Probab. Theory Relat. Fields 128, 255–321 (2004)
Digital Object Identifier (DOI) 10.1007/s00440-003-0304-0

Liming Wu

Essential spectral radius for Markov semigroups (I): discrete time case

Received: 26 January 2003 / Revised version: 15 September 2003 /
Published online: 12 November 2003 –

Abstract. Using two new measures of non-compactness $\beta_\tau(P)$ and $\beta_w(P)$ for a positive kernel P on a Polish space E, we obtain a new formula of Nussbaum-Gelfand type for the essential spectral radius $r_{ess}(P)$ on $b\mathcal{B}$. Using that formula we show that different known sufficient conditions for geometric ergodicity such as Doeblin's condition, drift condition by means of Lyapunov function, geometric recurrence etc lead to variational formulas of the essential spectral radius. All those can be easily transported on the weighted space $b_u\mathcal{B}$. Some related results on $L^2(\mu)$ are also obtained, especially in the symmetric case. Moreover we prove that for a strongly Feller and topologically transitive Markov kernel, the large deviation principle of Donsker-Varadhan for occupation measures of the associated Markov process holds if and only if the essential spectral radius is zero; this result allows us to show that the sufficient condition of Donsker-Varadhan for the large deviation principle is in fact necessary. The knowledge of $r_{ess}(P)$ allows us to estimate eigenvalues of P in L^2 in the symmetric case, and to estimate the geometric convergence rate by means of that in the metric of Wasserstein. Applications to different concrete models are provided for illustrating those general results.

1. Introduction and questions

Let $P(x, dy)$ be a Markov kernel on some Polish space E. Regarding it as an operator acting on some Banach lattice $\mathbb{B}$ of measurable functions on E via $Pf(x) := \int_E f(y)P(x, dy)$, we are mainly interested in estimating the essential radius

$$r_{ess}(P|_{\mathbb{B}}) := \sup\{|\lambda|;\ \lambda \in \sigma_{ess}(P|_{\mathbb{B}})\}\ \ (\text{convention: } \sup\emptyset := 0)$$

where $\sigma_{ess}(P|_{\mathbb{B}})$ denotes the Wolf essential spectrum of $P|_{\mathbb{B}}$, i.e., the set of $\lambda \in \mathbb{C}$ such that $\lambda - P$ is not a Fredholm operator in $\mathbb{B}$ (a bounded linear operator $A : \mathbb{B} \to \mathbb{B}$ is said *Fredholm* if its range $Ran(A)$ is closed and, its kernel $Ker A$ and the quotient space $\mathbb{B}/Ran A$ are both finite dimensional). Here $\mathbb{B}$, in this paper may be one of

1) $b\mathcal{B}$, the space of all real measurable and bounded functions equipped with sup-norm $\|f\| := \sup_{x\in E} |f(x)|$;

L. Wu: Department of Mathematics, Wuhan University, 430072 Hubei, China
and R. Laboratoire de Mathématiques Appliquées, CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France. e-mail: Li-Ming.Wu@math.univ-bpclermont.fr

Mathematics Subject Classification (2000): 60J05, 60F10, 47A10, 47D07

Key words or phrases: Essential spectral radius – Markov processes – Large deviations

2) $C_b(E)$, the space of all real continuous and bounded functions equipped with the sup-norm; or
3) $L^p(\mu) := L^p_{\mathbb{R}}(E, \mathcal{B}, \mu)$ equipped with the L^p-norm $\|f\|_p$, where μ is an invariant probability measure μ of P.

Much more popular objects in probability than $r_{ess}(P|_{\mathbb{B}})$ are

$$Gap(P|_{\mathbb{B}}) := \inf\{|\lambda - 1|;\ 1 \neq \lambda \in \sigma(P|_{\mathbb{B}})\} \tag{1.1}$$

$$r_{exp}(P|_{\mathbb{B}}) := \sup\{|\lambda|;\ 1 \neq \lambda \in \sigma(P|_{\mathbb{B}})\} \tag{1.2}$$

where $\sigma(P|_{\mathbb{B}})$ is the spectrum of $P|_{\mathbb{B}}$. For an irreducible (in the probabilistic sense defined in [39], [34], [31]) Markov kernel P, $Gap(P|_{\mathbb{B}}) > 0$ if and only if (iff in short) $r_{ess}(P|_{\mathbb{B}}) < 1$, and iff $r_{exp}(P|_{\mathbb{B}}) < 1$ in the case where P is moreover aperiodic (for $\mathbb{B}$ being one of the three spaces above, see Section 2 for explanation). Recall that for an irreducible Markov kernel P,

$$\begin{aligned} r_{exp}(P|_{\mathbb{B}}) &= \lim_{n\to\infty} \left(\sup_{\|f\|_{\mathbb{B}} \le 1} \|P^n f - \mu(f)\|_{\mathbb{B}} \right)^{1/n} \\ &= \inf\{r > 0;\ \exists C > 0, \forall n \ge 1 :\ \|P^n - \mu\|_{\mathbb{B}\to\mathbb{B}} \le Cr^n\} \end{aligned} \tag{1.3}$$

(where $\mu(f) := \int f d\mu$) is the geometric convergence rate of P^n to the invariant probability measure μ in the geometric ergodicity.

When does $r_{ess}(P|_{\mathbb{B}}) < 1$ (or equivalently the spectral gap exists for $P|_{\mathbb{B}}$)? It is a basic question both in Analysis and Probability, whose answer depends sensibly on the choice of $\mathbb{B}$. For example, the finite dimensional Ornstein-Uhlenbeck semigroup $(P_t)_{t>0}$ on $\mathbb{R}^d$ (generated by $\Delta - x \cdot \nabla$) verifies $r_{ess}(P_t|_{L^2(\mu)}) = 0$ but $r_{ess}(P_t|_{b\mathcal{B}}) = 1$ for each $t > 0$ (see Example 3.4 below).

Quite surprisingly in probability the first study was realized on the "bad" space $b\mathcal{B}$, due to the pioneering work of Doeblin [14], [15] (see Doob [17] or Revuz [39] for still updated treatment):

Theorem 1.1 (Doeblin). *Assume that P is an irreducible and aperiodic Markov kernel. If there exist some $N \ge 1$, some probability ν on E and some $\eta > 0$ such that*

$$\sup_{A \in \mathcal{B}:\ \nu(A) < \eta} P^N(x, A) < 1. \tag{1.4}$$

Then $r_{ess}(P|_{b\mathcal{B}}) < 1$ or equivalently $r_{exp}(P|_{b\mathcal{B}}) < 1$.

In particular condition (1.4) is fulfilled if

$$P^N(x, dy) \ge c\nu(dy) \tag{1.5}$$

for some $N \ge 1$, $c > 0$ and probability measure ν.

Under (1.5) (which implies also the irreducibility and the aperiodicity), it is even known that

$$r_{exp}(P|_{b\mathcal{B}}) = \sup\{|\lambda|; \lambda \in \sigma(P|_{b\mathcal{B}})\backslash\{1\}\} \le (1-c)^{1/N}$$

see Meyn and Tweedie [31], Theorem 16.0.2, (16.11) and for the exact references. In that theorem, condition (1.5) is shown even to be necessary to $r_{ess}(P|_{b\mathcal{B}}) < 1$ for an irreducible and aperiodic Markov kernel.

In other words (1.4), (1.5), $r_{exp}(P|_{b\mathcal{B}}) < 1$ and $r_{ess}(P|_{b\mathcal{B}}) < 1$ are all equivalent (called often "*uniform geometric ergodicity*) for an irreducible and aperiodic Markov kernel P. If P is irreducible but not necessarily aperiodic, $r_{ess}(P|_{b\mathcal{B}}) < 1$ is equivalent to the Doeblin recurrence. And for an arbitrary Markov kernel, $r_{ess}(P|_{b\mathcal{B}}) < 1$ is equivalent to the quasi-compactness of $P|_{b\mathcal{B}}$, developed in Revuz [39].

That raises a very natural

Question 1. *how to estimate $r_{ess}(P|_{b\mathcal{B}})$ by means of the data given by Doeblin's condition (1.4)?*

A very useful criterion for the Doeblin recurrence is by means of the so called Lyapunov function:

Theorem 1.2. *Assume that the Markov kernel P is Feller, topologically transitive (see (A3) in Section 4) and P^N is strongly Feller (i.e., $P^N(b\mathcal{B}) \subset C_b(E)$) for some $N \geq 1$. If there is some $1 \leq u \in b\mathcal{B}$ such that*

$$\frac{Pu}{u} \leq r1_{K^c} + b1_K \tag{1.6}$$

for some compact $K \subset E$, some $0 \leq r < 1$ and some $b \geq 0$, then $r_{ess}(P|_{b\mathcal{B}}) < 1$. Moreover condition (1.6) is equivalent to

$$\sup_{x \in E} \mathbb{E}^x \left(\frac{1}{r}\right)^{\sigma_K} < +\infty \tag{1.7}$$

for some $0 < r < 1$ and compact $K \subset E$, where $\sigma_K := \inf\{n \geq 0; X_n \in K\}$ is the first hitting time to K of the Markov process $(\Omega, (\mathcal{F}_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbb{P}_x)_{x\in E})$ associated with transition probability kernel P.

That is contained in [31], Chap.15 and 16 (much more general results are known under the only assumption of irreducibility). Indeed by the assumption of the theorem above, P is irreducible (see Lemma 6.2 in this paper) and every compact subset K is "P-petite" in the language therein by [31], Proposition 6.2.8.

Both (1.6) and (1.7) are necessary to the Doeblin recurrence. Condition (1.6) is also called "drift condition" in [31]. Most known Doeblin recurrent Markov processes are verified by means of (1.6). That leads to

Question 2. *how to estimate $r_{ess}(P|_{b\mathcal{B}})$ by means of the data given by (1.6) or (1.7)?*

In most concrete situation, the Lyapunov function u in (1.6) fails to be bounded. That leads to a modified theory (developed in Nummelin [34] and Meyn and Tweedie [31]). The main instrument is to introduce a new Banach space

$$b_u\mathcal{B} := \left\{ f : E \to \mathbb{R};\ \|f\|_u := \sup_{x\in E} \frac{|f(x)|}{u(x)} < +\infty \right\}.$$

It is known that

Theorem 1.3. *Let P be a Markov kernel satisfying the same assumption as in Theorem 1.2, and aperiodic. If the drift condition (1.6) is verified by some measurable function $u : E \to [1, +\infty)$, then P has an invariant probability measure μ such that $\mu(u) < +\infty$ and there exist some $C > 0, 0 < r < 1$ such that*

$$\|P^n f - \mu(f)\|_u \le Cr^n \|f\|_u, \ \forall n \ge 0, \ f \in b_u\mathcal{B}. \tag{1.8}$$

See [31], Chap. 15 and 16 (and the references therein) for a complete theory about this type of geometric ergodicity, and especially for abundant examples. The key tool for proving it is the Kendall renewal theorem ([24], see [31] Theorem 15.1.1).

In the same way we may ask

Question 3. *How to estimate $r_{ess}(P|_{b_u\mathcal{B}})$ by means of the data given by the drift condition (1.6) for unbounded u ?*

If $r_{ess}(P|_{b\mathcal{B}}) < 1$, then $r_{ess}(P|_{L^\infty(\mu)}) < 1$ where μ is an invariant probability measure of P. By Riesz-Thorin's theorem, $r_{ess}(P|_{L^p(\mu)}) < 1$ for every $1 < p < +\infty$. But in the context of Theorem 1.3, it is unknown whether $r_{ess}(P|_{L^p(\mu)}) < 1$, where $1 < p < +\infty$, which is of great importance too. By Riesz-Thorin's theorem, $r_{ess}(P|_{L^p(\mu)}) < 1$ for each $1 < p < \infty$ is equivalent to that for $p = 2$. And Chen [3] obtains this spectral gap in $L^2(\mu)$ in the context of Theorem 1.3 in several important situations covering the symmetric case. That leads to the following natural

Question 4. *how to estimate $r_{ess}(P|_{L^2(\mu)})$ by means of the data given by the drift condition (1.6) with unbounded u or geometric recurrence?*

Our main purpose is to show that probabilistic tools such as Doeblin's condition (1.4), drift condition (1.6) and geometric recurrence condition (1.7) yield very useful information or even characterization on $r_{ess}(P)$; and inversely knowledge on $r_{ess}(P)$ is very helpful for estimating $r_{exp}(P)$ or even eigenvalues of P.

