ESAIM: PS 17 (2013) 195–218
DOI: 10.1051/ps/2011156

ESAIM: Probability and Statistics
www.esaim-ps.org

POLYNOMIAL DEVIATION BOUNDS FOR RECURRENT HARRIS PROCESSES HAVING GENERAL STATE SPACE

Eva Löcherbach[1] and Dasha Loukianova[2]

Abstract. Consider a strong Markov process in continuous time, taking values in some Polish state space. Recently, Douc *et al.* [*Stoc. Proc. Appl.* **119**, (2009) 897–923] introduced verifiable conditions in terms of a supermartingale property implying an explicit control of modulated moments of hitting times. We show how this control can be translated into a control of polynomial moments of abstract regeneration times which are obtained by using the regeneration method of Nummelin, extended to the time-continuous context. As a consequence, if a p-th moment of the regeneration times exists, we obtain non asymptotic deviation bounds of the form

$$P_\nu\left(\left|\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s-\mu(f)\right|\geq\varepsilon\right)\leq K(p)\frac{1}{t^{p-1}}\frac{1}{\varepsilon^{2(p-1)}}\|f\|_\infty^{2(p-1)},\quad p\geq 2.$$

Here, f is a bounded function and μ is the invariant measure of the process. We give several examples, including elliptic stochastic differential equations and stochastic differential equations driven by a jump noise.

Mathematics Subject Classification. 60J55, 60J35, 60F10, 62M05.

Received November 29, 2010. Revised August 26, 2011.

1. Introduction

Let X be a positive Harris recurrent strong Markov process in continuous time, having invariant probability measure μ. From the Ergodic theorem we know that for all $x\in I\!R$, $f\in L^1(\mu)$ and $\varepsilon>0$

$$P_x\left(\left|\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s-\mu(f)\right|\geq\varepsilon\right)\to 0 \tag{1.1}$$

as t goes to infinity. The purpose of this paper is to establish the rate of convergence in (1.1), for bounded functions f. In the existing literature, mainly the case of exponential rate of convergence (exponential ergodicity) has been considered. But recently, there has been growing interest in studying other possible rates such as

Keywords and phrases. Harris recurrence, polynomial ergodicity, Nummelin splitting, continuous time Markov processes, drift condition, modulated moment.

[1] CNRS UMR 8088, Département de Mathématiques, Université de Cergy-Pontoise, 95000 Cergy-Pontoise, France. eva.loecherbach@u-cergy.fr

[2] Département de Mathématiques, Université d'Evry-Val d'Essonne, Bd François Mitterrand, 91025 Evry Cedex, France. dasha.loukianova@univ-evry.fr

sub-geometric or polynomial rates. We follow this direction and study in this paper the case when the rate of convergence in (1.1) is polynomial. More precisely we use the so-called regeneration method and show that if a certain regeneration time admits a p-th moment, then we obtain non asymptotic deviation bounds of the form

$$P_x\left(\left|\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s-\mu(f)\right|\geq\varepsilon\right)\leq K(p,x)\frac{1}{t^{p-1}}\frac{1}{\varepsilon^{2(p-1)}}\|f\|_\infty^{2(p-1)},\quad p\geq 2. \tag{1.2}$$

Here, f is a bounded function and μ is the invariant measure of the process. Such a bound is of major importance for many applications, for example non asymptotic problems for statistics of diffusions, concentration for particular approximations of granular media equations, and many other examples.

Let us give some comments on the history of the problem and compare our result with known results on deviation inequalities for Markov processes. In the context of Markov chains, Clémençon [6] and Bertail and Clémençon [3] have obtained bounds in (1.1) which are exponential in time, using the regeneration method of Nummelin. They work under the conditions of geometric (exponential) ergodicity and stationarity, and within the space of bounded functions. Our work is close to this in spirit, since we use the regeneration method as well (however, we use it in a more complicated framework since we work in continuous time). Compared to their work we do not need to assume stationarity, our results hold for any starting point x or any starting measure provided it integrates the p-th moment of the regeneration time. Moreover, we weaken the assumption of exponential ergodicity to polynomial ergodicity. Still in the discrete framework of Markov chains let us also mention the work by Adamczak [1] who derives, using completely different techniques, concentration inequalities for empirical processes of Markov chains, in the regime of exponential ergodicity. Finally, Chazottes *et al.* [5] obtain concentration inequalities for finite valued random fields on $\mathbb{Z}^d$ *via* coupling both in the exponential and the sub-exponential regime. For their purposes, the finiteness of the phase space is crucial.

All above mentioned results hold either in discrete time or in discrete state space, and this is not what we are interested in. In this paper we concentrate on the framework of continuous time and general state space. For continuous time Markov processes there is a huge literature on the subject, and most of the results are based on functional inequalities and/or perturbation techniques which are used to obtain non-asymptotic bounds in (1.1). As a matter of fact, in contrary to our approach, most of these papers deal with the stationary case only or with the case when the initial law of the process is absolutely continuous with respect to the invariant measure, having a square integrable density. Wu [24] uses the Lumer–Phillips theorem in order to derive non-asymptotic deviation bounds which are expressed in terms of a large deviations rate function. He works under the assumption that the initial law of the process is absolutely continuous with respect to the invariant measure. Based on this, Cattiaux and Guillin [4] use functional inequalities like the Poincaré inequality in order to derive an exponential deviation bound; they work under the assumption of a spectral gap and with bounded functions. A small paragraph in Cattiaux and Guillin [4] is devoted to the polynomial regime as well, under an assumption imposing polynomial decay of the α-mixing coefficient of the process, but the rate which is obtained is not optimal. In the same spirit, let us cite Guillin *et al.* [10] who work in the space of Lipschitz functions under the assumption of a spectral gap. For bounded functions, they obtain a Hoeffding type inequality. Finally, Lezaud [14] uses Kato's theory of perturbation of operators, still in the exponential regime. Let us also mention that in a completely different setting and having different applications in mind, Pal [18] establishes concentration inequalities for diffusion laws on the path space $C([0,\infty))$, using quadratic transportation cost inequalities. He studies concentration around the median of the distribution, in the exponential regime, for Lipschitz functions on the path space with respect to the uniform norm.

In contrast to most of the above mentioned papers, we do not assume exponential ergodicity, nor the existence of a spectral gap nor stationarity. We do not need to assume that the process is μ-symmetric. The method we use is the so-called regeneration method. It appeals to the condition of integrability of regeneration times. Let us describe briefly what is the idea of regeneration times. In the easiest situation where the process X has a recurrent point x_0, we may introduce a sequence of stopping times R_n, called *regeneration times*, such that:

1. for all n, $R_n<\infty$, $R_{n+1}=R_n+R_1\circ\theta_{R_n}$, $R_n\to\infty$ as $n\to\infty$. (Here, θ denotes the shift operator);

2. for all n, $X_{R_n} = x_0$;
3. for all n, the process $(X_{R_n+t})_{t\geq 0}$ is independent of $\mathcal{F}_{R_n}$.

In this case, paths of the process can be decomposed into i.i.d. excursions $[R_i, R_{i+1}[, i = 1, 2, \ldots$, plus an initial segment $[0, R_1]$, and then limit theorems follow immediately from the strong law of large numbers.

In general, recurrent points exist only in one-dimensional models. For one-dimensional recurrent diffusions it has been shown in Löcherbach *et al.* [16] that, if for some $p > 1$ the p-th moment of the regeneration time exists, then (1.2) holds.

For general multidimensional Harris recurrent processes, there is no direct way of defining regeneration times. However, there is a well-known method of introducing regeneration times artificially, which is known as method of "Nummelin splitting" in the case of Markov chains and which has been extended to the case of processes in continuous time by Löcherbach and Loukianova [15]. This method consists of constructing a bigger process $Z = (Z^1, Z^2, Z^3)$ taking values in $E \times [0, 1] \times E$, along a sequence of jump times $0 = T_0 < T_1 < \ldots < T_n < \ldots$, such that:

1. Z^1 is a copy of the original process X, and the T_n are arrival times of a rate-1-Poisson process, independent of Z^1;
2. on each time interval $[T_n, T_{n+1}[$, Z^2 and Z^3 are constant;
3. the sequence $(Z^3_{T_n})_n$ is a copy of the resolvent chain $X_{T_{n+1}}$ (the process X observed after independent exponential times);
4. the sequence $(Z^2_{T_n})_n$ is a copy of independent random variables, which are uniformly distributed on $[0, 1]$.

The three co-ordinates and the sequence of jump times $(T_n)_n$ are constructed in a *coupled* way, inspired by the splitting technique of Nummelin [17] and Athreya and Ney [2] in discrete time. We recall the whole construction in Section 3. The main point of this construction is that there exist a measurable set C having $\mu(C) > 0$ (C will be a *petite set* in the Meyn-Tweedie terminology) and a parameter $\alpha \in]0, 1]$ such that the successive visits of Z_{T_n} to $C \times [0, \alpha] \times E$ induce regeneration times for the process Z.

To resume, for any Harris recurrent Markov process X, the following holds true: the process X can be embedded as first co-ordinate into a new Markov process Z. This new process Z possesses regeneration times. These regeneration times are closely related to the hitting times of a certain petite set C, or in other words: the moments of regeneration times are closely related to hitting time moments. Once we have a p-th moment for the regeneration times, we obtain a control on the speed of convergence in the ergodic theorem and (1.2) holds true.

Note that different coupling techniques in spirit of the so-called Doeblin- or Dobrushin-coupling have been considered in the literature, for example in the case of diffusions by Veretennikov [22, 23], and for Lévy-noise driven solutions of SDE's by Kulik [11]. These couplings are more specific to the concrete models the authors are interested in – the coupling technique presented in this paper has the advantage of being completely general, as far as Harris recurrent processes are concerned.

Once the coupling is constructed, it remains to establish sufficient conditions on the generator of the process ensuring that p-th moments for regeneration times exist. These conditions are inspired by a recent work of Douc *et al.* [7] on sub-geometric rates of convergence for strong Markov processes. In this work, the authors introduce a drift condition towards a closed petite set in the spirit of a condition of existence of a Lyapunov function. This condition provides an upper bound on the control of sub-geometric or polynomial moments of hitting times where the dependence on the starting point is precisely given. The drift condition also provides a verifiable condition ensuring positive Harris recurrence of the process. We recall these results in Section 2. Section 3 is devoted to give a self-contained description of the state of the art concerning the regeneration or Nummelin-splitting-method in the multidimensional case. Section 4 provides a link between the two approaches "Drift Condition" of Douc *et al.* [7] and "Nummelin splitting". We show that the drift condition of Douc *et al.* [7] provides an upper bound on the regeneration times introduced according to the method of Nummelin splitting. More precisely, we show in Theorem 4.1 that certain polynomial moments up to a precise order are bounded – the bound on the order being determined by the Lyapunov condition. The dependence upon the starting point is controlled by the Lyapunov function as usual. So even though the moments of regeneration times can not be explicitly calculated, we get at least upper bounds in the rate of convergence in (1.1). As a main application of this result, in Section 5 we state and give the proof of the deviation inequality (1.2). Section 6 is devoted to some examples: multi-dimensional diffusions and SDE's driven by a jump noise that are treated in the spirit of a recent work of Kulik [11]. We close the paper with an appendix which recalls the Fuk–Nagaev inequality in the framework needed in Section 5.

2. Drift-condition, Harris-recurrence and modulated moments

Consider a probability space $(\Omega, \mathcal{A}, (P_x)_x)$. Let $X = (X_t)_{t\geq 0}$ be a process defined on $(\Omega, \mathcal{A}, (P_x)_x)$ which is strong Markov, taking values in a locally compact Polish space $(E, \mathcal{E})$, with càdlàg paths. $(P_x)_{x\in E}$ is a collection of probability measures on $(\Omega, \mathcal{A})$ such that $X_0 = x$ P_x-almost surely. We write $(P_t)_t$ for the transition semigroup of X. Moreover, we shall write $(\mathcal{F}_t)_t$ for the filtration generated by the process.

Throughout this paper, we impose the following condition on the transition semigroup $(P_t)_t$ of X.

