \equiv 1$. Obviously,

$$\mathbb{E}_\xi e^{\alpha \gamma} = \sum_{q=0}^{+\infty} \frac{\alpha^q \mathbb{E}_\xi \gamma^q}{q!} = \sum_{q=0}^{+\infty} \frac{\alpha^q v_q(\xi)}{q!} < \infty,$$

provided all quantities $v_q(\xi)$ are finite and grow not too fast in q.

Step 6. Auxiliary results for polynomial moments.

In order to guarantee that the values v_q, indeed, may not grow too fast, let us find alternative representations for them. By virtue of the identity

$$\left(\int_0^\gamma 1 dt \right)^q = q \int_0^\gamma \left(\int_t^\gamma 1 ds \right)^{q-1} dt,$$

which holds both for finite and infinite γ, we get,

$$v_q(\xi) = q \mathbb{E}_\xi \int_0^\gamma v_{q-1}(X_t) dt$$

10

at least, for any $q \geq 1$ such that $v_q(\xi)$ is finite. Indeed, due to the Fubini's theorem (iii) and the Markov property (iv),

$$v_q(\xi) \equiv \mathbb{E}_\xi \gamma^q = q\mathbb{E}_\xi \int_0^\gamma \left(\int_t^\gamma 1 ds \right)^{q-1} dt = q\mathbb{E}_\xi \int_0^\infty 1 \, (\gamma > t) \left(\int_t^\gamma 1 ds \right)^{q-1} dt =$$

$$\overset{(iii)}{=} q \int_0^\infty \mathbb{E}_\xi 1 \, (\gamma > t) \left(\int_t^\gamma 1 ds \right)^{q-1} dt = q \int_0^\infty \mathbb{E}_\xi 1 \, (\gamma \geq t) \, \mathbb{E}_\xi \left(\left(\int_t^\gamma 1 ds \right)^{q-1} |\mathcal{F}_t^X \right) dt =$$

$$= q \int_0^\infty \mathbb{E}_\xi 1 \, (\gamma > t) \, \mathbb{E}_\xi \left((\gamma - t)^{q-1} |X_t \right) dt = q \int_0^\infty \mathbb{E}_\xi 1 \, (\gamma > t) \, \mathbb{E}_{X_t} \gamma^{q-1} dt =$$

$$\overset{(iv)}{=} q \int_0^\infty \mathbb{E}_\xi 1 \, (\gamma > t) \, v_{q-1}(X_t) dt \overset{(iii)}{=} q\mathbb{E}_\xi \int_0^\gamma v_{q-1}(X_t) dt.$$

Hence, if both quantities are finite, then

$$v_q(\xi) = q\mathbb{E}_\xi \int_0^\gamma v_{q-1}(X_t) dt, \quad v_0 = 1.$$

Note that for each $q \geq 1$, if v_q is finite, then it satisfies

$$Lv_q(x) = -qv_{q-1}(x), \quad x \geq K, \tag{25}$$

by virtue of the probabilistic representation of solutions of the elliptic equation with Dirichlet boundary condition, or equivalently by Duhamel's formula. Obviously $v_q(K) = 0$, as well as $v_q(x) = 0$, $0 \leq x \leq K$. Also, it is known that if $v_q(\xi) < \infty$ for some ξ then it is finite for any ξ. However, we are not going to use this equation directly since it lacks the "second boundary condition" normally required for the second order differential equation. Instead, we will find solutions to boundary problems that approximate v_q. In fact, what we shall need instead is the following Lemma.

Let $\hat{L} : (\hat{L}u)(x) = a(x)u''(x) + c(x)u'(x)$ be the generator of the diffusion process $(\zeta_t, t \geq 0)$ with locally bounded coefficients $a > 0$ (the diffusion) and c (the drift), and such that a is locally uniformly non-degenerate, and which process is a strong solution of a corresponding SDE. Let us fix a positive (non-negative) function ψ on $\mathbb{R}^1$. Let

$$\tau_K := \inf(t \geq 0 : \zeta_t \leq K, \ \zeta_0 = \xi) \quad (\xi > 0),$$

and

$$v(\xi) = \mathbb{E}_\xi \int_0^{\tau_K} \psi(\zeta_t) \, dt.$$

Lemma 2. *For any $N > K > 0$, consider the boundary problem*

$$\hat{L}v_N^+ = -\psi, \quad v_N^+(K) = 0, \ \left(v_N^+\right)'(N) = 0, \tag{26}$$

Then the function $v_N^+(\xi) \uparrow v(\xi)$ as $N \uparrow \infty$, for every $\zeta_0 = \xi$ with $\xi \geq K$.