This paper is organized as follows. In the next section we recall several known facts about $\sigma_{ess}(P)$, its analytic meaning and the famous Nussbaum formula of $r_{ess}(P)$ (for convenience of probabilist reader). Section 3, the central one, is devoted to the study of Questions 1, 2 and 3 both on $b\mathcal{B}$ and $b_u\mathcal{B}$, and contains the main new results of this paper. Some complementary results are presented in Section 4 without the topological hypothesis (A1) in Section 3. Especially we obtain the equivalence between $r_{ess}(P|_{b\mathcal{B}}) < 1$ and $\beta_\tau(P^N) < 1$ for some N for a Harris recurrent Markov kernel P, basing on a deep result of Horowitz [20], and explain the asymptotic behavior of a Markov kernel P on $b_u\mathcal{B}$ if its essential spectral radius is < 1.

The counterparts in L^p of some results in Section 3 are presented in Section 5, where Question 4 is solved (only) in the symmetric case. Sections 6, 7 and 8 are devoted to applications of estimates of $r_{ess}(P)$ obtained previously.

In Section 6, for a strong Feller and topologically transitive Markov chain, we prove the equivalence between $r_{ess}(P) = 0$ and the large deviation principle (LDP in short) of Donsker and Varadhan, and especially the classical sufficient condition of Lyapunov function type found by Donsker and Varadhan for LDP is shown to be necessary. In Section 7 we show at first that the eigenvalues λ of a symmetric Markov kernel P on discrete space E with $|\lambda| > r_{ess}(P)$ can be estimated by the degrees of geometric recurrence to finite subsets of E, and next establish that the geometric convergence in the metric of Wasserstein implies that in $b_u\mathcal{B}$, or in $L^2(\mu)$ in the symmetric case (under some topological assumption), by following the important works of Chen [1], [2] concerning sharp estimates of spectral gap.

As applications of general results, a sequence of widely used concrete models are studied in Section 8: 1) forward recurrence time model, 2) reflected random walk, 3) random perturbed linear systems, 4) auto-regressive model, 5) non-linear random dynamical systems on $\mathbb{R}^d$ etc (a critical case is treated also). For them explicit estimates of the essential spectral radius and even the geometric convergence rate $r_{exp}(P)$ are obtained.

To keep the continuity of presentation the proofs of several results in Sections 3, 4 and 6 are left to Sections 9 and 10. Throughout this paper, we speak "decreasing", "increasing" instead of "non-increasing" and "non-decreasing".

2. Analytic preparation

In this section we recall several useful known facts from the spectral theory of (nonnegative) operators (see [22], [30], [40] and [33]) for probabilist reader's convenience.

Let $\mathbb{B}$ be a real Banach lattice (e.g. $b\mathcal{B}$, $b_u\mathcal{B}$, $C_b(E)$, $L^p(\mu)$ ($p \in [1, +\infty]$) or one of their dual Banach spaces) and P a nonnegative, linear and bounded operator on $\mathbb{B}$. A complex number $\lambda \in \mathbb{C}$ does not belong to the (Wolf) essential spectrum $\sigma_{ess}(P|_{\mathbb{B}})$ of $P|_{\mathbb{B}}$, iff $\lambda - P$ is a Fredholm operator on the complexified Banach space $\mathbb{B}_{\mathbb{C}}$ of $\mathbb{B}$, by definition. For a point λ_0 in the spectrum $\sigma(P|_{\mathbb{B}})$, $\lambda_0 \notin \sigma_{ess}(P|_{\mathbb{B}})$ iff λ_0 is isolated in $\sigma(P|_{\mathbb{B}})$ and the associated eigen-projection

$$E_{\lambda_0} := \frac{1}{2\pi i} \int_\Gamma (\lambda - P)^{-1} d\lambda$$

(Dunford integral in the counter-clockwise way) is finite dimensional, where Γ is a circumference of sufficiently small radius: $|\lambda - \lambda_0| = \delta$ such that the disk $|\lambda - \lambda_0| \le \delta$ contains no other spectral point than λ_0. Recall that the dimension of the range $Range(E_{\lambda_0})$ of E_{λ_0} is the *algebraic multiplicity of* λ_0, denoted by m_a. Hence by Jordan's decomposition of finite matrix,

$$\begin{aligned} &P \cdot E_{\lambda_0} = \lambda_0 E_{\lambda_0} + N_{\lambda_0} \\ &(N_{\lambda_0})^m = 0 \ \text{ for some } \ 1 \le m \le m_a. \end{aligned} \tag{2.1}$$

An explicit expression of E_{λ_0} is given as follows: let $(f_k)_{1\le k\le m_a}$ be a basis of $Range(E_{\lambda_0})$, and $(\phi_k)_{1\le k\le m_a}$ a basis of the range of the adjoint eigen-projection $E^*_{\lambda_0}$ acting on the dual Banach space $(\mathbb{B}_{\mathbb{C}})^*$ such that $\langle \phi_k, f_l\rangle = \phi_k(f_l) = \delta_{kl}$ (Kronecker's notation) (such a basis $(\phi_k)_{1\le k\le m_a}$ is called a dual basis). Then

$$E_{\lambda_0} f = \sum_{k=1}^{m_a} \langle \phi_k, f\rangle f_k, \ \forall f \in \mathbb{B}_{\mathbb{C}}. \tag{2.2}$$

The essential spectral radius $r_{ess}(P|_{\mathbb{B}}) := \sup\{|\lambda|;\ \lambda \in \sigma_{ess}(P|_{\mathbb{B}})\}$ of $P|_{\mathbb{B}}$ can be characterized as the smallest $r \in \mathbb{R}^+$ such that any $\lambda \in \sigma(P|_{\mathbb{B}})$ with $|\lambda| > r$ is an isolated eigenvalue of P with finite algebraic multiplicity. In particular *for any $r > r_{ess}(P|_{\mathbb{B}})$ such that $\Gamma_r := \{\lambda \in \mathbb{C}|\ |\lambda| = r\}$ contain no spectral point*, the spectral points $\lambda \in \sigma(P|_{\mathbb{B}})$ with $|\lambda| > r$ constitute a finite set $\{\lambda_j\}_{0\le j\le N}$. Let Π_r be the sum of all eigenprojections E_{λ_j} associated with those λ_j, $j = 0, 1, \cdots, N$. Then

$$\limsup_{n\to\infty} \frac{\|P^n(I-\Pi_r)\|_{\mathbb{B}\to\mathbb{B}}}{r^n} = 0 \ \text{ and}$$
$$\limsup_{n\to\infty} \frac{\|P^n\Pi_r f\|_{\mathbb{B}}}{r^n} = +\infty \ \text{ if } \ \Pi_r f \neq 0. \tag{2.3}$$

This furnishes a very clear analytical meaning to $r_{ess}(P)$.

An equivalent characterization of $\sigma_{ess}(P|_{\mathbb{B}})$ is obtained through the Calkin algebra $(\mathcal{L}(\mathbb{B}_{\mathbb{C}})/\mathcal{K}(\mathbb{B}_{\mathbb{C}}), \|\cdot\|_{Calkin})$, where $\mathcal{L}(\mathbb{B}_{\mathbb{C}})$ (resp. $\mathcal{K}(\mathbb{B}_{\mathbb{C}})$) is the Banach space of all linear and bounded (resp. and compact) operators on $\mathbb{B}_{\mathbb{C}}$, and $\|P\|_{Calkin} := \inf_{C\in\mathcal{K}(\mathbb{B}_{\mathbb{C}})} \|P - C\|$. Indeed $\sigma_{ess}(P|_{\mathbb{B}})$ coincides with the spectrum of the canonical image of P in the Calkin algebra ([33], A-III, p73-74). That yields immediately three useful (and well known) consequences:

Lemma 2.1. *(i) If C is a compact operator on $\mathbb{B}$, then $\sigma_{ess}((P+C)|_{\mathbb{B}}) = \sigma_{ess}(P|_{\mathbb{B}})$ (Weyl's lemma).*

(ii) The essential spectral radius of $P|_{\mathbb{B}}$ is given by (Gelfand's formula)

$$r_{ess}(P|_{\mathbb{B}}) = \lim_{n\to\infty} \left(\|P^n\|_{Calkin}\right)^{1/n} = \inf_{n\ge 1}\left(\inf_{C\in\mathcal{K}(\mathbb{B})} \|P^n - C\|\right)^{1/n}. \tag{2.4}$$

(iii) (essential spectral mapping theorem) $\sigma_{ess}(f(P)|_{\mathbb{B}}) = \{f(\lambda);\ \lambda \in \sigma_{ess}(P|_{\mathbb{B}})\}$ for any holomorphic function $f : D \to \mathbb{C}$ where the open domain D contains $\sigma(P|_{\mathbb{B}})$.

Nussbaum [37](1970) found a very useful improvement over formula (2.4). The crucial point is the introduction of the following notion:

Definition 2.1 **([37]).** *Let $P : \mathbb{B} \to \mathbb{B}$ be a bounded operator on the Banach space $\mathbb{B}$. The measure of non-compactness of P is defined as*

$$\beta(P) := \inf\{r > 0;\ \exists x_1, \cdots, x_n \in \mathbb{B},\ P(B(1)) \subset \{x_i,\ 1 \le i \le n\} + r\cdot B(1)\} \tag{2.5}$$

where $B(1)$ is the unit ball in $\mathbb{B}$ centered at 0 *with radius* 1.

It is known that ([37] or [30], p274)

$$\beta(P) \le \|P\|_{\mathbb{B}};\quad \beta(PQ) \le \beta(P)\beta(Q). \tag{2.6}$$

Now the Nussbaum formula for essential spectral radius is read as (see also [30], Theorem 4.3.13)

Theorem 2.2 (Nussbaum [37](1970)). *For a bounded linear operator $A : \mathbb{B} \to \mathbb{B}$,*

$$r_{ess}(A) = \lim_{n\to+\infty} \left[\beta(A^n)\right]^{1/n} = \inf_{n\geq 1} \left[\beta(A^n)\right]^{1/n}. \tag{2.7}$$

It is the starting point of this work.

We turn now to some special features of nonnegative operators: a linear bounded operator is said *nonnegative* on $\mathbb{B}$, if $Pf \geq 0$ for any $0 \leq f \in \mathbb{B}$.

Proposition 2.3. *Let P be a nonnegative linear and bounded operator on the Banach lattice $\mathbb{B}$.*

(a) The spectral radius $r_{sp}(P|_{\mathbb{B}})$ is in the spectrum $\sigma(P|_{\mathbb{B}})$ ([40], Chap.V, Proposition 4.1).

(b) If $r_{sp}(P|_{\mathbb{B}}) > 0$ and $r_{sp}(P|_{\mathbb{B}})$ is a pole of the resolvent $R(\lambda, P) := (\lambda - P)^{-1}$, then it corresponds to a positive eigenvector $f \in \mathbb{B}$ ([30], Theorem 4.1.4 and its note), and the peripherical spectrum $\{\lambda \in \sigma(P|_{\mathbb{B}});\ |\lambda| = r_{sp}(P|_{\mathbb{B}})\}$ consists entirely of poles of the resolvent $R(\lambda, P) := (\lambda - P)^{-1}$, and it is a finite union of finite groups of roots of unity. ([40], Chap.V, Theorem 5.5). Moreover the pole $r_{sp}(P|_{\mathbb{B}}) > 0$ is of maximal order among the peripherical spectrum ([30], Proposition 4.1.3).

(c) If $r_{sp}(P|_{\mathbb{B}}) = 1$ and it does not belong to $\sigma_{ess}(P|_{\mathbb{B}})$, then the peripherical spectrum is contained in $\sigma(P|_{\mathbb{B}})\backslash\sigma_{ess}(P|_{\mathbb{B}})$. In particular $r_{ess}(P|_{\mathbb{B}}) < 1$.