Assumption 2.1. *There exists a sigma-finite positive measure Λ on $(E, \mathcal{E})$ such that for every $t > 0$, $P_t(x, \mathrm{d}y) = p_t(x, y)\Lambda(\mathrm{d}y)$, where $(t, x, y) \mapsto p_t(x, y)$ is jointly measurable.*

We are seeking for conditions ensuring that the process X is recurrent in the sense of Harris. The most popular conditions for Harris-recurrence are drift conditions or more generally conditions in terms of a supermartingale property for a functional of the Markov process. We follow Douc *et al.* [7] and impose a drift condition towards a closed petite set B which implies the Harris recurrence of the process. Recall that a set $B \in \mathcal{E}$ is *petite* if there exists a probability measure a on $\mathcal{B}(I\!R_+)$ and a measure ν_a on $(E, \mathcal{E})$ such that

$$\int_0^\infty P_t(x, \mathrm{d}y)a(\mathrm{d}t) \geq 1_B(x)\nu_a(\mathrm{d}y). \tag{2.3}$$

Assumption 2.2. *There exists a closed petite set B, a continuous function $V : E \to [1, \infty[$, an increasing differentiable concave positive function $\Phi : [1, \infty) \to (0, \infty)$ and a constant $b < \infty$ such that for any $s \geq 0$, $x \in E$,*

$$E_x(V(X_s)) + E_x\left(\int_0^s \Phi \circ V(X_u)\mathrm{d}u\right) \leq V(x) + bE_x\left(\int_0^s 1_B(X_u)\mathrm{d}u\right). \tag{2.4}$$

Remark 2.3. If $V \in \mathcal{D}(\mathcal{A})$ belongs to the domain of the extended generator $\mathcal{A}$ of the process X, then Theorem 3.4 of Douc *et al.* [7] shows that

$$\mathcal{A}V(x) \leq -\Phi \circ V(x) + b1_B(x) \tag{2.5}$$

implies the above Assumption 2.2.

By Proposition 3.1 of Douc *et al.* [7], we know that under Assumption 2.2, the process X is positive recurrent in the sense of Harris. We write μ for its invariant probability measure. Hence, for any set $A \in \mathcal{E}$ such that $\mu(A) > 0$, we have $\limsup_{t\to\infty} 1_A(X_t) = 1$ almost surely. In particular the process is μ-irreducible.

Under Assumption 2.2, Douc *et al.* [7] give estimates on modulated moments of hitting times. Modulated moments are expressions of the type

$$E_x \int_0^\tau r(s)f(X_s)\mathrm{d}s,$$

where τ is a certain hitting time, r a rate function and f any positive measurable function. Knowledge of the modulated moments permits to interpolate between the maximal rate of convergence (taking $f \equiv 1$) and the maximal shape of functions f that can be taken in the ergodic theorem (taking $r \equiv 1$). In the present paper we are interested in the maximal rate of convergence and hence we shall always take $f \equiv 1$.

For the function Φ of (2.4) put

$$H_\Phi(u) = \int_1^u \frac{\mathrm{d}s}{\Phi(s)},\ u \geq 1,\ r_\Phi(s) = r(s) = \Phi \circ H_\Phi^{-1}(s). \tag{2.6}$$

We are interested in choices of the function Φ that yield a polynomial rate function r. This is achieved by the choice $\Phi(v) = cv^\alpha$ for $0 \leq \alpha < 1$ giving rise to polynomial rate functions

$$r(s) \sim Cs^{\frac{\alpha}{1-\alpha}}.$$

We suppose from now on that Assumption 2.2 is satisfied with such a kind of function $\Phi(v) = cv^\alpha$ for $0 \leq \alpha < 1$. The most important technical feature about the rate function that will be useful in the sequel is then the following sub-additivity property

$$r(t+s) \leq c(r(t) + r(s)), \tag{2.7}$$

for $t, s \geq 0$ and c a positive constant. We shall also use that

$$r(t+s) \leq r(t)r(s),$$

for all $t, s \geq 0$.

We are interested in regeneration time moments. We will see in Section 3 below that regeneration times are almost hitting times. Concerning hitting times, the following result is known in the literature. Fix $\delta > 0$ and define for any closed set $A \in \mathcal{E}$ the delayed hitting time

$$\tau_A(\delta) := \inf\{t \geq \delta : X_t \in A\}.$$

Then Theorem 4.1 and Proposition 4.5, (ii) of Douc *et al.* [7] imply the following two statements. Firstly, for the rate function r of (2.6) and for the petite set B of Assumption 2.2,

$$E_x \int_0^{\tau_B(\delta)} r(s)\mathrm{d}s \leq V(x) - 1 + \frac{b}{\Phi(1)} \int_0^\delta r(s)\mathrm{d}s. \tag{2.8}$$

Second, for the rate function r of (2.6) and for any closed set A with $\mu(A) > 0$, for any $\delta' > 0$,

$$E_x \int_0^{\tau_A(\delta')} r(s)\mathrm{d}s \leq c(A, \delta') \left[V(x) - 1 + \frac{b}{\Phi(1)} \int_0^\delta r(s)\mathrm{d}s \right]. \tag{2.9}$$

Remark 2.4. Suppose that $E = I\!R$ and that the process X has continuous trajectories. Fix a recurrent point $a \in I\!R$. Then we can choose $A = [a, \infty[$, if $x < a$, $A =]-\infty, a]$, if $x > a$ in (2.9) above. In this case, the successive visits

$$R_1 := \tau_{\{a\}}(\delta),\ R_{n+1} := \inf\{t \geq R_n + \delta : X_t = a\}$$

of the point a are regeneration times of the process. Hence, (2.9) gives a control of regeneration time moments in the one-dimensional case.

In the general multidimensional case, the times $\tau_A(\delta)$ do not define regeneration times any more. In this case, at least in general, regeneration times can only be introduced in an artificial manner, using the technique of Nummelin splitting in continuous time, as developed in Löcherbach and Loukianova [15]. However, the estimates (2.8) and (2.9) can be translated into bounds on moments of these new extended regeneration times of the process. This is the main issue of this paper and will be treated in Section 4 below.

In the next section we recall the technique of Nummelin splitting and then give the bounds on moments of the regeneration times. But before doing this we first recall some known facts about modulated moments of the resolvent chain from Douc *et al.* [8].

2.1. Modulated moments for the resolvent chain

Observing the continuous time process after independent exponential times gives rise to the resolvent chain and allows to use known results in discrete time instead of working with the continuous time process. This trick is quite often used in the theory of processes in continuous time.

Write $U^1(x, \mathrm{d}y) := \int_0^\infty \mathrm{e}^{-t} P_t(x, \mathrm{d}y)\mathrm{d}t$ for the resolvent kernel associated to the process. Introduce a sequence $(\sigma_n)_{n\geq 1}$ of i.i.d. exp(1)-waiting times, independent of the process X itself. Let $T_0 = 0$, $T_n = \sigma_1 + \ldots + \sigma_n$ and $\bar{X}_n = X_{T_n}$. Then the chain $\bar{X} = (\bar{X}_n)_n$ is recurrent in the sense of Harris, having the same invariant measure μ as the continuous time process, and its one-step transition kernel is given by $U^1(x, \mathrm{d}y)$.

Since X is Harris, it can be shown (Revuz [19], see also Prop. 6.7 of Höpfner and Löcherbach [13]), that the resolvent satisfies

$$U^1(x, \mathrm{d}y) \geq \alpha 1_C(x)\nu(\mathrm{d}y), \tag{2.10}$$

where $0 < \alpha \leq 1$, $\mu(C) > 0$ and ν a probability measure equivalent to $\mu(\cdot \cap C)$. The set C is in general not the petite set of Assumption 2.2. It can be chosen to be compact. In particular, (2.10) implies that the resolvent chain is aperiodic.

It is interesting to note that the drift condition (2.4) on the process in continuous time implies a similar drift condition on the resolvent chain. More precisely, Theorem 4.9 of Douc *et al.* [7], item (ii), implies that under Assumption 2.2 the resolvent chain satisfies a drift condition as well, with a different petite set and different functions $\bar{\Phi}$ and $\bar{V}$, but giving rise to the same rate function r since $\bar{\Phi}(t(1+\Phi'(1))) \sim \Phi(t)$ for $t \to \infty$. Moreover,

$$\|\bar{V} - V(1 + \Phi'(1))\|_\infty < \infty.$$

Now for any measurable set A with $\mu(A) > 0$, write $\bar{\tau}_A := \inf\{n \geq 1 : \bar{X}_n \in A\}$. Then, by Douc *et al.* (2004), proof of Theorem 2.8, second formula,

$$E_x\left[\sum_{k=0}^{\bar{\tau}_A - 1} r(k)\right] \leq c_1(A)\bar{V}(x) + c_2(A) \leq c_1 V(x) + c_2, \tag{2.11}$$

since $\bar{V}(x) \leq c_1 V(x) + c_2$.

After these preliminaries on resolvent chains we now turn to the description of the regeneration method in the case of a general state space.

3. Nummelin splitting and regeneration times

Regeneration times can be introduced for any Harris recurrent strong Markov process under the Assumption 2.1 – without any further assumption. We make once more use of the resolvent chain. Recall the definition of the resolvent kernel U^1 and the lower bound (2.10) which holds under the only assumption of Harris recurrence:

$$U^1(x, \mathrm{d}y) \geq \alpha 1_C(x)\nu(\mathrm{d}y),$$

where C is a fixed compact petite set with $\mu(C) > 0$. Note that since $\mu(C) > 0$, (2.9) and (2.11) hold for the hitting time of this set C.

Remark 3.1. Fort and Roberts [9] and Douc *et al.* [7] impose quite systematically the condition of irreducibility of some skeleton chain, see *e.g.* Theorems 3.2 and 3.3 of Douc *et al.* [7]. This implies the existence of some m such that P_m satisfies

$$P_m(x, \mathrm{d}y) \geq \alpha 1_C(x)\nu(\mathrm{d}y).$$

This condition is obviously stronger than (2.10) and implies that the process is not only positive Harris recurrent but also ergodic, *i.e.* for all $x \in E$,

$$||P_t(x, .) - \mu||_{TV} \to 0.$$

We do not impose this additional condition.

We now show how to construct regeneration times in continuous time by using the technique of Nummelin splitting which has been introduced for Harris recurrent Markov chains in discrete time by Nummelin [17] and Athreya and Ney [2]. The idea is to define on an extension of the original space $(\Omega, \mathcal{A}, (P_x))$ a Markov process $Z = (Z_t)_{t\geq 0} = (Z_t^1, Z_t^2, Z_t^3)_{t\geq 0}$, taking values in $E \times [0,1] \times E$ such that the times T_n are jump times of the process and such that $((Z_t^1)_t, (T_n)_n)$ has the same distribution as $((X_t)_t, (T_n)_n)$. We recall the details of this construction from Löcherbach and Loukianova [15].

First of all, define the split kernel $Q((x,u), \mathrm{d}y)$. This is a transition kernel $Q((x,u), \mathrm{d}y)$ from $E \times [0,1]$ to E defined by

$$Q((x,u),\mathrm{d}y) = \begin{cases} \nu(\mathrm{d}y) & \text{if } (x,u) \in C \times [0,\alpha] \\ \frac{1}{1-\alpha}\left(U^1(x,\mathrm{d}y) - \alpha\nu(\mathrm{d}y)\right) & \text{if } (x,u) \in C \times]\alpha,1] \\ U^1(x,\mathrm{d}y) & \text{if } x \notin C. \end{cases} \tag{3.12}$$

Remark 3.2. This kernel is called split kernel since $\int_0^1 \mathrm{d}u Q((x,u),\mathrm{d}y) = U^1(x,\mathrm{d}y)$. Thus Q is a splitting of the resolvent kernel by means of the additional "colour" u.

Write $u^1(x,x') := \int_0^\infty \mathrm{e}^{-t} p_t(x,x')\mathrm{d}t$. We now show how to construct the process Z recursively over time intervals $[T_n, T_{n+1}[, n \geq 0$. We start with some initial condition $Z_0^1 = X_0 = x$, $Z_0^2 = u \in [0,1]$, $Z_0^3 = x' \in E$. Then inductively in $n \geq 0$, on $Z_{T_n} = (x,u,x')$:

1. choose a new jump time σ_{n+1} according to
$$\mathrm{e}^{-t}\,\frac{p_t(x,x')}{u^1(x,x')}\,dt \text{ on } \mathbb{R}_+,$$
where we define $0/0 := a/\infty := 1$, for any $a \geq 0$, and put $T_{n+1} := T_n + \sigma_{n+1}$;
2. on $\{\sigma_{n+1} = t\}$, put $Z^2_{T_n+s} := u$, $Z^3_{T_n+s} := x'$ for all $0 \leq s < t$;
3. for every $s < t$, choose
$$Z^1_{T_n+s} \sim \frac{p_s(x,y)p_{t-s}(y,x')}{p_t(x,x')}\,\Lambda(\mathrm{d}y).$$
Choose $Z^1_{T_n+s} := x_0$ for some fixed point $x_0 \in E$ on $\{p_t(x,x') = 0\}$. Moreover, given $Z^1_{T_n+s} = y$, on $s+u < t$, choose
$$Z^1_{T_n+s+u} \sim \frac{p_u(y,y')p_{t-s-u}(y',x')}{p_{t-s}(y,x')}\Lambda(\mathrm{d}y').$$
Again, on $\{p_{t-s}(y,x') = 0\}$, choose $Z^1_{T_n+s+u} = x_0$;
4. at the jump time T_{n+1}, choose $Z^1_{T_{n+1}} := Z^3_{T_n} = x'$. Choose $Z^2_{T_{n+1}}$ independently of $Z_s, s < T_{n+1}$, according to the uniform law U. Finally, on $\{Z^2_{T_{n+1}} = u'\}$, choose $Z^3_{T_{n+1}} \sim Q((x',u'),\mathrm{d}x'')$.