Proof of Lemma 2. For any $0 \leq K \leq \xi \leq N$, let us consider a family of stochastic processes ζ_t^N, given by the SDE with reflection,

$$d\zeta_t^N = \sqrt{2a(\zeta_t^N)}dw_t + c(\zeta_t^N)dt + d\phi_t^N, \quad \zeta_0^N = \xi,$$

with values on $[0, N]$, with a non-sticky reflection at N and an absorbtion at zero, where ϕ_t^N is its local time at N. Applying Itô's formula (or, in fact, more precisely Itô–Krylov's formula if continuity of ψ, a, and c is not assumed) to $v_N(\zeta_t^N)$, we get the following representations:

$$v_N(x) = \mathbb{E}_x \int_0^{\tau_{K,N}} \psi(\zeta_s^N)ds,$$

where $\tau_{K,N} = \inf\left(t \geq 0: |\zeta_t^N| \leq K\right)$ is the moment when the process ζ_t^N first hits the interval $[0, K]$. Note that $\tau_{K,N}$ monotonically increases as N increases. Also note that if $\tau_K < \infty$ then, obviously, $\tau_{K,N} \uparrow \tau_K$; and if $\tau_K = \infty$ then still $\tau_{K,N} \uparrow \infty = \tau_K$. These all follow from the comparison theorem for one-dimensional SDEs possessing strong solutions with the same coefficients and different initial data. This comparison theorem can be shown as follows. Consider two SDEs with the same initial value ξ but with two different $N_1 < N_2$, say. Denote the corresponding solutions by $\zeta_t^{N_1}$ and $\zeta_t^{N_2}$. Assuming that $\xi \in [K, N_1]$, due to the strong uniqueness $\zeta_t^{N_1} = \zeta_t^{N_2}$ until $\hat{\tau} := \inf(t \geq 0: \zeta_t^{N_1} = K$ or $N_1)$. If at this moment – which is a stopping time – $\zeta_t^{N_1} = \zeta_t^{N_2} = K$, then the claim is justified because K is the absorbtion point. If, however, $\zeta_t^{N_1} = \zeta_t^{N_2} = N_1$, then the first process $\zeta_t^{N_1}$ will remain less than or equal to N_1 all the time, while the second will exceed this level N_1 with probability one on any right interval of $\hat{\tau}$. This follows easily from the "reverse" time change which makes diffusion back equal to one and from the Girsanov theorem about eliminating the drift via change of measure, because for the standard Wiener process this property is well-known (e.g., it follows from Khintchin's iterated logarithm law for WP [14] along with the strong Markov property. Thus, on any small right neighbourhood of the moment $\hat{\tau}$ we would have $\zeta^{N_1} \leq \zeta^{N_2}$, with strict inequality at least at infinitely many moments of time arbitrarily close to $\hat{\tau}$. Yet, both solutions are strong Markov. So, if we now start two processes with the same generator a new at two distinct initial value $\xi_1 < \xi_2$, then due to continuity the two solutions will satisfy $1(t > \hat{\tau})1(\zeta_t^{N_1} < \zeta_t^{N_2}) = 1(t > \hat{\tau})$, at least, until they meet again, i.e., for all $t < \bar{\tau} := \inf(s \geq \hat{\tau}: \zeta_s^{N_1} = \zeta_s^{N_2})$ (here, of course, $\inf(\emptyset) = \infty$, and, in fact, they will never meet again). But then, if we assume that $\bar{\tau} < \infty$, they will again coincide until the next moment when they touch the level N_1, after which we have again $\zeta^{N_1} \leq \zeta^{N_2}$, and the cycle can repeated indefinitely. This shows that $\zeta_t^{N_1} \leq \zeta_t^{N_2}$ for all $t \geq 0$. Hence, we have $\tau_{K,N} \uparrow \tau_K$, $N \uparrow \infty$, and so, the monotonic convergence Theorem yields the assertion of the Lemma, as required. $\square$

Similar calculi in similar situations yielding various close claims can be found in [22, 26, 28].