Here (c) is an easy consequence of (b). Indeed by part (b) about the cyclic property of the peripherical spectrum of P, for some $N \geq 1$, the peripherical spectrum of P^N is reduced to the singleton $\{1\}$ and 1 does not belong to $\sigma_{ess}(P^N|_{\mathbb{B}})$ by Lemma 2.1(iii). Hence the eigenprojection of P^N associated with 1 is finite dimensional, but it is also the sum of all eigenprojections E_{λ_j} of P associated with the peripherical spectral points λ_j. In other words the peripherical spectral points of P are all of finite algebraic multiplicity, the desired conclusion.

Given now a Markov kernel P and let $\mathbb{B}$ be one of $b\mathcal{B}$, $C_b(E)$, $L^p(\mu)$ ($p \in [1, +\infty]$ and $\mu P = \mu$) such that $Gap(P|_{\mathbb{B}}) = \sup\{|\lambda - 1|;\ 1 \neq \lambda \in \sigma(P|_{\mathbb{B}})\} > 0$. As $\{(\lambda-1)(\lambda-P)^{-1};\ \lambda > 1\}$ are still Markov, then uniformly bounded on $\mathbb{B}$. Developing $(\lambda - P)^{-1}$ in terms of series of Laurent, we see that 1 is a pole of $(\lambda - P)^{-1}$ of order 1. Thus $(1 - P)E_1 = 0$ (by [50], Chap.9, §8), i.e., the geometric multiplicity $dim(Ker(1 - P))$ of 1 coincides with the algebraic multiplicity $dim(Range(E_1))$ (consequently the Markov property is a sufficient condition for the diagonalization of the peripherical spectrum). When $dim(Ker(1 - P))$ is finite (equal to 1 if P is irreducible), we have $r_{ess}(P) < 1$: a remark in Introduction.

The above argument does not work on $\mathbb{B} = b_u\mathcal{B}$, because $\{(\lambda-1)(\lambda-P)^{-1};\ \lambda > 1\}$ are not necessarily uniformly bounded on $b_u\mathcal{B}$.

3. Several formulas of essential spectral radius on $b\mathcal{B}$ and on $b_u\mathcal{B}$

We now go to the job. This section, the central one, is devoted to the study of Questions 1,2 and 3. The key tool is two new parameters β_w, β_τ for measure of non-compactness of P. A great difference from the framework in the Introduction is here we do not impose neither the Markov property, nor the irreducibility.

3.1. Two new parameters

We introduce several necessary notations. Our states space E is Polish with a compatible metric d (i.e., (E, d) is complete and separable), whose Borel σ-field is denoted by $\mathcal{B}$. The notation "$K \subset\subset E$" means that K is compact in E. Let $M_b(E)$ (resp. $M_b^+(E)$, $M_1(E)$) be the space of all σ-additive (resp. and nonnegative; probability) measures of bounded variation on $(E, \mathcal{B})$. The pair relation between $\nu \in M_b(E)$ and $f \in b\mathcal{B}$ is

$$\langle \nu, f \rangle := \nu(f) := \int_E f(x) d\nu(x).$$

Using the pair above, $M_b(E)$ is a subspace of the dual Banach space $(b\mathcal{B})^*$. For a nonnegative kernel $P(x, dy)$, bounded on $b\mathcal{B}$, its adjoint operator P^* on $(b\mathcal{B})^*$ keeps $M_b(E)$ stable, i.e., for each $\nu \in M_b(E)$,

$$P^*\nu(\cdot) = (\nu P)(\cdot) := \int_E \nu(dx) P(x, \cdot) \in M_b(E).$$

Besides the variation norm $\|\nu\|_{var}$-topology, we shall also consider the following two weak topologies on $M_b(E)$. The weak topology $\sigma(M_b(E), b\mathcal{B})$ (i.e., the weakest topology on $M_b(E)$ for which $\nu \mapsto \nu(f)$ is continuous for all $f \in b\mathcal{B}$), according to the usual language ([13]), will be called *τ−topology, denoted simply by τ*. And the weak topology $\sigma(M_b(E), C_b(E))$ (the usual *weak convergence* topology) will be denoted by "*w*". The following result is well known for "*w*", perhaps less for τ:

Lemma 3.1. *Let $\mathcal{M} \subset M_b^+(E)$ be bounded (i.e., $\sup_{\nu \in \mathcal{M}} \|\nu\|_{var} = \sup_{\nu \in \mathcal{M}} \nu(E) < +\infty$).*

(a) These properties are equivalent:
  (a.i) $\mathcal{M}$ is relatively compact for the weak convergence topology;
  (a.ii) for any sequence $(f_n)_{n \ge 0}$ in $C_b(E)$ decreasing pointwise to zero over E,

$$\lim_{n\to\infty} \sup_{\nu \in \mathcal{M}} \nu(f_n) = 0;$$

  (a.iii) $\inf_{K \subset\subset E} \sup_{\nu \in \mathcal{M}} \nu(K^c) = 0$ (Prokohov's criterion).

(b) $\mathcal{M}$ is relatively compact for the τ-topology if and only if for any sequence $(A_n) \subset \mathcal{B}$ decreasing to empty set $\emptyset$, $\lim_{n\to\infty} \sup_{\nu \in \mathcal{M}} \nu(A_n) = 0$.

Proof of Part (b). **Necessity:** The functionals $F_n(\nu) := \nu(A_n)$ on the τ−closure $\bar{\mathcal{M}}^\tau$ of $\mathcal{M}$ is τ-continuous, decreasing to 0 (as n goes to infinity) for each $\nu \in \bar{\mathcal{M}}^\tau$. Since $\bar{\mathcal{M}}^\tau$ is τ-compact, that convergence is uniform over $\bar{\mathcal{M}}^\tau$ by Dini's monotone convergence theorem.

Sufficiency: Let $R := \sup_{\nu \in \mathcal{M}} \|\nu\|_{var}$, which is finite by assumption. Let $\bar{B}'(0, R)$ be the closed ball of radius R (centered at 0) in the topological dual space $(b\mathcal{B})'$. It is compact w.r.t. the weak*-topology $\sigma((b\mathcal{B})', b\mathcal{B})$, which, restricted to

$M_b(E) \subset (b\mathcal{B})'$, coincides with the τ-topology. Consequently to show the τ-relative compactness of $\mathcal{M}$, it is enough to show that each element ϕ belonging to the closure of $\mathcal{M}$ in $((b\mathcal{B})', \sigma((b\mathcal{B})', b\mathcal{B}))$ is a measure.

Indeed, a such continuous linear form ϕ on $b\mathcal{B}$ is finite additive, nonnegative. And by our sufficiency assumption, $\phi(1_{A_n}) \to 0$ for each sequence (A_n) decreasing to empty. Thus ϕ is well a σ-additive measure of bounded variation. □

That motivates us to introduce

Definition 3.1. *(a) For a bounded sub-family $\mathcal{M}$ of $M_b^+(E)$, define*

$$\begin{aligned}\beta_w(\mathcal{M}) &:= \inf_{K\subset\subset E}\ \sup_{\nu\in\mathcal{M}} \nu(K^c)\\ \beta_\tau(\mathcal{M}) &:= \sup_{(A_n)} \lim_{n\to\infty} \sup_{\nu\in\mathcal{M}} \nu(A_n)\end{aligned} \tag{3.1}$$

where $\sup_{(A_n)}$ is taken over all sequences $(A_n) \subset \mathcal{B}$ decreasing to $\emptyset$.

(b) Let $P(x, dy)$ be a nonnegative kernel on E such that $sup_{x\in E} P(x, E) = \|P1\| < +\infty$ (i.e., the boundedness of kernel P). We call

$$\beta_w(P) := \beta_w(\mathcal{M});\quad \beta_\tau(P) := \beta_\tau(\mathcal{M}) \tag{3.2}$$

where $\mathcal{M} = \{P(x,\cdot);\ x \in E\}$, measure of non-$\tau$-compactness and measure of non-"w"-compactness of P, respectively.

The following proposition, whose proof is postponed to §9, summarizes several elementary properties of β_w and β_τ: the last part (g) is the less obvious but crucial for our basic result, Theorem 3.5 below.

Proposition 3.2. *Let P and Q be two nonnegative bounded kernels on E. Then*

(a) $\beta_w(P) \le 2\beta(P|_{b\mathcal{B}})$; and

$$\beta_w(P) \ge \sup_{(f_n)} \lim_{n\to\infty} \|Pf_n\| \tag{3.3}$$

where $\sup_{(f_n)}$ is taken over all sequences $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ decreasing pointwise to zero over E with $\|f_0\| \le 1$. The equality in (3.3) is true if moreover E is locally compact.

(b) $\beta_\tau(P) = \sup_{(f_n)} \lim_{n\to\infty} \|Pf_n\|$ where $\sup_{(f_n)}$ is taken over all sequences $(f_n)_{n\in\mathbb{N}} \subset b\mathcal{B}$ tending pointwise to zero over E with $\sup_n \|f_n\| \le 1$.

(c)

$$\beta_\tau(P) \le 2\beta(P|_{b\mathcal{B}}).$$

(d) Let $\tilde\beta = \beta_w$ or β_τ. For any $a, b \ge 0$,

$$\tilde\beta(aP + bQ) \le a\tilde\beta(P) + b\tilde\beta(Q).$$

(e) It is always true that

$$\beta_\tau(QP) \le \beta_\tau(Q)\beta_\tau(P).$$

If $\beta_w(1_K P) = 0,\ \forall K \subset\subset E$, then

$$\beta_w(QP) \le \beta_w(Q)\beta_w(P).$$

(f) If $\beta_\tau(1_K P) = 0$ for all $K \subset\subset E$ (it is true if P is strongly Feller), then

$$\beta_\tau(QP) \le \beta_w(Q)\beta_\tau(P);$$

(g) If $\beta_\tau(1_K P) = 0 = \beta_\tau(Q)$ for all $K \subset\subset E$, then QP is compact in $b\mathcal{B}$.

By part (a), we give at first a contrary result.

Corollary 3.3. *Let P be a Markov kernel. Assume that (E, d) is locally compact and unbounded. If for each $N \ge 1$ and $K \subset\subset E$, there is a sequence $(x_k) \subset E$ such that* $\lim_{k\to\infty} P^N(x_k, K) = 0$, *then $r_{ess}(P|_{b\mathcal{B}}) = 1$. In particular $r_{ess}(P|_{b\mathcal{B}}) = 1$ if*

$$P(x, B(x, r)^c) \le h(r), \ \forall x \in E, r \ge 0 \tag{3.4}$$

where $h(r) : \mathbb{R}^+ \to [0, 1]$ decreases to zero as r goes to infinity with $h(0) = 1$, and $B(x, r) := \{y \in E; d(x, y) < r\}$.

Condition (3.4) means that P is of asymptotically bounded range.

Proof. For the first claim, notice that for each $N \ge 1$ and $K \subset\subset E$, $\sup_{x\in E} P^N(x, K^c) = 1 - \inf_{x\in E} P^N(x, K) = 1$. Then $\beta_w(P^N) = 1$. Thus by Proposition 3.2(a),

$$2\beta(P^N|_{b\mathcal{B}}) \ge \beta_w(P^N) = 1.$$

Consequently the Nussbaum formula (2.7) implies that $r_{ess}(P|_{b\mathcal{B}}) \ge 1$.

For the second claim, we show at first that under (3.4),

$$P^n(x, B(x, nr)) \ge (1 - h(r))^n, \ \forall n \ge 1, x \in E, r > 0.$$

Indeed it is true for $n = 1$ by (3.4). Assume it for n. Then for $n + 1$, noting that $d(x, y) < r$ implies $B(y, nr) \subset B(x, (n + 1)r)$, we have

$$\begin{aligned} P^{n+1}(x, B(x, (n+1)r)) &\ge \int_{y\in B(x,r)} P(x, dy)P^n(y, B(x, (n + 1)r)) \\ &\ge \int_{y\in B(x,r)} P(x, dy)P^n(y, B(y, nr)) \\ &\ge P(x, B(x, r))(1 - h(r))^n \ \text{ (recurrence assumption)} \\ &\ge (1 - h(r))^{n+1}, \end{aligned}$$

the desired result.