Note that by construction, given the initial value of Z at time T_n, the evolution of the process Z^1 during $[T_n, T_{n+1}[$ does not depend on the chosen value of $Z^2_{T_n}$.

We will write P_π for the measure related to X, under which X starts from the initial measure $\pi(\mathrm{d}x)$, and $\mathbb{P}_\pi$ for the measure related to Z, under which Z starts from the initial measure $\pi(\mathrm{d}x)\otimes U(\mathrm{d}u)\otimes Q((x,u),\mathrm{d}y)$. Hence, $\mathbb{P}_{x_0}$ denotes the measure related to Z under which Z starts from the initial measure $\delta_{x_0}(\mathrm{d}x)\otimes U(\mathrm{d}u)\otimes Q((x,u),\mathrm{d}y)$. In the same spirit we denote E_π the expectation with respect to P_π and $\mathbb{E}_\pi$ the expectation with respect to $\mathbb{P}_\pi$. Moreover, we shall write $\mathbb{F}$ for the filtration generated by Z, $\mathbb{G}$ for the filtration generated by the first two co-ordinates Z^1 and Z^2 of the process, and $\mathbb{F}^X$ for the sub-filtration generated by X interpreted as first co-ordinate of Z.

The new process Z is a Markov process with respect to its filtration $\mathbb{F}$. For a proof of this result, the interested reader is referred to Theorem 2.7 of Löcherbach and Loukianova [15]. In general, Z will no longer be strong

Markov. But for any $n \geq 0$, by construction, the strong Markov property holds with respect to T_n. Thus for any $f, g : E \times [0,1] \times E \to \mathbb{R}$ measurable and bounded, for any $s > 0$ fixed, for any initial measure π on $(E, \mathcal{E})$,

$$\mathbb{E}_\pi(g(Z_{T_n}) f(Z_{T_n+s})) = \mathbb{E}_\pi(g(Z_{T_n}) \mathbb{E}_{Z_{T_n}}(f(Z_s))).$$

Finally, an important point is that by construction,

$$\mathcal{L}((Z_t^1)_t | \mathbb{P}_x) = \mathcal{L}((X_t)_t | P_x)$$

for any $x \in E$, thus the first co-ordinate of the process Z is indeed a copy of the original Markov process X, when disregarding the additional colours (Z^2, Z^3).

However, adding the colours (Z^2, Z^3) allows to introduce regeneration times for the process Z (not for X itself). More precisely, write

$$A := C \times [0, \alpha] \times E$$

and put

$$S_0 := 0,\ R_0 := 0, S_{n+1} := \inf\{T_m > R_n : Z_{T_m} \in A\}, R_{n+1} := \inf\{T_m : T_m > S_{n+1}\}. \tag{3.13}$$

The sequence of $\mathbb{F}-$stopping times R_n generalises the notion of life-cycle decomposition in the following sense.

Proposition 3.3 (Props. 2.6 and 2.13 of Löcherbach and Loukianova [15]).

(a) *under $\mathbb{P}_x$, the sequence of jump times $(T_n)_n$ is independent of the first co-ordinate process $(Z_t^1)_t$ and $(T_n - T_{n-1})_n$ are i.i.d.* $\exp(1)-$*variables;*
(b) *at regeneration times, we start from a fixed initial distribution which does not depend on the past: $Z_{R_n} \sim \nu(\mathrm{d}x) U(\mathrm{d}u) Q((x,u), \mathrm{d}x')$ for all $n \geq 1$;*
(c) *at regeneration times, we start afresh and have independence after a waiting time: $Z_{R_n+\cdot}$ is independent of $\mathcal{F}_{S_n-}$ for all $n \geq 1$;*
(d) *the sequence of $(Z_{R_n})_{n \geq 1}$ is i.i.d.*

Since the original process X – under Assumption 2.2 – is Harris with invariant measure μ, the new process Z will be Harris, too. We shall write Π for its invariant probability measure. Π can be written in terms of an occupation time formula which is a consequence of Chacon–Ornstein's ratio limit theorem. In order to state this theorem, let us recall that an additive functional of the process Z is a $\bar{\mathbb{R}}_+-$valued, $\mathbb{F}-$adapted process $A = (A_t)_{t \geq 0}$ such that:

1. almost surely, the process is non-decreasing, right-continuous, having $A_0 = 0$;
2. for any $s, t \geq 0$, $A_{s+t} = A_t + A_s \circ \theta_t$ almost surely. Here, θ denotes the shift operator.

The additive functional is called integrable if $\mathbb{E}_\Pi(A_1) < +\infty$. Examples for integrable additive functionals are $A_t = \int_0^t f(Z_s)\mathrm{d}s$, where f is a positive measurable function, integrable with respect to the invariant measure Π.

Proposition 3.4 (Chacon–Ornstein's ratio limit theorem). *Let A_t, B_t be any positive additive functionals of Z such that $\mathbb{E}_\Pi(B_1) > 0$. Then*

$$\frac{A_t}{B_t} \to \frac{\mathbb{E}_\Pi(A_1)}{\mathbb{E}_\Pi(B_1)} \quad \mathbb{P}_x - \text{ almost surely, as } t \to \infty,$$

for any $x \in E$. Moreover, Z is recurrent in the sense of Harris and its unique invariant probability measure Π is given by

$$\Pi(f) = \ell\ \mathbb{E}_\pi \int_{R_1}^{R_2} f(Z_s)\mathrm{d}s, \tag{3.14}$$

where $\ell = \mathbb{E}(R_2 - R_1)^{-1} > 0$.

Proof. The proof follows easily from the regeneration property with respect to the regeneration times R_n. □

The invariant measure μ of the original process X is the projection onto the first co-ordinate of Π. From this we deduce that the invariant probability measure μ of the original process X must be given by

$$\mu(f) = \ell\, I\!E_\pi \int_{R_1}^{R_2} f(X_s)\mathrm{d}s, \tag{3.15}$$

where we recall that $\ell = I\!E(R_2 - R_1)^{-1} > 0$. In the above formula we interpret X as first co-ordinate of Z, under $I\!P_\pi$ [3]. $R_2 - R_1$ is the length of one regeneration period. Under assumption (2.2), the process is positive recurrent and hence the expected length ℓ of one regeneration period is finite.

We now turn to the main issue of this article which is the control of the speed of convergence in the ergodic theorem. As a consequence of the above considerations, we can write

$$P_x\left(\left|\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s - \mu(f)\right| > \delta\right) = I\!P_x\left(\left|\frac{1}{t}\int_0^t f(Z_s^1)\mathrm{d}s - \ell\, I\!E_\pi \int_{R_1}^{R_2} f(Z_s^1)\mathrm{d}s\right| > \delta\right), \tag{3.16}$$

where we recall that $I\!P_x$ denotes the measure related to Z under which $Z_0 \sim \delta_x \otimes U(\mathrm{d}u) \otimes Q((x,u),\mathrm{d}y)$. The more moments of the regeneration period $R_2 - R_1$ exist, the more the process is recurrent and the more the convergence in (3.16) is fast.

We first give estimates on the polynomial moments

$$I\!E_x \int_0^{R_1} r(s)\mathrm{d}s,$$

depending on the starting point x. Integrating this against $\nu(\mathrm{d}x)$ gives then a control on the corresponding moment of the regeneration period. This integration does not pose any problems because the support of the measure ν is the compact set C. Since our regeneration times are built based on the resolvent chain, the main technical ingredient that allows such a control will be the estimate (2.11) rather than (2.9).

4. Polynomial moments of regeneration times

The aim of this section is to show that the results of Douc *et al.* [7] can be translated immediately into a control of moments of regeneration times. This yields somehow a link between the two different approaches "Drift conditions" *versus* "Nummelin". Recall the definition of $r(s) = r_\Phi(s)$ in (2.6).

Theorem 4.1. *Grant Assumptions 2.1 and 2.2 with a function $\Phi(v) = cv^\alpha$, where $0 \le \alpha < 1$. Then there exist constants c_1 and c_2, such that*

$$I\!E_x \int_0^{R_1} r(s)\mathrm{d}s \le c_1 V(x) + c_2.$$

Remark 4.2. For $\Phi(v) = cv^\alpha$, it can be easily shown that there exists a constant c such that $r(s) = r_\Phi(s) \ge c\, s^{\frac{\alpha}{1-\alpha}}$. Hence the above theorem implies the control of polynomial moments of the regeneration time, *i.e.*

$$I\!E_x R_1^{\frac{1}{1-\alpha}} \le \tilde{c}_1 V(x) + \tilde{c}_2. \tag{4.17}$$

Proof. Recall the definition of the regeneration times in (3.13). Let

$$\tilde{S}_1 := \inf\{T_n : Z^1_{T_n} \in C\},\ \tilde{S}_{n+1} := \inf\{T_k > \tilde{S}_n : Z^1_{T_k} \in C\}.$$

[3] Actually, we should write $I\!E_\pi \int_{R_1}^{R_2} f(Z_s^1)\mathrm{d}s$ – but if not otherwise indicated, this identification will always be implicitly assumed.

Obviously, $R_1 \geq \tilde{S}_1$.
1) In the following, c will denote a constant that might change from line to line. We first show how to control

$$I\!E_x \int_0^{\tilde{S}_1} r(s) \mathrm{d}s.$$

In a first step we show that

$$I\!E_x \int_0^{\tilde{S}_1} r(s) \mathrm{d}s = I\!E_x \int_0^\infty \mathrm{e}^{-\int_0^t 1_C(Z_s^1)\mathrm{d}s} r(t) \mathrm{d}t = E_x \int_0^\infty \mathrm{e}^{-\int_0^t 1_C(X_s)\mathrm{d}s} r(t) \mathrm{d}t. \tag{4.18}$$

This can be seen as follows. First, in order to obtain the law of $\tilde{S}_1$, we evaluate for any $a > 0$,

$$\begin{aligned} I\!P_x(\tilde{S}_1 > a) &= \sum_{n \geq 1} I\!P_x(\tilde{S}_1 = T_n, T_n > a) \\ &= \sum_{n \geq 1} I\!P_x(Z^1_{T_1} \in C^c, \ldots, Z^1_{T_{n-1}} \in C^c, Z^1_{T_n} \in C, T_n > a) \\ &= \sum_{n \geq 1} P_x(X_{T_1} \in C^c, \ldots, X_{T_{n-1}} \in C^c, X_{T_n} \in C, T_n > a) \\ &= E_x \left(\sum_{n \geq 1} (1 - 1_C(X_{T_1})) \cdots (1 - 1_C(X_{T_{n-1}})) f(X_{T_n}, T_n) \right), \end{aligned}$$

where $f(t, x) = 1_{t>a} 1_C(x)$.