Step 7. Bounds for polynomial moments.

Let us prove that

$$v_q(\xi) \leq q! \cdot C^q \tag{27}$$

for all $q \geq 1$, with

$$C = \frac{a^{-1}}{m} \cdot A_m, \quad A_m := \int_K^\infty (1+w)^{-m} dw.$$

does not depend on q. We argue by induction. Now our particular generator is $\hat{L} = L$ from (18).

Base: Let $q = 1$. Fix $N > 0$. Notice that $v^0 = 1$ and consider a boundary value problem,

$$Lv_N^1 = -1, \quad v_N^1(K) = 0, \quad \left(v_N^1\right)'(N) = 0. \tag{28}$$

Since $L = f^2 L_0$, where $L_0 u = \frac{1}{2}u'' + \frac{1}{2}\nabla \ln \pi(x)u'$, this problem admits a unique solution

$$v_N^1(\xi) = 2 \int_K^\xi \pi^{-1}(w_1)dw_1 \int_{w_1}^N \frac{\pi(w_2)}{f^2(w_2)}dw_2, \tag{29}$$

and due to (1) and (12) we estimate replacing N by infinity in the upper limit of the integral,

$$v_N^1(\xi) \leq 2c \int_K^\xi (1+w_1)^m dw_1 \int_{w_1}^\infty c^{-1} \cdot (1+w_2)^{-m} \cdot a^{-1}(1+w_2)^{-m-1} dw_2 \leq$$

$$\leq 2a^{-1} \int_K^\xi (1+w_1)^m dw_1 \int_{w_1}^\infty (1+w_2)^{-m}(1+w_2)^{-m-1} dw_2 =$$

$$= 2a^{-1} \int_K^\xi (1+w_1)^m \frac{(1+w_1)^{-2m}}{2m} dw_1 \leq \frac{a^{-1}}{m} \int_K^\infty (1+w_1)^{-m} dw_1 = A_m \cdot \frac{a^{-1}}{m} =: C.$$

By virtue of Lemma 2, this implies $v_1(\xi) = \lim_{N\to\infty} v_N^1(\xi) \leq C$.

Induction Step: Note that if the right hand side in the equation (28) is multiplied by a constant $R > 0$, then, given the specific boundary conditions, the bound for the solution will be also multiplied by this R, so that instead of the upper bound C there will be a new upper bound RC.

Suppose that for some q and for $n = q - 1$ we have

$$v_n(\xi) \leq C^n n!$$

with the same constant C as above. Then, by the remark in the beginning of the induction step with $R = C^n n! \times q \equiv C^{q-1} q!$, we immediately obtain

$$v_q(\xi) \leq C^{q-1} q! \times C = C^q q!,$$

as required. Hence, the inequality (27) follows. Note that a similar simple argument with a reference to the induction method and without using explicitly the second barrier N can be found in [22, Lemma 3.1]; practically the same calculus, yet with unbounded growing in x moments was used in [28].

Now, take any $\alpha \in (0, C^{-1})$. Then due to (27) one has

$$\mathbb{E}_\xi e^{\alpha \gamma} = \sum_{q=0}^{+\infty} \frac{\alpha^q \mathbb{E}_\xi \gamma^q}{q!} \leq \sum_{q=0}^{+\infty} \alpha^q C^q = \frac{1}{1 - \alpha C} < \infty. \tag{30}$$

It may be argued now that the desired "exponential coupling" can be arranged via the exponential moment bound (30) and due to the elliptic Harnack inequality for divergent

13

type equations [12, Theorem 8.20] in the way similar to [28], see the next step. Note that it is a "common knowledge" that the bound (27) suffices for the Theorem claim. The reader who knows the exact reference may skip the rest of the proof.