Let now $K \subset\subset E$ be arbitrary but fixed. For any $N \ge 1, r > 0$, pick points $x_k \in E$ $(k \ge 1)$ such that $d(x_k, K) > Nk$. Thus $B(x_k, Nk) \subset K^c$. We get so $P^N(x_k, K) \le 1 - P^N(x_k, B(x_k, Nk)) \le 1 - (1 - h(k))^N$. Letting k tend to infinity, we see that the condition in the first claim is verified. □

Example 3.4. Let $P_t f(x) := \int_{\mathbb{R}} f\left(e^{-t}x + \sqrt{1 - e^{-2t}}y\right)\mu(dy)$ be the Ornstein-Uhlenbeck operators on $E = \mathbb{R}$, where μ is the standard Gaussian law $\mathcal{N}(0, 1)$. For each $t > 0$, it is well known that $r_{ess}(P_t|_{L^2(\mu)}) = 0$. But $r_{ess}(P_t|_{b\mathcal{B}}) = 1$ by the Corollary above for $\lim_{|x|\to\infty} P_t(x, K) = 0$ for each compact $K \subset\subset \mathbb{R}$.

3.2. Gelfand-Nussbaum type formula on $b\mathcal{B}$

A meta-physical feeling is that the essential spectrum of P comes from "what happens very far away". It may be true only if one imposes a condition which excludes the influence of "what happens locally" over the essential spectrum. For this purpose, we shall choose the following assumption

$$\beta_w(1_K P) = 0 \text{ and } \exists N \geq 1 : \ \beta_\tau(1_K P^N) = 0, \ \forall K \subset\subset E. \qquad \textbf{(A1)}$$

That is justified by the following Gelfand-Nussbaum type formula for the essential spectral radius.

Theorem 3.5. *Let P be a bounded nonnegative kernel on E.*

(a) Assume (A1). Then

$$\begin{aligned} r_{ess}(P|_{b\mathcal{B}}) &= \lim_{n\to\infty} \left[\beta_w(P^n)\right]^{1/n} = \inf_{n\geq 1} \left[\beta_w(P^n)\right]^{1/n} \\ &= \inf_{K\subset\subset E} r_{sp}\left(1_{K^c} P 1_{K^c}|_{b\mathcal{B}}\right). \end{aligned} \tag{3.5}$$

(b) Without (A1), we have always

$$\begin{aligned} r_{ess}(P|_{b\mathcal{B}}) &\geq \max\{\lim_{n\to\infty} [\beta_\tau(P^n)]^{1/n}, \limsup_{n\to\infty} \left[\beta_w(P^n)\right]^{1/n}\} \\ &\geq \inf_{K\subset\subset E} r_{sp}\left(1_{K^c} P 1_{K^c}|_{b\mathcal{B}}\right). \end{aligned} \tag{3.6}$$

It is basic for the whole paper and its proof will be given in §9.

Remarks (3.i). Without our extra condition (A1), (3.5) is false. For instance for any Markov Feller kernel P on a compact metric space E, we have $\beta_w(P) = 0$ but in general $r_{ess}(P|_{b\mathcal{B}})$ is not zero, as well as $r_{ess}(P|_{C_b(E)})$ which coincides with $r_{ess}(P|_{b\mathcal{B}})$, as shown in §4. A typical example is given by $P(x, dy) := \delta_{x^2}$ on the circle S^1 of the complex plane $\mathbb{C}$, for which $\beta_w(P) = 0$, but $r_{ess}(P|_{b\mathcal{B}}) = 1$ by Theorem 3.5(b) since $\beta_\tau(P^n) = 1$ for all n.

Condition (A1) is satisfied if

(A2) P is Feller and P^N is strongly Feller for some $N \geq 1$,

a very current situation. Here P is said *Feller* if $P(C_b(E)) \subset C_b(E)$, i.e., $x \to P(x, \cdot)$ is continuous from E to $(M_1(E), w)$; *strongly Feller*, if $P^N(b\mathcal{B}) \subset C_b(E)$, i.e., $x \to P(x, \cdot)$ is continuous from E to $(M_1(E), \tau)$.

Remarks (3.ii). The last equality in (3.5) is an extension of Persson's principle, well known in the L^2-theory of Schrödinger operators ([7], Theorem 3.12).

The measure of non-"w"-compactness $\beta_w(P)$ is smaller than $2\beta(P|_{b\mathcal{B}})$. Unlike the parameter $\beta(P|_{b\mathcal{B}})$ of Nussbaum, it can be estimated by means of the Lyapunov function method or other probabilistic methods.

We present now two immediate corollaries. The first is a partial answer to Question 1:

Corollary 3.6. *Assume that the nonnegative bounded kernel P satisfies (A1). Then*

$$r_{ess}(P|_{b\mathcal{B}}) \le \inf_{N\ge 1}\left(\inf_{\nu\in M_1(E)}\inf_{\eta>0}\sup_{A:\ \nu(A)<\eta}\sup_{x\in E} P^N(x,A)\right)^{1/N}.$$

Proof. Notice that for any probability measure ν and $\eta > 0$ and $N \ge 1$,

$$\beta_w(P^N|_{b\mathcal{B}}) \le \limsup_{\nu(A)\to 0}\|P^N 1_A\| \le \sup_{A:\ \nu(A)<\eta}\sup_{x\in E} P^N(x,A),$$

then the desired formula follows immediately from (3.5). □

Corollary 3.7. *Let* $0 \le P \le Q$. *Assume that P satisfies (A1) and the kernel Q is bounded. Then*

$$r_{ess}(P|_{b\mathcal{B}}) \le r_{ess}(Q|_{b\mathcal{B}}).$$

Even under (A1), a such comparison result is false for

$$r_{exp}(P|_{b\mathcal{B}}) := \sup\{|\lambda|;\ \lambda \ne r_{sp}(P|_{b\mathcal{B}}), \lambda\in\sigma(P|_{b\mathcal{B}})\}, \tag{3.7}$$

where $\sigma(P|_{b\mathcal{B}})$ denotes the spectrum, $r_{sp}(P|_{b\mathcal{B}}) := \sup\{|\lambda|; \lambda\in\sigma(P|_{b\mathcal{B}})\}$ is the spectral radius of $P|_{b\mathcal{B}}$.

Proof. By Theorem 3.5, we have

$$r_{ess}(P|_{b\mathcal{B}}) = \lim_{n\to\infty}\left[\beta_w(P^n)\right]^{1/n} \le \limsup_{n\to\infty}\left[\beta_w(Q^n)\right]^{1/n} \le r_{ess}(Q|_{b\mathcal{B}}).$$ □

Remarks (3.iii). A complete answer to Question 1 should be

Conjecture. *For any bounded nonnegative kernel $P(x,dy)$ on a general Polish space E,*

$$r_{ess}(P|_{b\mathcal{B}}) = r_\tau(P) := \lim_{n\to\infty}\left[\beta_\tau(P^n)\right]^{1/n} = \inf_{n\ge 1}\left[\beta_\tau(P^n)\right]^{1/n}. \tag{3.8}$$

In the locally compact case, we have by Proposition 3.2(a) and (b) that $\beta_w(P) \le \beta_\tau(P)$. Thus in that case and under (A1), the formula (3.8) holds true. We shall present some partial results supporting this conjecture in Sections 4 and 6.

We now present two examples to illustrate the power of formula (3.5).

Example 3.8 (the so called basic example in [23]). A sequence of tasks is performed in a certain order indexed by $\mathbb{N}$. At step k, *success*, with probability $p_k \in (0,1)$, means that the process goes to the next step $k+1$; but *failure*, of probability $q_k = 1 - p_k$, signifies that the process must start over at state 0.

This example is modelled as: $E = \mathbb{N}$ and $P(k,k+1) = p_k$, $P(k,0) = q_k$ for all $k \ge 0$. It is obvious that

$$\inf_{N\ge 1}\|\left(1_{[0,N]^c}P1_{[0,N]^c}\right)^n\| = \inf_{N\ge 1}\sup_{k>N}\prod_{j=0}^{n-1} p_{k+j}.$$

Thus by Theorem 3.5, we have for this Markov process,

$$r_{ess}(P|_{l^\infty(\mathbb{N})}) = \inf_{n\ge 1}\limsup_{k\to\infty}\left(p_k p_{k+1}\cdots p_{k+n-1}\right)^{1/n}.$$

Example 3.9 (taken from [31], p400 and [47]). Let $E = \mathbb{N}$. The Markov transition kernel is given by

$$P(0,i) = a_i \ (\forall i \geq 0),\ \ P(j,j) = p_j,\ \ P(j,0) = 1 - p_j =: q_j, \forall j \geq 1.$$

For this example we see that for each $N \geq 1$

$$\| \left(1_{[0,N]^c} P 1_{[0,N]^c}\right)^n \| = \sup_{j>N} p_j^n.$$

Hence by Theorem 3.5,

$$r_{ess}(P|_{l^\infty(\mathbb{N})}) = \limsup_{j\to+\infty} p_j.$$

3.3. *Lyapunov function type criteria for $r_{ess}(P|_{b\mathcal{B}})$*

The following result, being an easy consequence of Theorem 3.5, answers Question 2 in the framework of (A1).

Theorem 3.10. *Let P be a nonnegative bounded kernel on E satisfying (A1). Given $r > 0$, the following properties (a), (b), (c) are equivalent*

(a) $r_{ess}(P|_{b\mathcal{B}}) < r$;
(b) there are $1 \leq u \in b\mathcal{B}$ *and* $K \subset\subset E$ *such that*

$$\sup_{x\notin K} \frac{P(1_{K^c}u)}{u}(x) < r. \tag{3.9}$$

(c) there exist some $1 \leq u \in b\mathcal{B}$ *and* $K \subset\subset E$ *such that*

$$\sup_{x\notin K} \frac{Pu}{u}(x) < r. \tag{3.10}$$

In particular we have

$$\begin{aligned} r_{ess}(P|_{b\mathcal{B}}) &= \inf_{K\subset\subset E}\ \inf_{1\leq u\in b\mathcal{B}}\ \sup_{x\notin K} \frac{P(1_{K^c}u)(x)}{u(x)} \\ &= \inf_{K\subset\subset E}\ \inf_{1\leq u\in b\mathcal{B}}\ \sup_{x\notin K} \frac{Pu(x)}{u(x)}. \end{aligned} \tag{3.11}$$

Proof. $(c) \implies (b)$. Obvious.
$(b) \implies (a)$. Consider the kernel operator

$$P^u f = \frac{1}{u}P(uf),\ \ P^u(x,dy) = \frac{u(y)}{u(x)}P(x,dy).$$

It is similar to P on $b\mathcal{B}$. P^u satisfies again (A1). Thus by Theorem 3.5,

$$r_{ess}(P|_{b\mathcal{B}}) = r_{ess}(P^u|_{b\mathcal{B}}) \leq \|1_{K^c}P^u 1_{K^c}\| < r$$

where the last inequality is exactly our condition (3.9).

$(a) \Rightarrow (c)$(without (A1)). Fix some $\lambda : r_{ess}(P|_{b\mathcal{B}}) < \lambda < r$. By Theorem 3.5(b), there exists some compact $K \subset\subset E$ such that

$$r_{sp}(1_{K^c}P1_{K^c}) = \lim_{n\to\infty} \|(1_{K^c}P1_{K^c})^n\|^{1/n} < \lambda.$$

Hence

$$h_\lambda := 1_{K^c} + \sum_{k=1}^{\infty} \lambda^{-n}(1_{K^c}P1_{K^c})^n 1 \in b\mathcal{B}.$$

Notice that $1_{K^c}Ph_\lambda = (1_{K^c}P1_{K^c})h_\lambda \le \lambda h_\lambda$. Now for any $\varepsilon > 0$ so that $\lambda+\varepsilon < r$, letting $u(x) = 1 + Lh_\lambda$ where $L \ge \|P\|/\varepsilon$, we have

$$\frac{1_{K^c}Pu}{u} \le 1_{K^c}\frac{\|P\| + L1_{K^c}Ph_\lambda}{u} \le 1_{K^c}\frac{\|P\| + L\lambda h_\lambda}{1 + L\cdot h_\lambda} \le \varepsilon + \lambda < r,$$

the desired claim (c).