Now, we make use of the following very useful formula which is taken from Höpfner and Löcherbach [13], (5.29), page 59.

$$\begin{aligned} E_x \left(\sum_{n \geq 1} (1 - 1_C(X_{T_1})) \cdots (1 - 1_C(X_{T_{n-1}})) f(X_{T_n}, T_n) \right) &= E_x \left(\int_0^\infty f(t, X_t) \mathrm{e}^{-\int_0^t 1_C(X_s)\mathrm{d}s} \mathrm{d}t \right) \\ &= E_x \left(\int_a^\infty 1_C(X_t) \mathrm{e}^{-\int_0^t 1_C(X_s)\mathrm{d}s} \mathrm{d}t \right). \end{aligned}$$

Hence we obtain

$$I\!P_x(\tilde{S}_1 > a) = E_x \left(\int_a^\infty 1_C(X_t) \mathrm{e}^{-\int_0^t 1_C(X_s)\mathrm{d}s} \mathrm{d}t \right) = E_x \left(\mathrm{e}^{-\int_0^a 1_C(X_s)\mathrm{d}s} \right).$$

Writing finally that

$$I\!E_x \int_0^{\tilde{S}_1} r(s) \mathrm{d}s = I\!E_x \int_0^\infty 1_{s < \tilde{S}_1} r(s) \mathrm{d}s = \int_0^\infty r(s) I\!P_x(\tilde{S}_1 > s) \mathrm{d}s,$$

we get (4.18). No we apply once more formula (5.29) of Höpfner and Löcherbach [13] and obtain

$$E_x \int_0^\infty \mathrm{e}^{-\int_0^t 1_C(X_s)\mathrm{d}s} r(t) \mathrm{d}t = E_x \left(\sum_{n=1}^\infty (1 - 1_C(\bar{X}_1)) \cdots (1 - 1_C(\bar{X}_{n-1})) r(T_n) \right), \tag{4.19}$$

where we recall that $\bar{X}_n = X_{T_n}$ is the process observed at the n-th jump time of an independent rate one Poisson process. The expression at the right hand side of (4.19) is almost a modulated moment for the resolvent chain, except that we have to replace $r(T_n)$ by $r(n)$. This is not difficult since for n large we can use the law of large numbers. Since r is increasing we can write

$$E_x\left((1-1_C(\bar X_1))\cdots(1-1_C(\bar X_{n-1}))r(T_n)\right) \le E_x\left((1-1_C(\bar X_1))\cdots(1-1_C(\bar X_{n-1}))r(2n)\right) \tag{4.20}$$
$$+E_x\left((1-1_C(\bar X_1))\cdots(1-1_C(\bar X_{n-1}))1_{T_n>2n}r(T_n)\right). \tag{4.21}$$

Let us start with the control of the first term in this decomposition. Recall that $\bar\tau_C = \inf\{n \ge 1 : \bar X_n \in C\}$. Now, using that $r(2n) \le cr(n)$, which follows from $r(t+s) \le c(r(t)+r(s))$ by (2.7),

$$\mathbb{E}_x\left(\sum_{n=1}^{\infty}(1-1_C(\bar X_1))\cdots(1-1_C(\bar X_{n-1}))r(2n)\right) = \mathbb{E}_x\left(\sum_{n=1}^{\bar\tau_C} r(2n)\right)$$
$$\le c\mathbb{E}_x\left(\sum_{n=1}^{\bar\tau_C} r(n)\right) \le c\mathbb{E}_x\left(\sum_{n=1}^{\bar\tau_C-1} r(n)\right) + c\mathbb{E}_x r(\bar\tau_C). \tag{4.22}$$

Let $R(k) = \sum_{j=0}^{k-1} r(j)$. Since r is polynomial, $\lim_{k\to\infty} r(k)/R(k) = 0$. Hence there exists a constant c such that for all $k \ge 1$, $r(k) \le R(k) + c$. As a consequence,

$$\mathbb{E}_x r(\bar\tau_C) \le c + \mathbb{E}_x\left(\sum_{n=0}^{\bar\tau_C-1} r(n)\right).$$

Using (2.11), we can thus conclude that

$$\mathbb{E}_x\left(\sum_{n=1}^{\infty}(1-1_C(\bar X_1))\cdots(1-1_C(\bar X_{n-1}))r(2n)\right) \le c_1V(x) + c_2.$$

Now we turn to the second expression in (4.20) above: for any $1 \le p, q$ such that $\frac{1}{p}+\frac{1}{q}=1$,

$$\mathbb{E}_x\left((1-1_C(\bar X_1))\cdots(1-1_C(\bar X_{n-1}))1_{T_n>2n}r(T_n)\right) \le \left[\mathbb{E}_x r^p(T_n)\right]^{1/p}\cdot\left[\mathbb{P}_x(T_n>2n)\right]^{1/q}$$
$$\le \left[\mathbb{E}_x r^p(T_n)\right]^{1/p}\cdot \mathrm{e}^{-Cn} \tag{4.23}$$

for some suitable constant C. But $r^p(\cdot)$ is polynomial and T_n the sum of n independent $\exp(1)$ variables, hence $\sup_x \mathbb{E}_x r^p(T_n) \le P(n)$, where $P(.)$ is a polynomial in n. As a consequence,

$$\sum_{n\ge1}\sup_x \mathbb{E}_x\left((1-1_C(\bar X_1))\cdots(1-1_C(\bar X_{n-1}))1_{T_n>2n}r(T_n)\right) = C_2 < \infty.$$

Putting together (4.18), (4.19)–(4.23), we thus get that

$$\mathbb{E}_x\int_0^{\tilde S_1} r(s)\mathrm{d}s \le c_1V(x)+c_2. \tag{4.24}$$

This will be the main contribution to the control of $\mathbb{E}_x\int_0^{R_1} r(s)\mathrm{d}s$. In the sequel, we shall also use that (4.24) implies in particular

$$\sup_{x\in C}\mathbb{E}_x\int_0^{\tilde S_1} r(s)\mathrm{d}s < +\infty, \tag{4.25}$$

since C is compact.

2) Recall the definition of S_1 in (3.13). We now show how to use the control of $\tilde S_1$ in order to obtain a control of S_1. We have, since $r(t+s) \le r(s)r(t)$,

$$\begin{aligned} I\!\!E_x \int_0^{S_1} r(s)\mathrm{d}s &= I\!\!E_x \int_0^{\tilde S_1} r(s)\mathrm{d}s + \sum_{n\ge 1} I\!\!E_x \left(\int_{\tilde S_n}^{\tilde S_{n+1}} r(s)\mathrm{d}s 1_{\tilde S_n < S_1} \right) \\ &= I\!\!E_x \int_0^{\tilde S_1} r(s)\mathrm{d}s + \sum_{n\ge 1} I\!\!E_x \left(\int_0^{\tilde S_{n+1}-\tilde S_n} r(\tilde S_n + s)\mathrm{d}s 1_{\tilde S_n < S_1} \right) \\ &\le I\!\!E_x \int_0^{\tilde S_1} r(s)\mathrm{d}s + \sum_{n\ge 1} I\!\!E_x \left(\left[\int_0^{\tilde S_{n+1}-\tilde S_n} r(s)\mathrm{d}s \right] r(\tilde S_n) 1_{\tilde S_n < S_1} \right). \end{aligned} \tag{4.26}$$

The first term in this expression can be controlled using (4.24). We study the second term in the above expression

$$I\!\!E_x \left(r(\tilde S_n) 1_{\tilde S_n < S_1} \int_0^{\tilde S_{n+1}-\tilde S_n} r(s)\mathrm{d}s \right).$$

We know that $I\!\!P_x(\tilde S_n < S_1) = (1-\alpha)^n$ (see for example the proof of Prop. 2.16 in Löcherbach and Loukianova [15]). A first idea would be to use Markov's property with respect to $\tilde S_n$:

$$I\!\!E_x \left(r(\tilde S_n) 1_{\tilde S_n < S_1} \int_0^{\tilde S_{n+1}-\tilde S_n} r(s)\mathrm{d}s \right) = I\!\!E_x \left(r(\tilde S_n) 1_{\tilde S_n < S_1} I\!\!E_{Z_{\tilde S_n}} \int_0^{\tilde S_1} r(s)\mathrm{d}s \right).$$

But unfortunately it is **not true** that

$$I\!\!E_{Z_{\tilde S_n}} \int_0^{\tilde S_1} r(s)\mathrm{d}s \le \sup_{x\in C} I\!\!E_x \int_0^{\tilde S_1} r(s)\mathrm{d}s,$$

we only have that on $\{\tilde S_n < S_1\}$,

$$I\!\!E_{Z_{\tilde S_n}} \int_0^{\tilde S_1} r(s)\mathrm{d}s \le \sup_{x\in C, u>\alpha, z\in E} I\!\!E_{(x,u,z)} \int_0^{\tilde S_1} r(s)\mathrm{d}s,$$

and this can not be directly controlled using (4.24).

Hence, we must be more careful. We use that $r(\tilde S_n)1_{\{\tilde S_n < S_1\}}$ is measurable with respect to $\mathcal{G}_{\tilde S_n}$ where we recall that $(\mathcal{G}_t)_t$ is the filtration generated by the first two co-ordinates Z^1 and Z^2 of Z. Hence we will condition on $\mathcal{G}_{\tilde S_n}$. Note that by construction of Z, this means that we condition on the whole history of the whole process, *i.e.* the three co-ordinates, up to the last jump time $\sup\{T_k : T_k < \tilde S_n\}$ strictly before $\tilde S_n$, and on the history of Z^1 and Z^2 up to time $\tilde S_n$. In other words, conditioning on $\mathcal{G}_{\tilde S_n}$, we know $Z^1_{\tilde S_n}$ and $Z^2_{\tilde S_n}$, but $Z^3_{\tilde S_n}$ has still to be chosen. Moreover, on $\{\tilde S_n < S_1\}$, $Z^2_{\tilde S_n} > \alpha$, and hence the second line in the definition of the kernel $Q((x,u),\mathrm{d}x')$ of (3.12) has to be applied.

Write $\nu(x)$ for the density of $\nu(\mathrm{d}x)$ with respect to the dominating measure $\Lambda(\mathrm{d}x)$ of Assumption 2.1. Then,

$$\begin{aligned} & I\!\!E_x \left(r(\tilde S_n) 1_{\tilde S_n < S_1} \int_0^{\tilde S_{n+1}-\tilde S_n} r(s)\mathrm{d}s \right) \\ &\quad = I\!\!E_x \left(r(\tilde S_n) 1_{\tilde S_n < S_1} \int_E \frac{1}{1-\alpha} \left[u^1(Z^1_{\tilde S_n}, x') - \alpha\nu(x') \right] \Lambda(\mathrm{d}x') I\!\!E_{(Z^1_{\tilde S_n}, Z^2_{\tilde S_n}, x')} \int_0^{\tilde S_1} r(s)\mathrm{d}s \right) \\ &\quad \le \frac{1}{1-\alpha} I\!\!E_x \left(r(\tilde S_n) 1_{\tilde S_n < S_1} \int_E u^1(Z^1_{\tilde S_n}, x')\Lambda(\mathrm{d}x') I\!\!E_{(Z^1_{\tilde S_n}, Z^2_{\tilde S_n}, x')} \int_0^{\tilde S_1} r(s)\mathrm{d}s \right). \end{aligned} \tag{4.27}$$

But for any x, u,

$$\int_E u^1(x,x')\Lambda(\mathrm{d}x')\mathbb{E}_{(x,u,x')}\int_0^{\tilde{S}_1} r(s)\mathrm{d}s = \int_0^1 \mathrm{d}u \int_E Q((x,u),\mathrm{d}x')\mathbb{E}_{(x,u,x')}\int_0^{\tilde{S}_1} r(s)\mathrm{d}s, \tag{4.28}$$

since $\mathbb{E}_{(x,u,x')}\int_0^{\tilde{S}_1} r(s)\mathrm{d}s$ does not depend on u. Moreover,

$$\int_0^1 \mathrm{d}u \int_E Q((x,u),\mathrm{d}x')\mathbb{E}_{(x,u,x')}\int_0^{\tilde{S}_1} r(s)\mathrm{d}s = \mathbb{E}_x \int_0^{\tilde{S}_1} r(s)\mathrm{d}s.$$

Hence, since $Z^1_{\tilde{S}_n} \in C$,

$$\begin{aligned}\mathbb{E}_x\left(r(\tilde{S}_n)1_{\tilde{S}_n<S_1}\int_0^{\tilde{S}_{n+1}-\tilde{S}_n} r(s)\mathrm{d}s\right) &\le \frac{1}{1-\alpha}\mathbb{E}_x\left(r(\tilde{S}_n)1_{\tilde{S}_n<S_1}\left(\sup_{x\in C}\mathbb{E}_x\int_0^{\tilde{S}_1} r(s)\mathrm{d}s\right)\right)\\ &\le \frac{c}{1-\alpha}\mathbb{E}_x\left(r(\tilde{S}_n)1_{\tilde{S}_n<S_1}\right). \end{aligned}\tag{4.29}$$

Hence we must control $\mathbb{E}_x(1_{\tilde{S}_n<S_1}r(\tilde{S}_n))$. We write $\tilde{S}_n = \tilde{S}_1 + (\tilde{S}_n - \tilde{S}_1)$ and use once more the sub-multiplicativity of r. We obtain

$$\mathbb{E}_x\left(r(\tilde{S}_n)1_{\tilde{S}_n<S_1}\right) \le \mathbb{E}_x\left(r(\tilde{S}_1)1_{\tilde{S}_1<S_1}r(\tilde{S}_n-\tilde{S}_1)1_{\tilde{S}_n<S_1}\right). \tag{4.30}$$

Here, we have cut $\tilde{S}_n = \tilde{S}_1 + (\tilde{S}_n - \tilde{S}_1)$ into two pieces in order to get a last term which does not depend on the starting point. The same arguments as above in (4.27) and (4.28) yield, when conditioning on $\mathcal{G}_{\tilde{S}_1}$, the following.