Step 8. Using Harnack inequality. The usage of coupling method assumes that glueing or meeting of two versions of the process – one stationary and another non-stationary – can be arranged with a positive probability bounded away from zero on each period of this construction. Here it suffices to consider a "symmetric" SDE on the whole line. By "period" in our case any finite interval may be taken; e.g., it is convenient to use $[0,2]$ which will be split into two equal parts, $[0,1]$ and $[1,2]$ (their intersection at one single point is not important). On the first half, according to the inequality (30) and Bienaymé – Chebyshev – Markov's inequality, both independent versions of the process will attain some (actually, any) bounded neighbourhood of zero. On the second half we want to glue them with a probability also bounded away from zero. Note that the standard 1D or finite state space idea just to wait until the two trajectories intersect here does not work straightforward as we would like it. Or, rather, it works but the bound obtained in such a way would use some bounds on the derivative π', which we want to avoid by all means.

There is a recent rather general tool based on regeneration period moments [33]. Yet, to verify the mild condition (*) required for this tool is probably no easier than – or, maybe, equivalent to – what we suggest instead in the next paragraphs.

One more approach which does not involve any properties of π' uses classical inequalities for divergence form PDEs. Indeed, this step justifies that it is possible by virtue of Moser's Harnack inequality for divergent type elliptic equation (see [12]) (cf., e.g., [30], [31] where a parabolic Harnack inequality was used for the same goal). Here it is convenient to return to an SDE on the whole line with symmetric coefficients (17) which solution is denoted by $\bar{X}_t$. Its modulus satisfies the equation (5) with a new Wiener process.

We argue that an elliptic Harnack inequality

$$\mathbb{E}_x g(\bar{X}_{\bar{\sigma}}) \leq C \mathbb{E}_{x'} g(\bar{X}_{\bar{\sigma}}) \tag{31}$$

for any non-negative function g and any $|x|, |x'| \leq 1$ with $\bar{\sigma} := \inf(t \geq 0 : |\bar{X}_t| \geq 2)$ follows from [12, Theorem 8.20] due to the equation $\mathrm{div}(\exp(2U(x)\nabla v(x)) = 0$, here $v(x) = \mathbb{E}_x g(\bar{X}_\sigma)$. This reasoning should be combined with the bound $\mathbb{P}_x(\bar{\sigma} > t) \leq Ct^{-1}$ for $t > 0$ with some $C > 0$ depending on the sup-norms of all coefficients in the ball $B := \{|x| \leq 2\}$. The latter bound follows from [12, Theorem 8.16] applied and from the Bienaymé – Chebyshev–Markov inequality $\mathbb{P}_x(\sigma > t) \leq t^{-1} \mathbb{E}_x \sigma$ since the function $v(x) := \mathbb{E}_x \bar{\sigma}$ is a solution to the equation $\frac{1}{2} \exp(-2U)\mathrm{div}(\exp(2U)\nabla v) + 1 = 0$, & $v|_{\partial B} = 0$, or, equivalently, to the (Poisson) equation

$$\mathrm{div}(\exp(2U)\nabla v) + 2\exp(2U) = 0, \quad v|_{\partial B} = 0,$$

to which the Theorem 8.16 [12] is applicable stating that solution $v(x)$ is bounded by a constant, say, N depending only on $\sup_{|x| \leq 2} |U(x)|$ (actually, even on some integral norm of $\exp(2U)$). This immediately implies that by choosing $t \geq 3N$ we have that $\mathbb{P}_x(\bar{\sigma} \leq t) \geq \frac{2}{3}$. The same estimate holds true for the process X_t, with the stopping time $\sigma = \inf(t \geq 0 : X_t \geq 2)$, and with a non-negative function g on $\mathbb{R}^1_+$, i.e.,

$$\mathbb{E}_x g(X_\sigma) \leq C \mathbb{E}_{x'} g(X_\sigma).$$

14

Along with (30), this suffices for a successful exponential coupling for the process X_t with its stationary version. Although it will not be used here, note that the obtained bound implies a stronger exponential inequality $\mathbb{P}_x(\sigma > t) \leq C \exp(-\lambda t)$ with some $C, \lambda > 0$ by the well-known property of homogeneous Markov processes and their exit times.

Finally, we can change the function U outside the ball $|x| \leq 3$ so that it becomes bounded, – the latter is possible without changing the process until σ.

Step 9. Exponential convergence. Let us return to the half-line $\mathbb{R}_+$ and to the process X_t, and let us fix some $K > 0$. It is known that – modulo the conclusion of the previous step – for the proof of the desired exponential convergence in total variation, it suffices to show that $\mathbb{E} \exp(\alpha \gamma_X^\xi)$ is finite for some $\alpha > 0$, and for some – *actually, for any* – $K > 0$ where $\gamma_X^\xi = \gamma_X^{\xi,K}$ was defined in (24).