It remains to show (3.11). By $(b) \implies (a)$, we have

$$r_{ess}(P|_{b\mathcal{B}}) \le \inf_{K\subset\subset E}\ \inf_{1\le u\in b\mathcal{B}}\ \sup_{x\notin K} \frac{P(1_{K^c}u)(x)}{u(x)} \le \inf_{K\subset\subset E}\ \inf_{1\le u\in b\mathcal{B}}\ \sup_{x\notin K} \frac{Pu(x)}{u(x)}.$$

And by $(a) \implies (c)$, the last term at the r.h.s. above is smaller than $r_{ess}(P|_{b\mathcal{B}})$. The proof is completed. □

Remarks (3.iv). Without assumption (A1), then by the proof of $(a) \Rightarrow (c)$ above, we have

$$r_{ess}(P|_{b\mathcal{B}}) \ge \inf_{K\subset\subset E}\ \inf_{1\le u\in b\mathcal{B}}\ \sup_{x\notin K} \frac{Pu(x)}{u(x)}. \tag{3.12}$$

3.4. Essential spectral radius on $b_U\mathcal{B}$: variational formula

Given a nonnegative kernel $P(x, dy)$, let

$$\mathcal{A}(P) := \left\{ U : E \to [1, +\infty) \ \begin{array}{l}\text{is measurable and locally bounded and}\\ \sup_{x\in E} \frac{PU(x)}{U(x)} < +\infty\end{array} \right\}. \tag{3.13}$$

Here U is said "*locally bounded*", if $\sup_{x\in K} |U(x)| < +\infty$ for all $K \subset\subset E$. In other words for a measurable locally bounded function $U \ge 1$ on E, $U \in \mathcal{A}(P)$ iff $P : b_U\mathcal{B} \to b_U\mathcal{B}$ is bounded w.r.t. the norm $\|\cdot\|_U$, where

$$b_U\mathcal{B} := \left\{ f : E \to \mathbb{R} \text{ measurable such that } \|f\|_U := \sup_{x\in E} \frac{|f(x)|}{U(x)} < +\infty \right\}.$$

Theorem 3.11. *Let P be a nonnegative bounded kernel on E satisfying (A1).*

(a) Let $U \in \mathcal{A}(P)$. Assume moreover that $P^U(x, dy) := (U(y)/U(x))P(x, dy)$ satisfies again (A1), or in particular

$$U^p \in \mathcal{A}(P) \ \textit{for some}\ \ p > 1. \tag{3.14}$$

Given $r > 0$, the following properties are equivalent:

(a.i) $r_{ess}(P|_{b_U\mathcal{B}}) < r$;

(a.ii) There exist some measurable function $V \sim U$ (i.e., $c^{-1}U \le V \le cU$ for some $c > 1$) and some compact $K \subset\subset E$ such that

$$\sup_{x\notin K} \frac{P(1_{K^c}V)(x)}{V(x)} < r; \tag{3.15}$$

(a.iii) There exist some measurable function $V \sim U$ and some compact $K \subset\subset E$ such that

$$\sup_{x\notin K} \frac{PV(x)}{V(x)} < r; \tag{3.16}$$

In particular,

$$\begin{aligned} r_{ess}(P|_{b_U\mathcal{B}}) &= \inf_{K\subset\subset E} \inf_{V\sim U} \sup_{x\notin K} \frac{P(1_{K^c}V)(x)}{V(x)} \\ &= \inf_{K\subset\subset E} \inf_{V\sim U} \sup_{x\notin K} \frac{PV(x)}{V(x)}. \end{aligned} \tag{3.17}$$

(b) It holds that

$$\begin{aligned} \inf_{U\in\mathcal{A}(P)} r_{ess}(P|_{b_U\mathcal{B}}) &= \inf_{K\subset\subset E} \inf_{u\in\mathcal{A}(P)} \sup_{x\notin K} \frac{P(1_{K^c}u)(x)}{u(x)} \\ &= \inf_{K\subset\subset E} \inf_{u\in\mathcal{A}(P)} \sup_{x\notin K} \frac{Pu(x)}{u(x)}. \end{aligned} \tag{3.18}$$

The quantity $\inf_{U\in\mathcal{A}(P)} r_{ess}(P|_{b_U\mathcal{B}})$ will be crucial in our investigation of locally uniform large deviation principle in §6.

Proof. **(a)** Notice that

$$P^U f(x) = \frac{1}{U(x)} P(Uf)(x) = (M_U)^{-1} P M_U f, \ \forall f \in b\mathcal{B}$$

where $M_U(f) := Uf$ is an isomorphism from $(b\mathcal{B}, \|\cdot\|)$ to $(b_U\mathcal{B}, \|\cdot\|_U)$. Thus $r_{ess}(P|_{b_U\mathcal{B}}) = r_{ess}(P^U|_{b\mathcal{B}})$.

1) At first let us show that (3.14) implies P^U satisfies (A1). We begin with showing that $\beta_\tau(1_K(P^U)^N) = 0$ for each $K \subset\subset E$, where N is fixed in (A1). Indeed for any sequence $(A_n) \subset \mathcal{B}$ decreasing to $\emptyset$, we have by Hölder's inequality $(1/p + 1/p' = 1)$,

$$\begin{aligned} \sup_{x\in K} (P^U)^N 1_{A_n}(x) &= \sup_{x\in K} \frac{P^N(U1_{A_n})(x)}{U(x)} \\ &\le \sup_{x\in K} \left(P^N(U^p)(x)\right)^{1/p} \sup_{x\in K} P^N(x, A_n)^{1/p'}. \end{aligned}$$

By assumption (3.14), it follows that $\lim_{n\to\infty} \sup_{x\in K} (P^U)^N 1_{A_n}(x) = 0$, the desired claim by Lemma 3.1.

In the same way we have $\beta_w(1_K P^U) = 0$ for any $K \subset\subset E$. In summary P^U satisfies (A1).

2) Since $1 \le u \in b\mathcal{B}$ satisfies (3.9) or (3.10) in Theorem 3.10 with respect to P^U (instead of P) iff $V = uU \sim U$ satisfies (3.15) or (3.16), the equivalence between (a.i), (a.ii), (a.iii) follows from Theorem 3.10 applied to P^U on $b\mathcal{B}$. Those equivalences entail (3.17).

(b) To establish (3.18), the key is to remove the technical condition (3.14). At first by (3.12) in Remarks (3.iv) and the proof above, we always have (without condition (3.14))

$$r_{ess}(P|_{b_U\mathcal{B}}) \ge \inf_{K\subset\subset E} \sup_{x\notin K} \frac{PU(x)}{U(x)}, \ \forall U \in \mathcal{A}(P). \tag{3.19}$$

Now for (3.18), it remains to show that *for any $U \in \mathcal{A}(P)$ and r satisfying*

$$\inf_{K\subset\subset E} \sup_{x\notin K} \frac{P(1_{K^c}U)(x)}{U(x)} < r,$$

and for any $\varepsilon > 0$, there exists some $V \in \mathcal{A}(P)$ such that $r_{ess}(P_{b_V\mathcal{B}}) \le r + \varepsilon$.

Indeed choose some compact K such that $\sup_{x\notin K} \frac{P(1_{K^c}U)(x)}{U(x)} < r$. Put $V = U^a$ where $0 < a < 1$. We have by Hölder's inequality that for any $x \notin K$,

$$P(1_{K^c}V)(x) \le [P(1_{K^c}U)(x)]^a \, \|P1\|^{1-a} \le r^a \|P1\|^{1-a} V(x).$$

The same argument shows that $V \in \mathcal{A}(P)$. Hence for a sufficiently close to 1 (but fixed), we have

$$\sup_{x\notin K} \frac{P(1_{K^c}V)(x)}{V(x)} \le r + \varepsilon$$

where it follows by part (a) (as $V^{1/a} = U \in \mathcal{A}(P)$),

$$r_{ess}(P|_{b_V\mathcal{B}}) \le r + \varepsilon$$

the desired claim. □

Remarks (3.v). Without assumption (A1), then by the proof above and by Remarks (3.iv), we always have

$$r_{ess}(P|_{b_U\mathcal{B}}) \ge \inf_{K\subset\subset E} \inf_{V\sim U} \sup_{x\notin K} \frac{PV(x)}{V(x)}, \ \forall U \in \mathcal{A}(P). \tag{3.20}$$

Remarks (3.vi). The technical condition (3.14) is used to guarantee that P^U satisfies again (A1) once does P. Another useful observation is the following:

if P is Feller and P^N is strongly Feller, and if $U \in \mathcal{A}(P)$ and U, PU are continuous, then P^U is Feller and $(P^U)^N$ is strongly Feller.

Its proof, based on Dini's monotone convergence theorem, is omitted.

3.5. $r_{ess}(P)$ by means of degree of geometric recurrence

In this paragraph we assume that $P(x, dy)$ is a Markov kernel (i.e., $P \geq 0$ and P1=1).

Let $(\Omega := E^{\mathbb{N}}, (X_n(\omega) = \omega_n)_{n\geq 0}, \mathcal{F}_n := \sigma(X_m;\ m \leq n), (\mathbb{P}_x)_{x\in E})$ be the associated Markov chain. For any measurable non-empty subset A of E, define

$$\sigma_A := \inf\{n \geq 0;\ X_n \in A\};\ \tau_A := \inf\{n \geq 1;\ X_n \in A\}$$

where $(\theta\omega)_n = \omega_{n+1}$ is the shift on $\Omega = E^{\mathbb{N}}$. Note that $\tau_A = \sigma_A \circ \theta + 1$ and if $x \notin A$, then $\tau_A = \sigma_A,\ \mathbb{P}_x - a.s.$.

The following lemma exhibits a close relation between the drift condition and the geometric recurrence of (X_n) (it is a variant of some well known results, see [31] and [34]):

Lemma 3.12. *Let P be a Markov kernel and $K \subset\subset E$.*

(a) Given $U \in \mathcal{A}(P)$.

$$\begin{aligned}\inf_{V\sim U} \sup_{x\notin K} \frac{PV(x)}{V(x)} &= \limsup_{n\to\infty} \left(\sup_{x\in E} \frac{1}{U(x)} \mathbb{E}^x 1_{[\sigma_K>n]} U(X_n) \right)^{1/n} \\ &= r_{sp}(1_{K^c} P 1_{K^c}|_{b_U\mathcal{B}}).\end{aligned} \tag{3.21}$$

(b)

$$\inf_{V\in\mathcal{A}(P)} \sup_{x\notin K} \frac{PV(x)}{V(x)} = \sup_{K'\subset\subset E} \limsup_{n\to\infty} \left(\sup_{x\in K'} \mathbb{P}_x(\tau_K > n) \right)^{1/n}. \tag{3.22}$$

A quite delicate point here is that in (3.22), τ_K can not be replaced by σ_K (see the example in §8.1).

The main bridge connecting the recurrence property and $r_{sp}(1_{K^c} P 1_{K^c})$ is the following simple but crucial fact:

$$\frac{1}{U(x)} \mathbb{E}^x 1_{[\sigma_K>n]} U(X_n) = \frac{1}{U(x)} \left[(1_{K^c} P 1_{K^c})^n\, U \right](x). \tag{3.23}$$

Proof. **(a). 1)** We begin with the proof of the equality

$$r_0 := \inf_{V\sim U} \sup_{x\notin K} \frac{PV(x)}{V(x)} = \limsup_{n\to\infty} \left(\sup_{x\in E} \frac{1}{U(x)} \mathbb{E}^x 1_{[\sigma_K>n]} U(X_n) \right)^{1/n} =: r_1.$$

For any $r > r_0$, there is some $V \sim U$ such that

$$\sup_{x\notin K} \frac{PV(x)}{V(x)} < r.$$

Consider the martingale

$$M_0 := 1,\ M_n := \prod_{k=1}^{n} \frac{V(X_k)}{PV(X_{k-1})} = \frac{V(X_n)}{V(X_0)} \cdot \prod_{k=0}^{n-1} \frac{V}{PV}(X_k),\ n \geq 1. \tag{3.24}$$

Noting that $\frac{V}{PV}(X_k) \geq 1/r$ on $[k < \sigma_K]$, we have for $n \geq 1$,

$$\frac{1}{V(x)}\mathbb{E}^x V(X_n)1_{[\sigma_K > n-1]}\left(\frac{1}{r}\right)^n \leq \mathbb{E}^x M_n = 1$$

where it follows by using the equivalence $V \sim U$,

$$r_1 \leq \limsup_{n\to\infty}\left(\sup_{x\in E}\frac{1}{V(x)}\mathbb{E}^x 1_{[\sigma_K > n-1]}V(X_n)\right)^{1/n} \leq r.$$

As $r > r_0$ is arbitrary, we have shown $r_1 \leq r_0$.