$$\begin{aligned}\mathbb{E}_x\left(r(\tilde{S}_n)1_{\tilde{S}_n<S_1}\right) &\le \mathbb{E}_x\left(r(\tilde{S}_1)1_{\tilde{S}_1<S_1}r(\tilde{S}_n-\tilde{S}_1)1_{\tilde{S}_n<S_1}\right)\\ &\le \mathbb{E}_x\left(r(\tilde{S}_1)1_{\tilde{S}_1<S_1}\frac{1}{1-\alpha}\int_E u^1(Z^1_{\tilde{S}_1},x')\Lambda(dx')\mathbb{E}_{(Z^1_{\tilde{S}_1},Z^2_{\tilde{S}_1},x')}[r(\tilde{S}_{n-1})1_{\tilde{S}_{n-1}<S_1}]\right)\\ &\le \frac{1}{1-\alpha}\mathbb{E}_x\left(r(\tilde{S}_1)1_{\tilde{S}_1<S_1}\,\mathbb{E}_{Z^1_{\tilde{S}_1}}[r(\tilde{S}_{n-1})1_{\tilde{S}_{n-1}<S_1}]\right)\\ &\le \frac{1}{1-\alpha}\sup_{y\in C}\mathbb{E}_y\left(r(\tilde{S}_{n-1})1_{\tilde{S}_{n-1}<S_1}\right)\ \mathbb{E}_x\left(r(\tilde{S}_1)1_{\tilde{S}_1<S_1}\right). \end{aligned}\tag{4.31}$$

Concerning the last term in the above expression, we use that $r(t) \le \int_0^t r(s)\mathrm{d}s + c$ for some constant c and obtain

$$\begin{aligned}\mathbb{E}_x\left(r(\tilde{S}_1)1_{\tilde{S}_1<S_1}\right) &\le c + \mathbb{E}_x\left(\int_0^{\tilde{S}_1} r(s)\mathrm{d}s\right)\\ &\le c + c_1V(x) + c_2 = c_1V(x) + \tilde{c}_2, \end{aligned}\tag{4.32}$$

using (4.24).

Concerning the first term in (4.31), for $p, q \ge 1$ such that $\frac{1}{p}+\frac{1}{q}=1$, we obtain

$$\begin{aligned}\sup_{y\in C}\mathbb{E}_y\left(r(\tilde{S}_{n-1})1_{\tilde{S}_{n-1}<S_1}\right) &\le \sup_{y\in C}\left(\mathbb{E}_y r^p(\tilde{S}_{n-1})\right)^{1/p}\mathbb{P}_y(\tilde{S}_{n-1}<S_1)^{1/q}\\ &\le (1-\alpha)^{(n-1)/q}\left(\sup_{y\in C}\mathbb{E}_y r^p(\tilde{S}_{n-1}))\right)^{1/p}. \end{aligned}\tag{4.33}$$

We have to control this last expression. We claim the following: there exists a $\kappa > 0$ and a constant c such that for $p > 1$ sufficiently small,

$$\left(\sup_{y \in C} I\!E_y r^p(\tilde{S}_{n-1})) \right)^{1/p} \le c n^{\kappa}. \tag{4.34}$$

Once (4.34) is proven, we obtain, using (4.26), (4.29) and (4.31)–(4.34) the following:

$$\begin{aligned} I\!E_x \int_0^{S_1} r(s) \mathrm{d}s &\le (c_1 V(x) + c_2) + \frac{c}{(1-\alpha)^2} (c_1 V(x) + \tilde{c}_2) \sum_{n \ge 1} (1-\alpha)^{(n-1)/q} n^{\kappa} \\ &= \bar{c}_1 V(x) + \bar{c}_2. \end{aligned} \tag{4.35}$$

It remains to show (4.34): by our assumptions, r is polynomial and $r(x) \sim C x^{\frac{\alpha}{1-\alpha}}$ as $x \to \infty$, hence $r^p(x) \le c x^{\kappa p}$, where $\kappa = \alpha/(1-\alpha)$. We now fix the choice of p and q in (4.33). We choose

$$p \in \left] \frac{1}{\kappa}, 1 + \frac{1}{\kappa} \right[\; = \; \left] \frac{1-\alpha}{\alpha}, \frac{1}{\alpha} \right[.$$

Then $\kappa p \ge 1$, and we use Jensen's inequality to obtain

$$r^p(\tilde{S}_{n-1}) \le c \tilde{S}_{n-1}^{\kappa p} \le (n-1)^{p\kappa - 1} \left(\tilde{S}_1^{\kappa p} + \ldots + (\tilde{S}_{n-1} - \tilde{S}_{n-2})^{\kappa p} \right). \tag{4.36}$$

Now since $p < 1 + \frac{1}{\kappa} = \frac{1}{\alpha}$, we have $t^{\kappa p} \le c \int_0^t r(s) \mathrm{d}s$ for some constant c. Then for any of the above terms $(k \ge 2)$, by (4.25),

$$\sup_{y \in C} I\!E_y (\tilde{S}_k - \tilde{S}_{k-1})^{\kappa p} \le c \sup_{y \in C} I\!E_y \int_0^{\tilde{S}_1} r(s) \mathrm{d}s < \infty.$$

As a consequence, coming back to (4.36),

$$\sup_{y \in C} I\!E_y r^p(\tilde{S}_{n-1}) \le c(n-1)^{p\kappa} \sup_{y \in C} I\!E_x \int_0^{\tilde{S}_1} r(s) \mathrm{d}s = \tilde{c}(n-1)^{p\kappa},$$

and this yields (4.34).
3) Finally we proceed to the control of R_1. Clearly,

$$I\!E_x \int_0^{R_1} r(s) \mathrm{d}s \le I\!E_x \int_0^{S_1} r(s) \mathrm{d}s + I\!E_x \left[r(S_1) \int_0^{R_1 - S_1} r(s) \mathrm{d}s \right].$$

We have to control the last term above. We condition on $\mathcal{G}_{S_1}$, notice that $Z_{S_1}^2 \le \alpha$ and use step 1. of the construction of Z, hence

$$I\!E_x \left[r(S_1) \int_0^{R_1 - S_1} r(s) \mathrm{d}s \right] = I\!E_x \left[r(S_1) \left(\int_E \nu(x') \Lambda(\mathrm{d}x') \int_0^{\infty} \mathrm{e}^{-t} \frac{p_t(Z_{S_1}^1, x')}{u^1(Z_{S_1}^1, x')} \mathrm{d}t \int_0^t r(s) \mathrm{d}s \right) \right].$$

But by (2.10), $\nu(x') \le \frac{1}{\alpha} u^1(Z_{S_1}^1, x')$, since $Z_{S_n}^1 \in C$, thus

$$\begin{aligned} I\!E_x \left[r(S_1) \int_0^{R_1 - S_1} r(s) \mathrm{d}s \right] &\le \frac{1}{\alpha} I\!E_x \left[r(S_1) \left(\int_E \Lambda(\mathrm{d}x') \int_0^{\infty} \mathrm{e}^{-t} p_t(Z_{S_1}^1, x') \mathrm{d}t \int_0^t r(s) \mathrm{d}s \right) \right] \\ &= \frac{1}{\alpha} I\!E_x \left[r(S_1) \int_0^{\infty} \mathrm{e}^{-t} \mathrm{d}t \left(\int_E p_t(Z_{S_1}^1, x') \Lambda(\mathrm{d}x') \right) [\int_0^t r(s) \mathrm{d}s] \right] \\ &= \frac{1}{\alpha} I\!E_x \left[r(S_1) \int_0^{\infty} \mathrm{e}^{-t} \mathrm{d}t \int_0^t r(s) \mathrm{d}s \right] \\ &= \frac{c}{\alpha} I\!E_x (r(S_1)), \end{aligned}$$

since $\int_0^\infty \mathrm{e}^{-t} \int_0^t r(s)\mathrm{d}s\mathrm{d}t < \infty$. Finally, $r(t) \le \int_0^t r(s)\mathrm{d}s + c$ gives

$$\mathbb{E}_x(r(S_1)) \le \mathbb{E}_x \int_0^{S_1} r(s)\mathrm{d}s + c,$$

which is controlled due to (4.35). This concludes the proof. □

Remark 4.3. The fact that the rate function is polynomial was crucial at two points in the above proof: in equations (4.23) and (4.34). The general sub-geometrical case could probably be handled by paying in particular attention to the constants that arrive in expressions like $\mathbb{E}_x r^p(T_n) \le [\mathbb{E}_x r^p(T_1)]^n$.

5. Polynomial deviation inequality

We impose Assumption 2.2 with a function $\Phi(v) = cv^\alpha$, where $0 \le \alpha < 1$. As a consequence, we obtain a control for polynomial moments $\mathbb{E}_x R_1^p$ of the regeneration time for all $p \le 1/(1-\alpha)$, due to (4.17). Since V is continuous and since the measure ν of (2.10) which is used in order to construct the regeneration periods is of compact support, also $\mathbb{E}_\nu R_1^p$ is finite for all $p \le 1/(1-\alpha)$.

In order to derive the deviation inequality we first state a deviation inequality for the counting process associated to the life cycle decomposition

$$N_t = \sup\{n : R_n \le t\} = \sum_{n=1}^\infty 1_{\{R_n \le t\}},\ N_0 = 0.$$

We have almost surely, as $t \to \infty$, $N_t/t \to \mathbb{E}_\Pi N_1 = \ell$, where we recall that

$$\ell = (\mathbb{E}_\nu R_1)^{-1} = (\mathbb{E}(R_2 - R_1))^{-1},$$

see Proposition 3.4 and equation (3.14).

The deviation inequality for the counting process associated to the life cycle decomposition is the following.

Theorem 5.1. *Grant Assumptions 2.1 and 2.2 with $\Phi(v) = cv^\alpha$, $0 \le \alpha < 1$. Let $x \in E$ be any starting point and $0 < \varepsilon < 1$. Then for any $1 < p \le 1/(1-\alpha)$ there exists a positive constant $C(\ell, p, \nu)$ such that the following inequality holds:*

$$\mathbb{P}_x\left(\left|\frac{N_t}{t} - \ell\right| > \ell\varepsilon\right) \le C(\ell, p, \nu)V(x)\ \frac{1}{\varepsilon^p \wedge \varepsilon^{2(p-1)}}\ \frac{1}{t^{p-1}}. \tag{5.37}$$

Here $C(\ell, p, \nu)$ is given by

$$C(\ell, p, \nu) = \left\{ \begin{array}{ll} C(p)\left[(1 + (1/\ell)^{p-1}) + (m_p + \sigma^{2(p-1)})\right]\left(\ell^{p-1} \vee \ell\right) & \text{if } p \ge 2 \\ C(p)\left[(1 + (1/\ell)^{p-1}) + m_p \ell C_p^p\right] & \text{if } p \in]1, 2[\end{array} \right\},$$

where $C(p)$ is a constant depending only on p, C_p is the constant of the Burkholder–Davis–Gundy inequality, $m_p = \mathbb{E}(R_2 - R_1 - \frac{1}{\ell})^p$ and $\sigma^2 = \mathrm{Var}(R_2 - R_1)$, in the case $p \ge 2$.

Proof. The proof is basically the same as in Löcherbach *et al.* [16], proof of Theorem 3.1. Put in contrary to there we use the Fuk–Nagaev inequality given in the appendix (Thm. A.1) in the case $p \ge 2$. We decompose:

$$\mathbb{P}_x\left(|N_t/t - \ell| > \ell\varepsilon\right) \le \mathbb{P}_x\left(N_t/t > \ell(1+\varepsilon)\right) + \mathbb{P}_x\left(N_t/t < \ell(1-\varepsilon)\right). \tag{5.38}$$

Put for $k \ge 1$, $\bar\eta_k = -1(R_{k+1} - R_k - 1/\ell)$. For the first term of (5.38), we have

$$\mathbb{P}_x\left(N_t/t > \ell(1+\varepsilon)\right) \le \mathbb{P}_x\left(R_1 - 1/\ell \le -t\varepsilon/2\right) + \mathbb{P}_x\left(\sum_{k=1}^{[t\ell(1+\varepsilon)]} \bar\eta_k \ge t\varepsilon/2\right). \tag{5.39}$$

In an analogous way,

$$I\!P_x\left(N_t/t<\ell(1-\varepsilon)\right)\le I\!P_x\left(R_1-\frac{1}{\ell}\ge t\varepsilon/2\right)+I\!P_x\left(\sum_{k=1}^{[t\ell(1-\varepsilon)]-1}\bar\eta_k\le -t\varepsilon/2\right). \tag{5.40}$$

The random variables $\bar\eta_k, k\ge 1$, are identically distributed centred random variables such that $I\!E_x|\bar\eta_k|^p<\infty$. Moreover, they are two-dependent. Indeed, $\bar\eta_k$ is not independent of $\mathcal{F}_{R_k}$, but only independent of $\mathcal{F}_{R_{k-1}}$. This is due to step 1 of the construction of Z, where the waiting time for the new jump is chosen depending on the actual value of Z at time R_k.