Let X_t^{st} be the *independent* stationary version of the Markov process X_t, i.e., X_t^{st} is the process with the same generator and initial distribution with the density π, if necessary, on some extended probability space with another independent Wiener process. (However, we will not change our notations for $\mathbb{P}$ and $\mathbb{E}$.) Naturally, the couple (X_t, X_t^{st}) is considered on some extension of the original probability space. For $\xi > K$ let $\tau \equiv \tau^\xi$ be the moment of the first intersection of X_t started from $X_0 = \xi$ with the stationary version X_t^{st}, i.e.,

$$\tau := \inf \left(t \geq 0: \; X_t = X_t^{st} \right).$$

As the random variable τ is a stopping time and X_t has strong Markov property, we can define a new strong Markov process

$$\hat{X}_t := X_t 1\,(t < \tau) + X_t^{st} 1\,(t \geq \tau), \tag{32}$$

with the property

$$\text{Law}\,(\hat{X}_t) = \text{Law}\,(X_t).$$

Obviously on $\{t > \tau\} \cap \{\tau < \infty\}$ the trajectories of X_t^{st} and $\hat{X}_t$ "after τ" coincide. Then for any Borel set A one has (we drop the initial value ξ since the final estimate is uniform in it)

$$\left| \mathbb{P}\,(X_t \in A) - \mathbb{P}\,(X_t^{st} \in A) \right| \overset{(32)}{=} \left| \mathbb{P}\left(\hat{X}_t \in A\right) - \mathbb{P}\,(X_t^{st} \in A) \right| =$$

$$= \left| \mathbb{E}\left(\left(1\left(\hat{X}_t \in A\right) - 1\left(X_t^{st} \in A\right) \right) \times \left(1\,(t < \tau) + 1\,(t \geq \tau) \right) \right) \right| =$$

$$= \left| \mathbb{E}\left(\left(1\left(\hat{X}_t \in A\right) - 1\left(X_t^{st} \in A\right) \right) \times 1\,(t < \tau) \right) \right|$$

$$\leq \mathbb{E}\left| \left(1\left(\hat{X}_t \in A\right) - 1\left(X_t^{st} \in A\right) \right) \right| \times 1\,(t < \tau).$$

Since $\left| \left(1\left(\hat{X}_t \in A\right) - 1\,(X_t^{st} \in A) \right) \right| \leq 1$, taking into account the exponential version of Bienaymé – Chebyshev–Markov's inequality, we conclude that

$$\left| \mathbb{P}\,(X_t \in A) - \mathbb{P}\,(X_t^{st} \in A) \right| \leq \mathbb{E} 1\,(t < \tau) = \mathbb{P}\,(t < \tau) \leq \exp\,(-\alpha t)\,\mathbb{E} \exp(\alpha \tau).$$

Passing to the supremum over Borel sets A we obtain due to (30),

$$\|\mu_t - \mu_{st}\|_{TV} \leq 2 \exp\left(-\alpha t\right) \mathbb{E} \exp(\alpha \tau) \leq \frac{2}{1 - \alpha C} \exp\left(-\alpha t\right),$$

where $\mu_t = Law\,(X_t)$ and $\mu_{st}(dx) = \pi(x)dx$. This bound does not depend on the initial value ξ which was dropped in the notation $\mathbb{E}_\xi$. Hence, the proof of the Theorem is completed. $\square$

Remark 1. *It is likely that for the symmetric density π on $\mathbb{R}^1$ and for the equation (2) on $\mathbb{R}^1$ this method is applicable, too, with the function $f(y) \equiv f(|y|)$, and that it provides a similar convergence rate bound (11) as in the Theorem 1. We leave it till further papers.*

Remark 2. *The form of the process X_t is a result of an educated guess. We were looking for the process X_t whose generator L would be the generator of the Langevin diffusion multiplied by a positive function $F = f^2$, which is needed for applying a random time change. At the same time, we wanted the new process X_t to have the same invariant density π. From this condition the function F is determined up to two constants $c_1, c_2 > 0$*

$$F(x) = c_1 + \int_0^x \frac{c_2 dy}{\pi(y)}, \quad x \geq 0.$$

It can be checked by an explicit computation that with this choice of the function F, the new process X_t would still have an exponential rate of convergence to the invariant measure (we choose $c_1 = c_2 = 1$ but in fact any strictly positive c_1 and c_2 give the same result).