We now prove the inverse inequality $r_0 \leq r_1$. Notice at first that

$$r_1 = \limsup_{n\to\infty}\left(\sup_{x\in E}\frac{1}{U(x)}\mathbb{E}^x 1_{[\sigma_K \geq n]}U(X_n)\right)^{1/n}$$

because $\mathbb{E}^x 1_{[\sigma_K \geq n]}U(X_n) = \mathbb{E}^x 1_{[\sigma_K > n-1]}PU(X_{n-1}) \leq C\mathbb{E}^x 1_{[\sigma_K > n-1]}U(X_{n-1})$ where $C = \|PU/U\|$. For any $r > r_1$, put

$$V(x) := \sum_{n=0}^{\infty} r^{-n}\mathbb{E}^x 1_{[\sigma_K \geq n]}U(X_n).$$

By the definition of r_1, $V \sim U$. For initial point $x \in K^c$, noting that $1_{[\sigma_K \geq n]} \circ \theta = 1_{[\sigma_K \geq n+1]}$, $\mathbb{P}_x - a.s.$, we have for all $x \in K^c$,

$$\begin{aligned} PV(x) &= \sum_{n=0}^{\infty} r^{-n}\mathbb{E}^x 1_{[\sigma_K \geq n]} \circ \theta \cdot U(X_{n+1}) \\ &= \sum_{n=0}^{\infty} r^{-n}\mathbb{E}^x 1_{[\sigma_K \geq n+1]}U(X_{n+1}) \leq rV(x). \end{aligned}$$

i.e., $\sup_{x\in K^c}(PV/V)(x) \leq r$. Thus $r_0 \leq r$, the desired inequality.

2) It remains to prove the second equality in (3.21). Indeed,

$$\begin{aligned} &\limsup_{n\to\infty}\left(\sup_{x\in E}\frac{1}{U(x)}\mathbb{E}^x 1_{[\sigma_K > n]}U(X_n)\right)^{1/n} \\ &= \limsup_{n\to\infty}\left(\sup_{x\in E}\frac{1}{U(x)}\left[(1_{K^c}P1_{K^c})^n\, U\right](x)\right)^{1/n} \quad \text{(by (3.23))} \\ &= \limsup_{n\to\infty}\left(\left\|(1_{K^c}P1_{K^c})^n\right\|_{b_U\mathcal{B}}\right)^{1/n} = r_{sp}(1_{K^c}P1_{K^c}|_{b_U\mathcal{B}}). \end{aligned}$$

(b). We show at first that for any $V \in \mathcal{A}(P)$ fixed,

$$\sup_{x\notin K}\frac{PV(x)}{V(x)} \geq \sup_{K'\subset\subset E}\limsup_{n\to\infty}\left(\sup_{x\in K'}\mathbb{P}_x(\tau_K > n)\right)^{1/n}. \tag{3.25}$$

For any r strictly greater than the l.h.s. above, the martingale (M_n) constructed in (3.24) verifies

$$M_n := \prod_{k=1}^{n} \frac{V(X_k)}{PV(X_{k-1})} = \frac{PV(X_n)}{PV(X_0)} \cdot \prod_{k=1}^{n} \frac{V}{PV}(X_k) \geq \frac{1}{PV(X_0)} 1_{[\tau_K > n]} \left(\frac{1}{r}\right)^n .$$

Then

$$\sup_{x \in K'} \mathbb{P}_x(\tau_K > n) \leq \sup_{x \in K'} PV(x) \mathbb{E}^x M_n \cdot r^n \leq \sup_{x \in K'} PV(x) \cdot r^n$$

where the above desired inequality (3.25) follows for $\sup_{x \in K'} PV(x) < +\infty$.

We now establish the inverse inequality of (3.22). For any r strictly greater than the r.h.s. of (3.22), consider

$$V(x) := \mathbb{E}^x \left(\frac{1}{r}\right)^{\sigma_K} .$$

Then $V \geq 1$ over E and V is bounded on any $K' \subset\subset E$. Moreover, using $\sigma_K \circ \theta + 1 = \tau_K$ and $\sigma_K = \tau_K$, $\mathbb{P}_x - a.s.$ for $x \notin K$, we have

$$PV(x) = \mathbb{E}^x \left(\frac{1}{r}\right)^{\sigma_K \circ \theta} = r\mathbb{E}^x \left(\frac{1}{r}\right)^{\tau_K} = r 1_{x \in K^c} V(x) + r 1_{x \in K} \sup_{z \in K} \mathbb{E}^z \left(\frac{1}{r}\right)^{\tau_K} .$$

It follows that $V \in \mathcal{A}(P)$ and $\sup_{x \in K^c} (PV/V)(x) \leq r$. Therefore the inverse inequality of (3.22) is proved. The equality (3.22) is established. □

From the previous lemma and Theorem 3.11(b), we immediately obtain

Corollary 3.13. *Assume that the Markov kernel satisfies (A1). Then*

$$\begin{aligned} &\inf_{U \in \mathcal{A}(P)} r_{ess}(P|_{b_U \mathcal{B}}) \\ &\quad = \inf\{r > 0;\ \exists K \subset\subset E,\ \sup_{x \in K'} \mathbb{E}^x \left(\frac{1}{r}\right)^{\tau_K} < +\infty,\ \forall K' \subset\subset E\} \end{aligned} \tag{3.26}$$

Formula (3.26) will be the basis for applying the comparison technique to estimates of $r_{ess}(P)$, see §8.3 and §8.4.

3.6. $r_{ess}(P)$ and concentration of invariant measure

Proposition 3.14. *Let P be a Markov kernel such that $r_0 := \inf_{U \in \mathcal{A}(P)} r_{ess}(P|_{b_U \mathcal{B}}) < 1$. Assume the following near-neighbor condition*

$$P(x, \{y; d(y,x) > 1\}) = 0,\ \forall x \in E. \tag{3.27}$$

Then for any invariant probability measure μ of P, i.e., $\mu P = \mu$, (whose existence is assured by Theorem 4.6 below) and each fixed $x_0 \in E$,

$$\limsup_{N \to \infty} [\mu(\{x; d(x, x_0) > N\})]^{1/N} \leq r_0. \tag{3.28}$$

Proof. (following that of [31], Theorem 16.3.2) For any $r \in (r_0, 1)$, by Lemma 3.12(3.21) and Remarks (3.v), there are some compact $K \subset\subset E$, $U \in \mathcal{A}(P)$ and $V \sim U$ such that

$$\limsup_{n\to\infty} \left(\sup_{x\in E} \frac{1}{U(x)} \mathbb{E}^x 1_{[\sigma_K > n]} U(X_n) \right)^{1/n} \le \sup_{x\notin K} \frac{PV(x)}{V(x)} < r. \tag{3.29}$$

Fix an invariant probability measure μ of P. Let us show that V is μ-integrable by following a classical argument.

Indeed since $\left(\sum_{k=1}^n (V(X_k) - PV(X_{k-1}))\right)_{n\ge1}$ is a $\mathbb{P}_x$-martingale for each $x \in E$, we have by Doob's stopping time theorem that for any $n \ge 1$,

$$\begin{aligned} 0 &= \mathbb{E}^x \sum_{k=1}^{\tau_K\wedge n} (V(X_k) - PV(X_{k-1})) \\ &= \mathbb{E}^x (PV)(X_{\tau_K\wedge n}) - PV(x) + \mathbb{E}^x \sum_{k=1}^{\tau_K\wedge n} (V(X_k) - PV(X_k)) \\ &\ge -PV(x) + (1-r)\mathbb{E}^x \sum_{k=1}^{\tau_K\wedge n} V(X_k) \end{aligned}$$

where it follows by letting $n \to \infty$,

$$\mathbb{E}^x \sum_{k=1}^{\tau_K} V(X_k) \le \frac{1}{1-r} PV(x), \ \forall x \in E. \tag{3.30}$$

The compact K is then Harris recurrent ([31], p200), hence by [31], Theorem 10.4.7 (in the notation there, $\sum_{n\ge1} \mathbb{P}_x(\tau_K \ge n,\ X_n \in B) =_K U(x, B)$ and $\bar{K} = E$),

$$\mu(B) = \int_K \mu(dx) \sum_{n\ge1} \mathbb{P}_x(\tau_K \ge n,\ X_n \in B) = \int_K \mu(dx) \sum_{k=1}^{\tau_K} 1_B(X_k),\ \forall B \in \mathcal{B}.$$

This implies by (3.30)

$$\int_E V d\mu = \int_K \mu(dx)\mathbb{E}^x \sum_{k=1}^{\tau_K} V(X_k) \le \frac{1}{1-r} \sup_{x\in K} PV(x) < +\infty.$$

Since $U \sim V$, U is μ-integrable too. Hence by (3.29), $\limsup_{n\to\infty} \left(\mathbb{P}_\mu(\sigma_K > n)\right)^{1/n} < r$. Then by Jensen's inequality,

$$\int_E \left(\frac{1}{r}\right)^{\mathbb{E}^x \sigma_K} d\mu(x) \le \mathbb{E}^\mu \left(\frac{1}{r}\right)^{\sigma_K} < +\infty.$$

But by the near-neighbor condition (3.27), with $\mathbb{P}_x$-probability one, $\sigma_K \ge d(x, K) := \inf_{y\in K} d(x, y)$ (distance between x and K). Thus for each $x_0 \in E$ fixed, using $d(x, x_0) \le d(x, K) + \sup_{y\in K} d(x_0, y)$, we have

$$\int_E \left(\frac{1}{r}\right)^{d(x,x_0)} d\mu(x) < +\infty$$

where it follows $\limsup_{N\to\infty}[\mu(B(x_0, N))]^{1/N} \le r$. As $r > r_0$ is arbitrary, (3.28) is proved. □

Remarks (3.vii). For an irreducible and geometric recurrent Markov kernel satisfying (3.27), it is known that $\limsup_{N\to\infty}[\mu(B(x_0, N)^c)]^{1/N} < 1$ by [31], Theorem 16.3.2. In other words (3.28) gives a quantitative exponential estimate of the tail probability of μ. Moreover the lower bound (3.28) of $r_{ess}(P|_{b_u\mathcal{B}})$ is almost sharp by Example 8.3.

A partial inverse implication holds too, see [31], Theorem 16.3.1 in which another function is used in the near-neighbor condition (3.27) instead of the distance $d(x, x_0)$ (such important inverse result was at first obtained by Malyshev and Men'sikov [29](1982), see historical comments in Chap. 16 in [31]). Then it is very interesting to investigate whether $\limsup_{N\to\infty}[\mu(B(x_0, N)^c)]^{1/N} < 1$ together with (3.27) implies $\inf_{U\in\mathcal{A}(P)} r_{ess}(P|_{b_U\mathcal{B}}) < 1$ (and estimate of the last quantity).

4. Several complementary results without (A1)

4.1. A partial answer to the conjecture

We give at first a translation of a deep probabilistic result due to Horowitz [20].

Theorem 4.1. *Let P be a Harris recurrent ([39], p75) Markov kernel. Then $r_{ess}(P|_{b\mathcal{B}}) < 1$ iff $r_\tau(P) = \inf_{n\ge}(\beta_\tau(P^n))^{1/n} < 1$.*

Proof. The necessity follows from Theorem 3.5(b). The important sufficient part follows from Proposition 4.2(a) below, by [39], Chap.VI, Theorem 3.10 due to Horowitz [20] (see Revuz [39] page 179, and historical comments in p320). □

This is a partial answer to the conjecture in Remarks (3.iii). As $\beta_\tau(P^N) \le \lim_{\nu(A)\to 0} \|P^N 1_A\|$, the condition that $r_\tau(P) < 1$ might be easier to verify than Doeblin's condition (1.4).