If $p\ge 2$, we can apply Theorem A.1. Let $M_0=0$ and $M_n=\sum_{k=1}^n\bar\eta_k$. Denote $M_n^*=\sup_{k\le n}|M_k|$. As a consequence of (5.39) and (5.40) we can write

$$I\!P_x\left(|N_t/t-\ell|>\ell\varepsilon\right)\le I\!P_x\left(|R_1-1/\ell|\ge t\varepsilon/2\right)+I\!P_x\left(M^*_{[t\ell(1+\varepsilon)]}\ge t\varepsilon/2\right). \tag{5.41}$$

We use Theorem A.1 with $n=[t\ell(1+\varepsilon)]$ and $\lambda=t\varepsilon/8$ and obtain

$$\begin{aligned}I\!P_x\left(|N_t/t-\ell|>\ell\varepsilon\right)&\le\frac{2^{p-1}I\!E_x|R_1-1/\ell|^{p-1}}{(t\varepsilon)^{p-1}}+\ C(p)[m_p+\sigma^{2(p-1)}]\ \left(\ell^{p-1}\vee\ell\right)\varepsilon^{-2(p-1)}t^{-(p-1)}\\&\le\left(2^{p-1}I\!E_x|R_1-1/\ell|^{p-1}+C(p)[m_p+\sigma^{2(p-1)}]\ \left(\ell^{p-1}\vee\ell\right)\right)\frac{1}{\varepsilon^{2(p-1)}}t^{-(p-1)},\end{aligned}$$

since $\varepsilon<1$, where $m_p=I\!E_x|\bar\eta_1|^p$, $\sigma^2=\mathrm{Var}(\bar\eta_1)$. Finally we use that

$$I\!E_x|R_1-1/\ell|^{p-1}\le C(p)[I\!E_xR_1^{p-1}+(1/\ell)^{p-1}],$$

and that for some constants c and d,

$$I\!E_x(R_1^{p-1})\le 1+I\!E_xR_1^p\le cV(x)+d$$

to conclude that, since $V(.)\ge 1$,

$$I\!P_x\left(|N_t/t-\ell|>\ell\varepsilon\right)\le C(p)\,V(x)\left((1+(1/\ell)^{p-1})+[m_p+\sigma^{2(p-1)}]\ \left(\ell^{p-1}\vee\ell\right)\right)\frac{1}{\varepsilon^{2(p-1)}}t^{-(p-1)}.$$

This finishes the proof in the case $p\ge 2$.

In the case $1<p<2$, we apply the Burkholder–Davis–Gundy inequality. In order to produce independent random variables, we define

$$\eta_k^{(1)}=\begin{Bmatrix}\bar\eta_k & \text{if } k \text{ odd}\\ 0 & \text{elseif}\end{Bmatrix},\ \eta_k^{(2)}=\begin{Bmatrix}\bar\eta_k & \text{if } k \text{ even}\\ 0 & \text{elseif}\end{Bmatrix}. \tag{5.42}$$

Let $M_0^1=0$ and $M_n^1=\sum_{k=1}^n\eta_k^{(1)}$. In the same way, $M_0^2=0$ and $M_n^2=\sum_{k=1}^n\eta_k^{(2)}$.

We also introduce the following two sub-filtrations, associated to the sum of odd and the sum of even terms. Let

$$\mathcal{A}_n^{(1)}:=\sigma\left\{\eta_k^{(1)}:k\le n,k\text{ odd }\right\}=\sigma\left\{M_k^{(1)},k\le n\right\},$$

and

$$\mathcal{A}_n^{(2)}:=\sigma\left\{\eta_k^{(2)}:k\le n,k\text{ even }\right\}=\sigma\left\{M_k^{(2)},k\le n\right\}.$$

Then $(M_n^1)_n$ and $(M_n^2)_n$ are discrete $\mathcal{A}_n^{(1)}$-martingales ($\mathcal{A}_n^{(2)}$-martingales, respectively). Both martingales are L^p martingales such that $[M^{(i)}]_n=\sum_{k=1}^n(\eta_k^{(i)})^2$, for $i=1,2$. Denote $(M^{(i)})_n^*=\sup_{k\le n}|M_k^{(i)}|$, $i=1,2$. In this case, as a consequence of (5.39) and (5.40) we write

$$I\!P_x\left(|N_t/t-\ell|>\ell\varepsilon\right)\le I\!P_x\left(|R_1-1/\ell|\ge t\varepsilon/2\right)+I\!P_x\left((M^{(1)})^*_{[t\ell(1+\varepsilon)]}\ge t\varepsilon/4\right)+I\!P_x\left((M^{(2)})^*_{[t\ell(1+\varepsilon)]}\ge t\varepsilon/4\right). \tag{5.43}$$

We use the Burkholder–Davis–Gundy inequality to bound the last terms in (5.43): for all $p > 1$ there exists a constant C_p depending only p such that $\|(M^{(i)})^*_n\|_p \le C_p\|[M^{(i)}]^{1/2}_n\|_p$, hence $\mathbb{E}_x((M^{(i)})^*_n)^p \le C_p^p \mathbb{E}_x\left(\sum_{k=1}^n(\eta_k^{(i)})^2\right)^{p/2}$.

Notice that by definition, the term $\sum_{k=1}^n(\eta_k^{(1)})^2$ contains $[\frac{n+1}{2}]$ terms whereas $\sum_{k=1}^n(\eta_k^{(2)})^2$ contains $[n/2]$ terms. Since $1 < p < 2$, the sub-additivity of the function $x \mapsto x^{p/2}$ implies

$$\left(\sum_{k=1}^n(\eta_k^{(1)})^2\right)^{p/2} \le \sum_{k=1}^n|\bar\eta_k^{(1)}|^p, \quad \text{hence} \quad \mathbb{E}_x((M^{(1)})^*_n)^p \le C_p^p n \mathbb{E}|\bar\eta_1|^p. \tag{5.44}$$

The same kind of bound holds also for the even terms.

Now we can conclude similarly to Löcherbach *et al.* [16]: For $1 < p < 2$,

$$\begin{aligned}\mathbb{P}_x\left(|N_t/t-\ell| > \ell\varepsilon\right) &\le \frac{2^{p-1}\mathbb{E}_x|R_1-1/\ell|^{p-1}}{(t\varepsilon)^{p-1}} + 2\, 4^p C_p^p \mathbb{E}_x|\bar\eta_1|^p\,[t\ell(1+\varepsilon)]\frac{1}{(t\varepsilon)^p}\\ &\le \left(2^{p-1}\mathbb{E}_x|R_1-1/\ell|^{p-1} + 2^{2p+2}C_p^p\mathbb{E}_x|\bar\eta_1|^p\,\ell\right)\frac{1}{\varepsilon^p}\frac{1}{t^{p-1}}\\ &\le C(p)V(x)\left((1+(1/\ell)^{p-1}) + m_p\ell C_p^p\right)\frac{1}{\varepsilon^p}\frac{1}{t^{p-1}}.\end{aligned}$$

This concludes the proof. □

Once the deviation inequality for the counting process $(N_t)_t$ is proven, we obtain on the lines of Löcherbach *et al.* [16], Theorem 3.2, the following general deviation inequality for additive functionals of the original Markov process X, built of bounded functions.

Theorem 5.2. *Grant Assumptions 2.1 and 2.2 with $\Phi(v) = cv^\alpha$, $0 \le \alpha < 1$. Put $p = 1/(1-\alpha)$. Let $f \in L^1(\mu)$. Suppose that $\|f\|_\infty < \infty$. Let x be any initial point and $0 < \varepsilon < \|f\|_\infty$. Then for all $t \ge 1$ the following inequality holds:*

$$P_x\left(\left|\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s - \mu(f)\right| > \varepsilon\right) \quad \le \quad K(\ell,p,\nu,X)\,V(x)\,t^{-(p-1)}\left\{\begin{matrix}\frac{1}{\varepsilon^{2(p-1)}}\|f\|_\infty^{2(p-1)} & \text{if } p \ge 2\\ \frac{1}{\varepsilon^p}\|f\|_\infty^p & \text{if } 1 < p < 2\end{matrix}\right\}. \tag{5.45}$$

Here $K(\ell,p,\nu,X)$ is a positive constant, different in the two cases, which depends on ℓ, p, ν and on the process X through the life cycle decomposition, but which does not depend on f, t, ε.

Remark 5.3. The above result holds for any starting measure ν such that $\mathbb{E}_\nu(R_1^p)$ is finite, so a fortiori for any measure ν such that $V \in L^1(\nu)$. In contrary to most of the existing results in the literature (see *e.g.* Cattiaux and Guillin [4]) we do not need to assume absolute continuity of the initial law of the process with respect to the invariant measure μ.

Proof. First of all, since the law of X starting from a fixed point x is the same as the law of Z^1 starting from the initial measure $\mathbb{P}_x$, we certainly have that

$$P_x\left(\left|\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s - \mu(f)\right| > \varepsilon\right) = \mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t f(Z_s^1)\mathrm{d}s - \mu(f)\right| > \varepsilon\right).$$

Now for $p < 2$ the rest of the proof is exactly the same as the proof of Theorem 3.2 in Löcherbach *et al.* [16]. The only difference compared to there is that the variables $\xi_n = \int_{R_n}^{R_{n+1}}(f-\mu(f))(Z_s^1)\mathrm{d}s$ are no longer independent but only 2-dependent. Hence, the same trick as in the proof of Theorem 5.1 applies: one has to separate even and odd terms. But this does only change the constants in the upper bound.

For $p \geq 2$, we use the Fuk–Nagaev inequality again. We start as in the proof of Theorem 3.2 of Löcherbach *et al.* [16]. Denote

$$\Omega_t = \left\{ \left| \frac{N_t}{t} - \ell \right| \leq \ell\delta \right\}, \quad \delta = \varepsilon / \|f\|_\infty < 1.$$

Put $\bar{f} := f - \mu(f)$. Then

$$\begin{aligned}
&\mathbb{P}_x \left(\left| \int_0^t f(Z_s^1) \mathrm{d}s - t\mu(f) \right| > t\varepsilon \right) \\
&\leq \mathbb{P}_x \left(\left| \int_0^{R_1} \bar{f}(Z_s^1) \mathrm{d}s \right| > \frac{t\varepsilon}{3} \right) + \mathbb{P}_x \left(\left| \int_{R_1}^{R_{N_t+1}} \bar{f}(Z_s^1) \mathrm{d}s \right| > \frac{t\varepsilon}{3} \; ; \Omega_t \right) \\
&\quad + \mathbb{P}_x \left(\left| \int_t^{R_{N_t+1}} \bar{f}(Z_s^1) \mathrm{d}s \right| > \frac{t\varepsilon}{3} ; \Omega_t \right) + \mathbb{P}_x \left(\Omega_t^c \right) \\
&= A + B + C + D.
\end{aligned}$$

The terms A and C are handled as in Löcherbach *et al.* [16]. Term D is controlled thanks to Theorem 5.1. So the main term that has to be controlled is the term B, and we have

$$B \leq \mathbb{P}_x \left(\sum_{k=1}^{[t\ell(1+\delta)]} |\xi_k| \geq t\varepsilon/3 \right), \;\; \xi_k = \int_{R_k}^{R_{k+1}} \bar{f}(Z_s^1) \mathrm{d}s.$$

Put $\bar{\xi}_k = \frac{1}{\|f\|_\infty} \xi_k$, then $\bar{\xi}_k = \int_{R_k}^{R_{k+1}} (g - \mu(g))(Z_s^1) \mathrm{d}s$, where $g = f / \|f\|_\infty$, $\|g\|_\infty = 1$. We write

$$\mathbb{P}_x \left(\sum_{k=1}^{[t\ell(1+\delta)]} |\xi_k| \geq t\varepsilon/3 \right) = \mathbb{P}_x \left(\sum_{k=1}^{[t\ell(1+\delta)]} |\bar{\xi}_k| \geq \frac{t\varepsilon}{3\|f\|_\infty} \right) \leq \mathbb{P}_x \left(\sup_{k \leq [t\ell(1+\delta)]} S_k \geq t\delta/3 \right), \;\; \delta = \varepsilon / \|f\|_\infty < 1,$$

$S_k = \sum_{i-1}^k |\bar{\xi}_k|$, and apply the Fuk–Nagaev inequality of Theorem A.1 with $n = t\ell(1+\delta)$ and $\lambda = t\delta/12$. This gives the following upper bound

$$\mathbb{P}_x \left(\sum_{k=1}^{[t\ell(1+\delta)]} |\xi_k| \geq t\varepsilon/3 \right) \leq C(p)[m_p + \sigma^{2(p-1)}] \left(\ell^{p-1} \vee \ell \right) \delta^{-2(p-1)} t^{-(p-1)},$$

where

$$m_p = \mathbb{E}(|\bar{\xi}_1|^p) \leq 2^p \, \mathbb{E}((R_2 - R_1)^p), \sigma^2 = Var(\bar{\xi}_1) \leq 4 \, \mathbb{E}((R_2 - R_1)^2.$$

Therefore, since $\delta = \varepsilon / \|f\|_\infty$, there exists a constant $K(\ell, p, \nu, X)$ depending only on the process and the life cycle decomposition, but not on the function f, such that

$$\mathbb{P}_x \left(\sum_{k=1}^{[t\ell(1+\delta)]} |\xi_k| \geq t\varepsilon/3 \right) \leq K(\ell, p, \nu, X) \|f\|_\infty^{2(p-1)} \varepsilon^{-2(p-1)} t^{-(p-1)}.$$

This finishes the proof. □

6. Examples

We close our paper with two examples where the above deviation inequalities can be applied.