Remark 3. *Note that in the Theorem 1 the property of continuity of the state space is important. If the state space is discrete, the modification we consider (multiplication of the generator by a function) typically does not affect the rate of convergence. Indeed, let us consider a birth-death process Y_t, $t \geq 0$ with birth rates $\{\lambda_n, n \in \mathbb{N}_{\geq 0}\}$ and death rates $\{\mu_n, n \in \mathbb{N}\}$:*

$$\mathbb{P}\left(Y_{t+h} = m | Y_t = n\right) = \begin{cases} \lambda_n h + o(h), & m = n+1 \\ \mu_n h + o(h), & m = n-1 \\ 1 - \mu_n h - \lambda_n h + o(h), & m = n \\ o(h), & |m - n| > 1 \end{cases} \tag{33}$$

The generator A_0 of the process Y_t is given by

$$A_0 \varphi(n) = \lambda_n(\varphi(n+1) - \varphi(n)) + \mu_n(\varphi(n-1) - \varphi(n)).$$

If we want to apply the same transformation as in the continuous case, i.e. to consider a birth-death process X_t with the generator $A = f_n \cdot A_0$, then the new process X_t will have the birth and death rates $\{\lambda'_n = f_n \cdot \lambda_n, n \in \mathbb{N}_{\geq 0}\}$ and $\{\mu'_n = f_n \cdot \mu_n, n \in \mathbb{N}\}$ respectively.

The invariant distribution π of Y_t for the can be computed explicitly and equals

$$\pi(\{n\}) \equiv \pi_n = \pi_0 \frac{\lambda_0 \ldots \lambda_{n-1}}{\mu_1 \ldots \mu_n}, \quad \pi_0 = \left(1 + \sum_{n \geq 1} \frac{\lambda_0 \ldots \lambda_{n-1}}{\mu_1 \ldots \mu_n}\right)^{-1}.$$

Hence

$$\frac{\lambda_{n-1}}{\mu_n} = \frac{\pi_{n-1}}{\pi_n} \quad for \ each \quad n \in \mathbb{N}.$$

The assumption that X_t has the same invariant distribution as Y_t yields

$$\frac{\lambda_{n-1}^{'}}{\mu_n^{'}} = \frac{f_{n-1} \cdot \lambda_{n-1}}{f_n \cdot \mu_n} = \frac{\pi_{n-1}}{\pi_n} \quad for \ each \quad n \in \mathbb{N},$$

hence $f_n = f_0 = const$ for all values of n. So, such a transformation is just changing the time scale by multiplying it by a positive constant which doesn't influence the rate of convergence qualitatively.

Acknowledgements

References

[1] N.Abourashchi, A.Yu. Veretennikov, On exponential mixing bounds and convergence rate for reciprocal Gamma diffusion processes, Mathematical Communications, 14(2) (2009), 331-339.

[2] N. Abourashchi, A.Yu. Veretennikov, On exponential mixing and rate of convergence for Student processes, Theory Probab. Math. Statist. 81 (2010), 1-13.

[3] Y. Aït-Sahalia, Nonparametric pricing of interest rate derivative securities, Econometrica, 1996, 64(3), 527–560.

[4] B.M. Bibby, I.M. Skovgaard, M. Sørensen, Diffusion-type models with given marginal distribution and autocorrelation function, Bernoulli, 11(2) (2005), 191–220.

[5] V.I. Bogachev, N.V. Krylov, and M. Röckner, On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions, Communications in Partial Differential Equations, 26(11) (2001), 2037–2080.

[6] V.I. Bogachev, N.V. Krylov, M. Röckner, S.V. Shaposhnikov, Fokker–Planck–Kolmogorov Equations, AMS, Providence, RI, 2015.

[7] P. Cattiaux, A. Guillin, Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity, Journal of Functional Analysis, 272 (2017) 2361-2391.

[8] R. Douc, E. Moulines, P. Priouret, P. Soulier, Markov Chains, Springer, 2018.