Proposition 4.2. *Let P be a nonnegative bounded kernel on E with $r_{sp}(P|_{b\mathcal{B}}) = 1$ such that $r_\tau(P) = \lim_{n\to\infty}(\beta_\tau(P^n))^{1/n} < 1$.*

(a) Let $\lambda_0 \in \sigma(P|_{b\mathcal{B}})$ with $|\lambda_0| > r_\tau(P)$. If $\phi \in (b_{\mathbb{C}}\mathcal{B})^*$ (the dual Banach space of $b_{\mathbb{C}}\mathcal{B}$) is a generalized eigenvector of the adjoint operator P^*, i.e., $(P^* - \lambda_0)^m\phi = 0$ for some $m \ge 1$, then ϕ is a (complex valued) measure of bounded variation on E.
(b) (existence of invariant measure) There is $\nu \in M_1(E)$ such that $\nu P = \nu$.

Proof. (a) Let ϕ be a non-zero (complex) valued linear and continuous form on $b_{\mathbb{C}}\mathcal{B}$ such that $(P^* - \lambda_0)^m\phi = 0$ where $|\lambda_0| > r_\tau(P)$. Put $\phi_k := (P^* - \lambda_0)^{m-k}\phi$. We have $\phi_0 = 0$ and

$$P^*\phi_k = \lambda_0\phi_k + \phi_{k-1}, \ k = 1, \cdots, m.$$

Hence for proving that ϕ is a measure, by induction we need only to prove

Claim. *if* $\phi \in (b_{\mathbb{C}}\mathbb{B})^*$ *with* $\|\phi\|^* = 1$ *verifies* $P^*\phi = \lambda_0\phi + \nu$ *where* ν *is a complex valued measure of bounded variation, then* ϕ *is also a measure.*

To this end, notice that for each $N \geq 1$,

$$(P^*)^N\phi = \lambda_0^N\phi + \sum_{k=0}^{n-1}(\lambda_0)^k(P^*)^k\nu,$$

where $(P^*)^k\nu = \nu P^k, k \geq 1$ are again measure. Now for any sequence $(A_n)_{n\geq 0} \subset \mathcal{B}$ decreasing to empty set, we obtain by the expression above,

$$\begin{aligned}\limsup_{n\to\infty} |\lambda_0|^N |\langle\phi, 1_{A_n}\rangle| &= \limsup_{n\to\infty} |\langle(P^*)^N\phi, 1_{A_n}\rangle| \\ &= \limsup_{n\to\infty} |\langle\phi, P^N 1_{A_n}\rangle| \\ &\leq \limsup_{n\to\infty} \|P^N 1_{A_n}\| \leq \beta_\tau(P^N).\end{aligned}$$

Consequently for all $N \geq 1$,

$$\limsup_{n\to\infty} |\langle\phi, 1_{A_n}\rangle| \leq \frac{\beta_\tau(P^N)}{|\lambda_0|^N}$$

where it follows $\limsup_{n\to\infty} |\langle\phi, 1_{A_n}\rangle| = 0$ by letting $N \to \infty$ and using that fact that $|\lambda_0| > r_\tau(P)$. Hence ϕ is a measure, the desired claim.

We prove now part (b). By [30], Theorem 4.1.5, there is a nonnegative linear continuous form $\phi \in (b\mathcal{B})^*$ with $\langle\phi, 1\rangle = 1$ such that $P^*\phi = \phi$. Hence ϕ is a probability measure by part (a). □

With the same proof as above, we have

Proposition 4.3. *Let* P *be a Feller nonnegative bounded kernel on* E *with* $r_{sp}(P|_{C_b(E)}) = 1$ *such that* $r_w(P) = \lim_{n\to\infty}(\beta_w(P^n))^{1/n} < 1$.

(a) Let $\lambda_0 \in \sigma(P|_{C_b(E)})$ *with* $|\lambda_0| > r_w(P)$. *If* $\phi \in (C_b(E))^*$ *(the dual Banach space of the space* $C_b(E; \mathbb{C})$ *of all bounded complex valued continuous functions on* E*) is a generalized eigenvector of the adjoint operator* P^**, i.e.,* $(P^* - \lambda_0)^m\phi = 0$ *for some* $m \geq 1$*, then* ϕ *is a (complex valued) measure of bounded variation on* E.

(b) (existence of invariant measure) There is $\nu \in M_1(E)$ *such that* $\nu P = \nu$.

4.2. What means $r_{ess}(P) < 1$ *in* $b_u\mathcal{B}$*?*

We have evaluated $r_{ess}(P|_{b_U\mathcal{B}})$ by means of different formulas. But whether does, $r_{sp}(P|_{b_U\mathcal{B}}) = 1$ for a Markov kernel? It may seem very easy or obvious, but it is in reality a quite delicate matter. Indeed there may exist some $U \geq 1$ such that $PU = rU$ for some $r > 1$, and hence $r_{sp}(b_U\mathcal{B}) \geq r > 1$. A positive answer can be given by means of $r_\tau(P^U)$. To this end we prepare

Lemma 4.4 (sub-invariant function principle)**.** *Let P be a positive bounded kernel such that $r_\tau(P) \le \lambda$ and there is some everywhere strictly positive function $f \in b\mathcal{B}$ such that $Pf \le \lambda f$. Then $\lambda \ge r_{sp}(P|_{b\mathcal{B}})$.*

Proof. Assume in contrary that $\lambda < r_{sp}(P|_{b\mathcal{B}}) =: r_0$. Then $r_0 > \lambda \ge r_\tau(P)$. Thus by Proposition 4.2, there would be some probability measure ν such that $P^*\nu = \nu P = r_0 \nu$. We get so

$$\lambda\nu(f) \ge \nu(Pf) = (\nu P)(f) = r_0\nu(f).$$

As $\nu(f) > 0$, we obtain $\lambda \ge r_0$, a contradiction. □

Proposition 4.5. *Let P be a positive kernel, bounded on $b_U\mathcal{B}$ where $U \ge 1$ is a measurable function, and $P^U(x, dy) = (U(y)/U(x))P(x, dy)$. Then*

$$r_{sp}(P|_{b_U\mathcal{B}}) = \max\{r_\tau(P^U),\ \limsup_{n\to\infty} \left(\|P^n g\|_U\right)^{1/n}\} \tag{4.1}$$

for every strictly positive $g \in b_U\mathcal{B}$. In particular if P is moreover Markov, and $r_{ess}(P|_{b_U\mathcal{B}}) \le 1$, then $r_{sp}(P|_{b_U\mathcal{B}}) = 1$.

The quantity "$\limsup_{n\to\infty} (\|P^n g\|_U)^{1/n}$" is called "individual spectral radius of P at g" in [30]. There are Banach lattice and positive operator A on it such that every individual spectral radius of A is zero, but the spectral radius of A is > 0, see [30], Chap.4.

Proof. Since $r_{sp}(P|_{b_U\mathcal{B}}) \ge r_{ess}(P|_{b_U\mathcal{B}}) = r_{ess}(P^U|_{b\mathcal{B}}) \ge r_\tau(P^U)$, the inequality "$\ge$" in (4.1) is true. For the inverse inequality for each strictly positive $g \in b_U\mathcal{B}$ fixed, let r be an arbitrary number strictly greater than the r.h.s. of (4.1). Put

$$h := \sum_{k=0}^{\infty} \frac{1}{r^n} P^n g.$$

It belongs again to $b_U\mathcal{B}$ and $Ph \le rh$. By Lemma 4.4 (applied to P^U and $f := h/U$), $r \ge r_{sp}(P|_{b_U\mathcal{B}})$, where follows the inverse inequality. The last claim follows from (4.1) with $g = 1$ by noting that $r_\tau(P^U) \le r_{ess}(P|_{b_U\mathcal{B}}) \le 1$. □

We now extend a known difficult result, Theorem 3.7 in Revuz [39], Chap.6 from $b\mathcal{B}$ to $b_U\mathcal{B}$, whose proof is given in Section 9.

Theorem 4.6. *Let P be a Markov kernel bounded on $b_U\mathcal{B}$, where $U \ge 1$ is a measurable function. Assume that $r_{ess}(P|_{b_U\mathcal{B}}) < 1$. There exist $k \ge 1$, and for each $j = 1, \cdots, k$, $d_j \ge 1$ nonnegative functions $U_{i,j} \in b\mathcal{B}$ ($i = 1, \cdots, d_j$) such that $\sum_{j=1}^k \sum_{i=1}^{d_j} U_{i,j} = 1$ over E, and probability measures $\mu_{i,j}$ carried by the pairwise disjoint sets $E_{i,j} := [U_{i,j} = 1]$ such that, if d denotes the least common multiple of the d_j ($j = 1, \cdots, k$), then for every $l \in \mathbb{N}$,*

$$\left\| P^{nd+l} - \sum_{j=1}^{k}\sum_{i=1}^{d_j} U_{i-l(mod\ d_j),j} \otimes \mu_{i,j} \right\|_{b_U\mathcal{B}\to b_U\mathcal{B}} \to 0 \ (as\ n \to \infty) \tag{4.2}$$

and this convergence is geometric. The sets $E_j = \bigcup_{i=1}^{d_j} E_{i,j}$ *is* P*-absorbing (i.e.,* $P(x, E_j) = 1$ *for every* $x \in E_j$*), and* P *restricted to* E_j *is Harris recurrent with invariant measure* $\mu_j := \sum_{i=1}^{d_j} \mu_{i,j}$*.*

4.3. Coincidence of $r_{ess}(P)$ *on* $C_b(E)$ *and on* $b\mathcal{B}$

For a Feller nonnegative bounded kernel $P(x, dy)$, it might be more natural or more fruitful to study its essential spectral radius as an operator on $C_b(E)$. More natural, perhaps; more fruitful? No, as claimed by

Proposition 4.7. *Let* $P(x, dy)$ *be a Feller nonnegative bounded kernel on* E*. Then* $r_{ess}(P|_{C_b(E)}) = r_{ess}(P|_{b\mathcal{B}})$*, and for* $\lambda \in \mathbb{C}$ *with* $|\lambda| > r_{ess}(P|_{C_b(E)})$*,* $\lambda \in \sigma(P|_{C_b(E)})$ *iff* $\lambda \in \sigma(P|_{b\mathcal{B}})$*, and in the last case the eigenprojection* E_λ *of* $P|_{C_b(E)}$ *is a (complex valued) Feller kernel coinciding with that of* $P|_{b\mathcal{B}}$*.*

Proof. At first by Nussbaum formula (2.7) and the obvious fact that $\beta(P|_{C_b(E)}) \le \beta(P|_{b\mathcal{B}})$, we have $r_{ess}(P|_{C_b(E)}) \le r_{ess}(P|_{b\mathcal{B}})$.

For $\lambda \in \sigma(P|_{C_b(E)})$ with $|\lambda| > r_{ess}(P|_{C_b(E)})$, let E_λ be the associated eigenprojection. By Proposition 3.2(a), $r_w(P) \le r_{ess}(P|_{C_b(E)})$ (equality is false by the example in Remarks (3.i)). Hence $|\lambda| > r_w(P)$. By Proposition 4.3(a), every element in the range of the adjoint eigenprojection E_λ^* acting on $C_b(E, \mathbb{C})^*$, being a generalized eigenvector of P^*, is a complex valued measure of bounded variation. Then by the expression (2.2), E_λ is a Feller kernel on E.