6.1. Multi-dimensional diffusions

Consider the solution of the following stochastic differential equation in $\mathbb{R}^d$

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t,$$

where W_t is an n-dimensional Brownian motion, $n \geq d$, such that b is a locally bounded Borel measurable function $\mathbb{R}^d \to \mathbb{R}^d$ and σ is a bounded continuous function $\mathbb{R}^d \to \mathbb{R}^{d\times n}$ which is uniformly elliptic: Writing $a := \sigma\sigma^*$, we suppose that there exists $\varepsilon > 0$ such that

$$\langle a(x)\xi, \xi\rangle \ \geq \varepsilon \|\xi\|^2$$

for all $x \in \mathbb{R}^d$. Classical results on lower bounds for transition densities of diffusions (see for instance Kusuoka and Stroock [12]) imply that in this case any compact set of $\mathbb{R}^d$ is petite. We cite the following recurrence conditions from Fort and Roberts [9]. Suppose there exist $M, \beta, \gamma > 0$ and $l < 2$ such that

$$\sup_{x:\|x\|>M} \|x\|^{-(2+l)}\langle x, a(x)x\rangle = \beta, \qquad \sup_{x:\|x\|>M} \|x\|^{-l}\mathrm{tr}(a(x)) = \gamma,$$

$$\sup_{x:\|x\|>M} \|x\|^{-l}\langle b(x), x\rangle = -r, \text{ for some } r > (\gamma - \beta l)/2.$$

We choose

$$\kappa \in \left]0, l + \frac{2r-\gamma}{\beta}\right[$$

and put $m = 2-l+\kappa$, thus $2-m = l-\kappa$. Let $V(x) = \|x\|^m$ outside a compact set. Then $\sup_{x:\|x\|>M} \mathcal{A}V(x) < \infty$, and standard calculus shows that for all $\|x\| > M$,

$$\mathcal{A}V(x) \leq m\left(-r + \frac{1}{2}[\gamma + (m-2)\beta]\right)\frac{V(x)}{\|x\|^{2-l}}.$$

Then by our choice of κ, $\tilde{r} := r - \frac{1}{2}[\gamma + (m-2)\beta] > 0$. Hence for $\|x\| > M$,

$$\mathcal{A}V(x) \leq -\Phi \circ V(x),$$

where

$$\Phi(x) = m\tilde{r}\, x^{1-\alpha}, \quad \text{with} \quad \alpha = \frac{2-l}{m} < 1.$$

Hence we get polynomial moments of regeneration times up to the order $m/(2-l) = 1 + \kappa/(2-l)$.

6.2. Solutions to SDE's driven by a jump noise

This chapter is inspired by a recent work of Kulik [11] on exponential ergodicity for solutions to SDE's driven by a jump noise. More precisely, consider the solution of the following stochastic differential equation on $\mathbb{R}^d$ driven by a jump noise

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \int_{\|u\|\leq 1} c(X_{s-}, u)\tilde{\mu}(\mathrm{d}t, \mathrm{d}u) + \int_{\|u\|>1} c(X_{s-}, u)\mu(\mathrm{d}t, \mathrm{d}u). \tag{6.46}$$

Here μ is a Poisson random measure (PRM) on $\mathbb{R}_+ \times \mathbb{R}^q$, having compensator $\hat{\mu}(\mathrm{d}t, \mathrm{d}u) = \mathrm{d}t\nu(\mathrm{d}u)$, and $\tilde{\mu}(\mathrm{d}t, \mathrm{d}u) = \mu(\mathrm{d}t, \mathrm{d}u) - \mathrm{d}t\nu(\mathrm{d}u)$ denotes the compensated PRM. We follow Kulik [11] and impose the following conditions on the coefficients b and c. The drift function b belongs to $C^1(\mathbb{R}^d, \mathbb{R}^d)$ and satisfies a linear growth condition. The jump rate $c(x, u)$ is one times continuously differentiable with respect to x. Moreover,

$$\|c(x,u) - c(y,u)\| \leq K(1 + \|u\|)\|x - y\|, \quad \|c(x,u)\| \leq \psi(x)\|u\|, \ x, y \in \mathbb{R}^d, u \in \mathbb{R}^q,$$

where K is some constant and where $\psi : I\!R^d \to I\!R_+$ satisfies a linear growth condition. Finally we impose a moment condition on the Lévy measure ν. For all $R > 0$,

$$\int \sup_{x:\|x\|\leq R} (\|c(x,u)\| + \|\nabla_x c(x,u)\|)\,\nu(\mathrm{d}u) < +\infty.$$

Then for any fixed $x \in I\!R^d$, there exists a unique strong solution X_t to (6.46), which is a strong Markov process, having càdlàg trajectories.

We quote sufficient conditions implying that compact sets are petite from Kulik [11]. For this sake, we have to introduce some notation. Let $\mathcal{S}^q = \{v \in I\!R^q : \|v\| = 1\}$ be the unit sphere in $I\!R^q$. For any $w \in \mathcal{S}^q$ and for any $\varrho \in]0,1[$, let $V_+(w,\varrho) = \{y \in I\!R^q : \langle y, w\rangle \geq \varrho\|y\|\}$, and $V(w,\varrho) = \{y \in I\!R^q : |\langle y, w\rangle| \geq \varrho\|y\|\}$. Then Kulik [11] obtains the following result.

Proposition 6.1 ([11])**.** *Suppose that the following assumptions hold.*

1. *Cone condition: for every $w \in \mathcal{S}^q$, there exists $\varrho \in]0,1[$, such that for every $\delta > 0$,*

$$\nu\left(V(w,\varrho) \cap \{u : \|u\| \leq \delta\}\right) > 0.$$

2. *Non-degeneracy condition: there exists a point $x_* \in I\!R^d$ and a neighbourhood O_* of x_* such that $c(x,u) = \chi(x)u + \delta(x,u)$, for all $x \in O_*$, and*

$$\|\delta(x_*,u)\| + \|\nabla_x \delta(x_*,u)\| = o(\|u\|),$$

as $\|u\| \to 0$. Moreover, the functions $\tilde{b}(.) = b(.) - \int_{\|u\|\leq 1} c(.,u)\nu(du)$ and χ are one times continuously differentiable and satisfy the joint non-degeneracy condition

$$rank\left(\nabla\tilde{b}(x_*)\chi(x_*) - \nabla\chi(x_*)\tilde{b}(x_*)\right) = d.$$

3. *Support condition: for any $R > 0$ there exists t such that for all x with $\|x\| \leq R$,*

$$x_* \in \operatorname{supp} P_t(x,\cdot).$$

If the above conditions hold, then any compact set is petite.

Remark 6.2.

1. In the one-dimensional case $d = q = 1$, the above conditions can be stated in a simpler way. For example, condition 1. can then be written as follows: for all $\delta > 0$, $\nu(u : 0 < \|u\| \leq \delta) > 0$;
2. Simon [21], Theorem I, gives a sufficient condition for condition 3. above to hold, see also Proposition 4.7 in Kulik [11].

Proof. Theorem 1.3, Proposition 4.3 and 4.4 of Kulik [11] show that under the above conditions, the following Dobrushin condition holds: for all $R > 0$, there exists $t^* = t^*(R)$ such that

$$\inf_{x,y:\|x\|,\|y\|\leq R} \int \left[P_{t^*}(x,\cdot) \wedge P_{t^*}(y,\cdot)\right](\mathrm{d}z) > 0, \tag{6.47}$$

where for any two probability measures P and Q,

$$[P \wedge Q](\mathrm{d}z) := \left(\frac{\mathrm{d}P}{\mathrm{d}(P+Q)}(z) \wedge \frac{\mathrm{d}Q}{\mathrm{d}(P+Q)}(z)\right)(P+Q)(\mathrm{d}z).$$

From this the claim follows since (6.47) implies that any compact set is a petite set. □

It remains to give conditions that are sufficient for the recurrence condition (2.4), (2.5) respectively. There is a wide range of possible conditions and in what follows we restrict attention to a particular sufficient condition which is stated in the same spirit as the conditions of Proposition 4.1 of Kulik [11].

Proposition 6.3. *Suppose that the conditions of Proposition 6.1 hold. Suppose moreover that there exist* $M, \gamma > 0$ *and* $0 < l < 1$ *such that*

1. *Moment-condition: there exists* $m \geq 1$ *such that* $\int_{\|u\| \geq 1} \|u\|^m \nu(\mathrm{d}u) < \infty$.
2. *Moderate jumps: the function* c *can be decomposed into* $c = c_1 + c_2$ *such that*
   (*a*) $\|c_1(x,u)\| \leq \gamma \|x\|^l \|u\|, u \in \mathbb{R}^q, \|x\| > M$.
   (*b*) $\|x + c_2(x,u)\| \leq \|x\|$, $\|u\| > 1$, $\|x\| > M$, *and* $c_2(\cdot, u) = 0$, *if* $\|u\| \leq 1$.
3. *Drift-condition:* $\sup_{x:\|x\|>M} \|x\|^{-(1+l)} \langle b(x), x\rangle = -r$, *for some constant* r *satisfying* $r > 2\gamma \int_{\|u\|>1} \|u\|^m \nu(\mathrm{d}u)$.

Then there exists $M_0 \geq M$ *such that* (2.5) *holds with* $B = \{x : \|x\| \leq M_0\}$, B *petite,* $V(x) = \|x\|^m$ *and* $\Phi(x) = c x^{1-\alpha}$, *where* $\alpha = \frac{1-l}{m} < 1$.

Proof. We use the drift condition for the generator defined for all functions F in the extended domain of the generator

$$\mathcal{A}F(x) = \langle b(x), \nabla F(x)\rangle + \int_{\mathbb{R}^q} \left(F(x + c(x,u)) - F(x) - 1_{\{\|u\| \leq 1\}} \langle \nabla F(x), c(x,u)\rangle \right) \nu(\mathrm{d}u).$$

Applying this to $V(x) = \|x\|^m$ yields for all $\|x\| > M$,

$$\begin{aligned} \mathcal{A}V(x) &= m\langle b(x), x\rangle \|x\|^{m-2} + \int_{\|u\|>1} \left(\|x + c(x,u)\|^m - \|x\|^m\right) \nu(\mathrm{d}u) \\ &\quad + \int_{\|u\| \leq 1} \left(\|x + c(x,u)\|^m - \|x\|^m - m\langle x, c(x,u)\rangle \|x\|^{m-2} \right) \nu(\mathrm{d}u) \\ &\leq -m \cdot r \|x\|^{m-1+l} + \int_{\|u\|>1} \left(\|x + c(x,u)\|^m - \|x\|^m\right) \nu(\mathrm{d}u) \\ &\quad + \int_{\|u\| \leq 1} \left(\|x + c(x,u)\|^m - \|x\|^m - m\langle x, c(x,u)\rangle \|x\|^{m-2} \right) \nu(\mathrm{d}u). \end{aligned} \tag{6.48}$$

We start with the term in the last line. By Taylor's formula, writing $h = c(x,u) = c_1(x,u)$, since $\|u\| \leq 1$, we certainly have that

$$\begin{aligned} \Big| \, \|x + c(x,u)\|^m - \|x\|^m - m\langle x, c(x,u)\rangle \|x\|^{m-2} \, \Big| &\leq \frac{1}{2} \sup_{y \in]x, x+h[} |\langle h, \nabla^2 V(y) h\rangle| \\ &\leq \frac{1}{2} m \left[1 + |m-2|\right] \|h\|^2 \sup_{y \in]x, x+h[} \|y\|^{m-2}. \end{aligned}$$

Here, $]x, x+h[$ denotes the d-dimensional interval $]x_1, x_1 + h_1[\times \ldots \times]x_d, x_d + h_d[$.