[9] G. Fort, E. Moulines, P. Priouret, Convergence of adaptive and interacting Markov Chain Monte Carlo algorithms, Annals of Statistics, 39(6) (2011) 3262–3289.

[10] M. Fukushima, Dirichlet forms and Markov processes, Amsterdam; New York: North-Holland Pub. Co. & Elsevier North-Holland, 1980.

[11] I. Gikhman and A. V. Skorokhod, Stochastic Differential Equations, Springer, New York, 1972.

[12] D. Gilbarg, N.S. Trudinger, Elliptic partial differential equations of second order. 2nd ed., Springer, Berlin et al., 1983.

[13] I.A Ibragimov, Yu.V. Linnik, Independent and stationary sequences of random variables, Groningen, Wolters-Noordhoff, 1971.

[14] K. Itô, H.P. McKean, Diffusion Processes and Their Sample Paths. Springer, 1965.

[15] S.A. Klokov, Lower Bounds of Mixing Rate for a Class of Markov Processes, Theory Probab. Appl., 51(3) (2007) 528–535.

[16] R. Khasminskii, Stochastic Stability of Differential Equations, 2nd Ed., with contribution of G.N. Milstein and M.B. Nevelson, Springer, Berlin, 2012.

[17] Y. Kovchegov, N. Michalowski, A Class of Markov Chains with no Spectral Gap, Proc. AMS, 141(12) (2013), 4317–4326.

[18] N.V. Krylov, On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes, Math. USSR-Izv., 7(3) (1973), 691–709.

[19] A.M. Kulik, N.N. Leonenko, Ergodicity and mixing bounds for the Fisher–Snedecor diffusion. Bernoulli 19(5B), (2013), 2294–2329.

[20] E. Löcherbach, D. Loukianova, O. Loukianov, Polynomial bounds in the Ergodic theorem for one-dimensional diffusions and integrability of hitting times, Ann. Inst. Henri Poincare – prob. 47 (2011), 425–449.

[21] Z.-M. Ma, M. Röckner, Introduction to the Theory of (Non-Symmetric) Dirichlet Forms, 1st ed., Springer, 1992.

[22] Y. Mao, Convergence rates in strong ergodicity for Markov processes, Stoch. Proc. Appl. 116(2006), 1946-1976.

[23] H.P. McKean, Stochastic integrals, AMS Chelsea Publishing, 1969 (reprinted 2005).

[24] M.V. Menshikov, S.Y. Popov, Exact power estimates for countable Markov chains, Markov Process. Related Fields, 1 (1) (1995) 57–78.

[25] M. Menshikov, S. Popov, A. Wade, Non-homogeneous Random Walks: Lyapunov Function Methods for Near-Critical Stochastic Systems, Cambridge, CUP, 2016.

[26] A. Uglov, A. Veretennikov, Yet again on polynomial convergence for SDEs with a gradient-type drift, arXiv:1706.09374.

[27] A.Yu. Veretennikov, On strong and weak solutions of one - dimensional stochastic equations with boundary conditions. Theory Probab. Appl. 26(4) (1981), 670-686.

[28] A.Yu. Veretennikov, On polynomial mixing for SDEs with a gradient-type drift, Theory Probab. Appl., 45(1) (2001) 160–164.

[29] A.Yu. Veretennikov, On lower bounds for mixing coefficients of Markov diffusions, In: From Stochastic Calculus to Mathematical Finance; The Shiryaev Festschrift. Kabanov, Yu.; Liptser, R.; Stoyanov, J. (Eds.) Springer, Berlin et al. (2006), 623–633.

[30] A.Yu. Veretennikov, On polynomial mixing bounds for stochastic differential equations. Stochastic Processes and their Applications, 70 (1997) 115–127.

[31] A.Yu. Veretennikov, On polynomial mixing and convergence rate for stochastic difference and differential equations, Teoriya Veroyatnostej i ee Primenen., 44(2) (1999) 312–327.

[32] A.Yu. Veretennikov. Ergodic Markov processes and Poisson equations (lecture notes). In: Modern problems of stochastic analysis and statistics - Selected contributions in honor of Valentin Konakov (editor: V. Panov), Springer, 2017, 457 – 511.

[33] G.A. Zverkina, On strong bounds of rate of convergence for regenerative processes, Cham, Springer, 2016, 381-393.