Applying condition 2. (a) to $h = c_1(x,u)$, where $\|u\| \leq 1$, yields

$$\|h\|^2 \leq \gamma^2 \|x\|^{2l} \|u\|^2.$$

If $m - 2 > 0$, we choose $M_0 \geq M$ such that $(1 + \gamma M_0^{l-1})^{m-1} \leq 2$ (recall that $l < 1$). Then we obtain

$$\begin{aligned} \sup_{y \in]x, x+h[} \|y\|^{m-2} = \|x + h\|^{m-2} &\leq \|x\|^{m-2} \left[1 + \gamma \|x\|^{l-1}\right]^{m-2} \\ &\leq \|x\|^{m-2} \left[1 + \gamma M_0^{l-1}\right]^{m-2} \\ &\leq 2\|x\|^{m-2}. \end{aligned}$$

If $m < 2$, we can proceed similarly,

$$\begin{aligned}\sup_{y\in]x,x+h[} \|y\|^{m-2} &\le \|x\|^{m-2}\left[1-\gamma\|x\|^{l-1}\right]^{m-2}\\ &\le \|x\|^{m-2}\left[1-\gamma M_0^{l-1}\right]^{m-2}\\ &\le 2\|x\|^{m-2},\end{aligned}$$

where we choose M_0 such that $(1-\gamma M_0^{l-1})^{m-2} \le 2$.

As a consequence, for any $\|x\| \ge M_0$, the last line of (6.48) is bounded from above by

$$m\left([1+|m-2|]\gamma^2\int_{\|u\|\le 1}\|u\|^2\nu(\mathrm{d}u)\right)\|x\|^{m-2+2l} \le C\ M_0^{l-1}\|x\|^{m-1+l}, \tag{6.49}$$

since $\|x\|^{l-1} \le M_0^{l-1}$. Here, $M_0^{l-1} \to 0$ as $M_0 \to \infty$, and C is some constant. Hence the last term of (6.48) will be neglectable for our purposes.

Concerning the first jump term in (6.48) we proceed as Kulik [11], proof of Proposition 4.1: for $\|u\| > 1$, using condition 2. (b), we have

$$\|x+c(x,u)\|^m - \|x\|^m \le \|x+c(x,u)\|^m - \|x+c_2(x,u)\|^m = \|x(u)+c_1(x,u)\|^m - \|x(u)\|^m,$$

where $x(u) = x + c_2(x,u)$, and then, applying Taylor's formula,

$$\|x(u)+c_1(x,u)\|^m - \|x(u)\|^m \le m\|c_1(x,u)\| \sup_{y\in]x(u),x(u)+c_1(x,u)[}\|y\|^{m-1}.$$

Now, since $m \ge 1$, we argue as before and obtain, using successively condition 2. (a) and 2. (b) and $\|u\| > 1$,

$$\begin{aligned}m\|c_1(x,u)\| \sup_{y\in]x(u),x(u)+c_1(x,u)[}\|y\|^{m-1} &\le m\gamma\|x\|^l\|u\|\left(\|x(u)\|+\gamma\|x\|^l\|u\|\right)^{m-1}\\ &\le m\gamma\|x\|^l\|u\|\left(\|x\|+\gamma\|x\|^l\|u\|\right)^{m-1}\\ &\le m\gamma\|x\|^{m-1+l}\|u\|^m\left(1+\gamma M_0^{l-1}\right)^{m-1}\\ &\le 2m\gamma\|x\|^{m-1+l}\|u\|^m,\end{aligned}$$

by the choice of M_0. As a consequence, the first jump term in (6.48) can be upper bounded as follows:

$$\int_{\|u\|>1}\left(\|x+c(x,u)\|^m - \|x\|^m\right)\nu(\mathrm{d}u) \le m\|x\|^{m-1+l}\left[2\gamma\int_{\|u\|>1}\|u\|^m\nu(\mathrm{d}u)\right].$$

Collecting all the above results, we finally obtain that for all $\|x\| \ge M_0$,

$$\mathcal{A}V(x) \le m\left(-r+2\gamma\int_{\|u\|>1}\|u\|^m\nu(\mathrm{d}u)+CM_0^{l-1}\right)\frac{V(x)}{\|x\|^{1-l}}.$$

By condition 3, for M_0 sufficiently large, $-r+2\gamma\int_{\|u\|>1}\|u\|^m\nu(\mathrm{d}u)+CM_0^{l-1} < 0$ eventually, and this implies the assertion. □

Appendix

For the convenience of the reader we give in this section a Fuk–Nagaev inequality for sums of two-dependent identically distributed centred random variables admitting a moment of order p. This inequality is the key tool to our deviation inequalities, and its proof can be found in the book of Rio [20].

Let $X_1, X_2, \ldots$ be centred identically distributed random variables which are two-dependent and such that $E(|X_1|^p) < \infty$ for some $p \ge 2$. Put $S_k := X_1 + \ldots + X_k$. Then the following deviation inequality holds.

Theorem A.1. *For any $\lambda > 0$ we have that*

$$P\left(\sup_{k\le n}|S_k|\ge 4\lambda\right)\le C(p)\left(\sigma^{2(p-1)}\lambda^{-2(p-1)}n^{p-1}+m_p\, n\,\lambda^{-p}\right),$$

where $\sigma^2 = \operatorname{Var}(X_1)$ and $m_p = E(|X_1|^p)$.

Proof. We use the Fuk–Nagaev inequality presented in Rio [20], Theorem 6.2. First of all, since the variables are two-dependent, we certainly have the upper bound on the α-mixing coefficients $\alpha_n = \sup_{k\ge n}\alpha(\sigma(X_1),\sigma(X_{k+1}))$:

$$\alpha_0 = \frac{1}{2},\quad \alpha_1 \le \frac{1}{2},\quad \alpha_n \equiv 0 \text{ for all } n\ge 2.$$

Hence, using the notation (1.21) of Rio [20], we can upper bound

$$\alpha^{-1}(u)\le 2\, 1_{[0,\frac{1}{2}[}(u).$$

As a consequence the expression $R(u)$ of Theorem 6.2 of Rio [20] is given as

$$R(u)\le 2Q(u)1_{[0,\frac{1}{2}[}(u)\le 2Q(u),$$

where $Q(u)=\inf\{x : H_{X_1}(x)\le u\}$ is the quantile of $|X_1|$, $H_{X_1}(t) = P(|X_1|>t)$. But since X_1 admits a $p-$th moment, we certainly have that

$$H_{X_1}(t)\le m_p\, t^{-p},$$

by Markov's inequality (recall that $m_p = E(|X_1|^p)$). Since Q is the generalised inverse function of H_{X_1}, this implies that

$$Q(u)\le m_p^{1/p}\, u^{-1/p},$$

and this in turn leads to

$$H(u) = R^{-1}(u)\le 2^p\, m_p u^{-p}.$$

Now we can apply (6.5) of Rio [20]. First of all notice that by the two-dependency structure

$$s_n^2 = \sum_{i=1}^n\sum_{j=1}^n|\operatorname{Cov}(X_i,X_j)|\le 3n\sigma^2.$$

Thus we obtain, for any $r\ge 1$,

$$\begin{aligned}
P\left(\sup_{k\le n}|S_k|\ge 4\lambda\right) &\le 4\left(1+\frac{\lambda^2}{rs_n^2}\right)^{-r/2} + 4n\lambda^{-1}\int_0^{H(\lambda/r)}Q(u)\mathrm{d}u\\
&\le 4\left(1+\frac{\lambda^2}{rs_n^2}\right)^{-r/2} + 4n\lambda^{-1}\int_0^{H(\lambda/r)}m_p^{1/p}u^{-1/p}\mathrm{d}u\\
&= 4\left(1+\frac{\lambda^2}{rs_n^2}\right)^{-r/2} + \frac{4n\lambda^{-1}\, m_p^{1/p}}{(1-\frac{1}{p})}H(\lambda/r)^{1-\frac{1}{p}}\\
&\le 4\left(1+\frac{\lambda^2}{rs_n^2}\right)^{-r/2} + C(p)m_p\, n\,\lambda^{-p}\, r^{p-1}\\
&\le 4\left(1+\frac{\lambda^2}{3nr\sigma^2}\right)^{-r/2} + C(p)\, m_p\, n\,\lambda^{-p}\, r^{p-1}.
\end{aligned}$$

Here, $C(p)$ is a constant depending only on p. Now we choose $r = 2(p-1)$. By assumption on p, $r \geq 1$. Finally we get

$$P\left(\sup_{k\leq n} |S_k| \geq 4\lambda\right) \leq C(p)\left(\sigma^{2(p-1)}\lambda^{-2(p-1)}n^{p-1} + m_p\, n\, \lambda^{-p}\right).$$

□

Acknowledgements. Eva Löcherbach has been partially supported by the ANR project ANR-08-BLAN-0220-01.

References

[1] R. Adamczak, A tail inequality for suprema of unbounded empirical processes with applications to Markov chains. *Electron. J. Probab.* **13** (2008) 1000–1034.
[2] K.B. Athreya and P. Ney, A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245** (1978) 493–501.
[3] P. Bertail and S. Clémençon, Sharp bounds for the tails of functionals of markov chains, *Teor. Veroyatnost. i Primenen* **54** (2009) 609–619.
[4] P. Cattiaux and A. Guillin, Deviation bounds for additive functionals of Markov processes. *ESAIM: PS* **12** (2008) 12–29.
[5] J.-R. Chazottes, P. Collet, C. Külske and F. Redig, Concentration inequalities for random fields via coupling. *Probab. Theory Relat. Fields* **137** (2007) 201–225.
[6] S. Clémençon, Moment and probability inequalities for sums of bounded additive functionals of regular Markov chains *via* the Nummelin splitting technique. *Stat. Probab. Lett.* **55** (2001) 227–238.
[7] R. Douc, G. Fort and A. Guillin, Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stoch. Proc. Appl.* **119** (2009) 897–923.
[8] R. Douc, G. Fort, E. Moulines and P. Soulier, Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.* **14** (2004) 1353–1377.
[9] G. Fort and G.O. Roberts, Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.* **15** (2005) 1565–1589.
[10] A. Guillin, C. Léonard, L. Wu and N. Yao, Transportation-information inequalities for Markov processes. *Probab. Theory Relat. Fields* **144** (2009) 669–695.
[11] A.M. Kulik, Exponential ergodicity of the solutions to SDE's with a jump noise. *Stoch. Proc. Appl.* **119** (2009) 602–632.
[12] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus. III. *J. Fac. Sci., Univ. Tokyo, Sect. I A* **34** (1987) 391–442.
[13] R. Höpfner and E. Löcherbach, Limit theorems for null recurrent Markov processes. *Memoirs AMS* **161** (2003).
[14] P. Lezaud, Chernoff and Berry-Esséen inequalities for Markov processes. *ESAIM: PS* **5** (2001) 183–201.
[15] E. Löcherbach and D. Loukianova, On Nummelin splitting for continuous time Harris recurrent Markov processes and application to kernel estimation for multi-dimensional diffusions. *Stoch. Proc. Appl.* **118** (2008) 1301–1321.
[16] E. Löcherbach, D. Loukianova and O. Loukianov, Deviation bounds in ergodic theorem for positively recurrent one-dimensional diffusions and integrability of hitting times. *Ann. Inst. Henri Poincaré* **47** (2011) 425–449.
[17] E. Nummelin, A splitting technique for Harris recurrent Markov chains. *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **43** (1978) 309–318.
[18] S. Pal, Concentration for multidimensional diffusions and their boundary local times. To appear in *Probab. Theory Relat. Fields* (2011), DOI 10.1007/s00440-011-0368-1
[19] D. Revuz, *Markov chains*, Revised edition. Amsterdam, North Holland (1984).
[20] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants*. Springer. *Math. Appl.* **31** (2000).
[21] T. Simon, Support theorem for jump processes. *Stoch. Proc. Appl.* **89** (2000) 1–30.
[22] A.Yu. Veretennikov, On polynomial mixing bounds for stochastic differential equations. *Stoch. Proc. Appl.* **70** (1997) 115–127.
[23] A.Yu. Veretennikov and S.A. Klokov, On subexponential mixing rate for Markov processes. *Teor. Veroyatnost. i Primenen* **49** (2004) 21–35.
[24] L. Wu, A deviation inequality for non-reversible Markov process, *Ann. Inst. Henri Poincaré* **36** (2000) 435–445.