Now for any $r > r_{ess}(P|_{C_b(E)})$ such that $\Gamma_r := \{z \in \mathbb{C};\ |z| = r\}$ contains no spectral point of $P|_{C_b(E)}$, let Π_r be the sum of all eigenprojections E_λ associated with $\lambda \in \sigma(P|_{C_b(E)})$ such that $|\lambda| > r$. Π_r is again a finite dimensional projection, given by a Feller kernel, and there is some constant $C > 0$ such that

$$\|[P(I - \Pi_r)]^n\|_{C_b(E;\mathbb{C})} \le Cr^n, \ \forall n \ge 1.$$

The key remark is: for a Feller kernel such as $[P(I - \Pi_r)]^n$, its norm on $C_b(E; \mathbb{C})$ coincides with that on $b_{\mathbb{C}}\mathcal{B}$. Consequently

$$\|[P(I - \Pi_r)]^n\|_{b_{\mathbb{C}}\mathcal{B}} \le Cr^n, \ \forall n \ge 1.$$

This implies not only $r_{ess}(P|_{b\mathcal{B}}) \le r$ (then the desired inverse inequality "$r_{ess}(P|_{C_b(E)}) \ge r_{ess}(P|_{b\mathcal{B}})$"), but also all other conclusions of this proposition. □

Remarks (4.i). The above result is largely inspired by a well known result in [39]: a Feller kernel is compact in $C_b(E)$ iff it is so on $b\mathcal{B}$. It explains also why our study is concentrated on $b\mathcal{B}$ and $b_u\mathcal{B}$, for the structure of $C_b(E)$ hidden the role of the parameter β_τ (and (A1)), basic in the investigation of $r_{ess}(P)$ by Theorem 3.5 and Theorem 4.1 (and again more if our conjecture in Remark (3.iii) is true).

Let $\mathcal{A}_C(P) := \{u \in \mathcal{A}(P);\ u, Pu$ are continuous on $E\}$ and $C_u(E)$ the subspace of all $f \in b_u\mathcal{B}$ which are continuous. Given $u \in \mathcal{A}_C(P)$, using the similarity of $P|_{C_u(E)}$ and $P^u|_{C_b(E)}$ (P^u is still Feller by Remarks (3.vi)), one can transport the above result to $C_u(E)$.

5. Essential spectral radius on $L^p(\mu)$

5.1. A general formula: Persson type's principle

We begin with Persson type's principle ([7], Theorem 3.12), corresponding to Theorem 3.5.

Theorem 5.1. *Let $P(x, dy)$ be a nonnegative bounded kernel, bounded on $L^p(\mu) = L^p(E, \mathcal{B}, \mu)$, where $1 < p \le +\infty$. Assume that for any $K \subset\subset E$, $\beta_\tau(1_K P) = 0$ and $\|P\|_a < +\infty$ for some $1 \le a < p$. Then*

$$r_{ess}\left(P|_{L^p(\mu)}\right) = \inf_{K\subset\subset E} r_{sp}\left(1_{K^c} P 1_{K^c}|_{L^p(\mu)}\right). \tag{5.1}$$

Remarks (5.i). For continuous time and symmetric Markov semigroups (P_t), under the strong Feller condition, Grillo [19] (1998) proved the above Persson principle in L^2 (with an extra condition on the metric generated by the Dirichlet form), and F.Y. Wang [41] (2000) found an infinitesimal criterion for exactly evaluating $r_{ess}\left(P_t|_{L^2(\mu)}\right)$. We shall return to extend their works in the sequel paper.

Remarks (5.ii). Let

$$\|P\|_{tail(L^p(\mu))} := \lim_{L\to+\infty} \sup_{f:\|f\|_p\le 1} \|1_{[|Pf|>L]} Pf\|_p, \ \forall p \in [1, +\infty)$$
$$\|P\|_{tail(L^\infty(\mu))} := \limsup_{\mu(A)\to 0} \|P1_A\|_\infty := \lim_{\varepsilon\to 0+} \sup_{A\in\mathcal{B},\mu(A)<\varepsilon} \|P1_A\|_\infty$$

be the tail norm of P (see [18], [49]). For $1 < p < +\infty$, it coincides with the *measure of non-semi-compactness* given by de Pagter-Schep [12](1988) (see [18]). Weis [43](1984) and de Pagter-Schep [12](1988) proved the following Nussbaum formula of essential spectral radius:

Let $Pf(x) := \int_E p(x, y) f(y)\mu(dy)$ be a nonnegative absolute continuous kernel operator, bounded on $L^p(\mu)$, where $1 < p < +\infty$. Then

$$r_{ess}(P|_{L^p(\mu)}) := \lim_{n\to\infty} (\|P^n\|_{tail(L^p(\mu))})^{1/n}.$$

Here the absolute continuity of P is crucial (otherwise the infinite dimensional Ornstein-Uhlenbeck semigroup provides such a counter-example).

Remarks (5.iii). Without assumption of Theorem 5.1, the following inequality

$$r_{ess}(P|_{L^p(\mu)}) \ge \inf_{K\subset\subset E} r_{sp}\left(1_{K^c} P 1_{K^c}|_{L^p(\mu)}\right) \tag{5.2}$$

holds always. Indeed, it is obvious that the measure of non-compactness defined in (2.5) verifies $\beta(P^n) \ge \|P^n\|_{tail(L^p(\mu))}$. But by [18] or [49](Lemma 3.1),

$$\begin{aligned}\|P^n\|_{tail(L^p(\mu))} &= \limsup_{\mu(A),\mu(B)\to 0} \|1_A P(1_B\cdot)\|_p \\ &\ge \inf_{K\subset\subset E} \|1_{K^c} P^n 1_{K^c}\|_p \ge \inf_{K\subset\subset E} \|(1_{K^c} P 1_{K^c})^n\|_p.\end{aligned}$$

Hence (5.2) follows by Nussbaum's formula (2.7).

Proof of Theorem 5.1. Denote the measure of non-compactness of P in L^p by $\beta(P)$. For each compact $K \subset\subset E$ and any $n \geq 2$,

$$\begin{aligned} P^n &= (1_K P + 1_{K^c} P)^n \\ &= (1_{K^c} P)^n + (1_{K^c} P)^{n-1} 1_K P + \sum_{j \in \{0,1\}^n:\ j_k = 0 \text{ for some } k \leq n-1} A_{j_1} \cdots A_{j_n} \end{aligned}$$

where $A_0 := 1_K P$ and $A_1 := 1_{K^c} P$. By Proposition 3.2(g), if $j_k = 0$ for some $1 \leq k \leq n-1$, product $A_{j_1} \cdots A_{j_n}$ is compact on $b\mathcal{B}$, then on L^∞, hence on L^p by interpolation of compact operator (and its boundedness on L^a). In other words $\beta(A_{j_1} \cdots A_{j_n}) = 0$. Consequently by $\beta(P) \leq \|P\|_p$, we have

$$\beta(P^n) = \beta\left((1_{K^c} P)^n + (1_{K^c} P)^{n-1} 1_K P)\right) \leq 2\|(1_{K^c} P 1_{K^c})^{n-1}\|_p \cdot \|P\|_p.$$

Hence by Nussbaum's formula (2.7),

$$r_{ess}\left(P|_{L^p(\mu)}\right) \leq \lim_{n\to\infty} \left(2\|(1_{K^c} P 1_{K^c})^{n-1}\|_p \cdot \|P\|_p\right)^{1/n} = r_{sp}\left(1_{K^c} P 1_{K^c}|_{L^p(\mu)}\right).$$

This, together with (5.2), yields (5.1). □

For a Markov kernel P with invariant and ergodic probability measure μ, a quite natural question is to determine whether the spectral radius of $1_{K^c} P 1_{K^c}$ (the transition kernel killed at K) in $L^p(\mu)$ is < 1. The previous theorem yields a satisfactory answer:

Corollary 5.2. *Let P be a Markov kernel with invariant probability measure μ. Assume that P is ergodic w.r.t. μ, i.e., if $f \in b\mathcal{B}$ verifies $Pf = f$, then f is $\mu-a.s.$ constant. Assume that $\beta_\tau(1_K P) = 0$ for any $K \subset\subset E$.*

Let $A \in \mathcal{B}$ be a relatively compact subset of E charged by μ and $1 < p \leq +\infty$. Then $r_{sp}(1_{A^c} P 1_{A^c}|_{L^p(\mu)}) < 1$ iff $r_{ess}(P|_{L^p(\mu)}) < 1$.

Proof. P is contractive on $L^q(\mu)$ for every $q \in [1, +\infty]$. Since the closure $\overline{A}$ is compact, we can apply Theorem 5.1 and get

$$r_{sp}(1_{A^c} P 1_{A^c}|_{L^p(\mu)}) \geq r_{sp}(1_{\overline{A}^c} P 1_{\overline{A}^c}|_{L^p(\mu)}) \geq r_{ess}(P|_{L^p(\mu)}).$$

Then the necessity is true.

For the sufficiency, we observe by the inequality above that there are only two possibilities: either $r_{sp}(1_{A^c} P 1_{A^c}|_{L^p(\mu)}) = r_{ess}(P|_{L^p(\mu)})$ or $r_{sp}(1_{A^c} P 1_{A^c}|_{L^p(\mu)}) > r_{ess}(P|_{L^p(\mu)})$. In the first case $r_{sp}(1_{A^c} P 1_{A^c}|_{L^p(\mu)}) < 1$ by the sufficiency's assumption. It remains to treat the second case.

In such case $r(A^c) := r_{sp}(1_{A^c} P 1_{A^c}|_{L^p(\mu)})$ is strictly greater than $r_{ess}(1_{A^c} P 1_{A^c}|_{L^p(\mu)})$, because

$$r_{ess}(1_{A^c} P 1_{A^c}|_{L^p(\mu)}) = r_{ess}(P|_{L^p(\mu)})$$

by Theorem 5.1. Thus $r(A^c)$ is an isolated eigenvalue of $1_{A^c} P 1_{A^c}$ corresponding to some nonnegative eigenfunction $h \in L^p(\mu)$ with $\|h\|_p = 1$ by Proposition 2.3(b).

If in contrary $r(A^c) = 1$, we would get

$$Ph \geq 1_{A^c} P 1_{A^c} h = h.$$

Since $\mu(Ph) = \mu(h)$, $Ph = h,\ \mu - a.s.$. But by the assumed ergodicity, h should be μ-a.s. a positive constant. This is in contradiction with the fact that $h = 0$, $\mu - a.e.$ on A. □

5.2. Symmetric case.

Lemma 5.3. *Let $P : L^2(\mu) \to L^2(\mu)$ be a nonnegative, bounded and symmetric operator.*

(a) Then for any two $f, g > 0,\ \mu - a.e.$ in $L^2(\mu)$, we have

$$\|P\|_2 = r_{sp}(P|_{L^2(\mu)}) = \limsup_{n\to\infty} \langle P^n f, g\rangle^{1/n}. \tag{5.3}$$

(b) If moreover P is μ-essentially irreducible (i.e., for some $\lambda > r_{sp}(P|_{L^2(\mu)})$, $R_\lambda 1_A := \sum_{k=0}^{\infty} \lambda^{-n-1} P^n 1_A > 0,\ \mu - a.e.$ for any $A \in \mathcal{B}$ with $\mu(A) > 0$), then (5.3) holds for any $0 \leq f, g \in L^2(\mu)$ such that $\mu(f) \wedge \mu(g) > 0$.

Proof. **(a)** The first equality is an immediate consequence of spectral decomposition.

Let r_0 be the quantity at the r.h.s. of (5.3). Obviously $\|P\|_2 \geq r_0$. Below we show the inverse inequality which is the core of this lemma.

If $\|P\|_2 = 0$, it is obviously true. In the case where $\|P\|_2 > 0$, we may assume that $\|P\|_2 = 1$ without loss of generality. By absurd assume that $r_0 < 1$. Put

$$\mathcal{D} := \{h \in L^2(\mu);\ |h| \leq C \cdot (f \wedge g), \mu - a.e. \text{ for some } C > 0\}.$$

Then $\mathcal{D}$ is dense in $L^2(\mu)$ for $f, g > 0,\ \mu - a.e.$. By the definition of r_0 and the non-negativeness of P, we have for each $r \in (r_0, 1)$ fixed,

$$\lim_{n\to\infty} r^{-n} \langle P^n h, h\rangle = 0,\ \forall h \in \mathcal{D}.$$

Now write the spectral decomposition: $P^2 = \int_{[0,1]} \lambda^2 dE_\lambda$. We have for each $h \in \mathcal{D}$,

$$\|(E_1 - E_r)h\|_2^2 \leq r^{-2n} \int_{(r,1]} \lambda^{2n} d\langle E_\lambda h, h\rangle \leq r^{-2n} \langle P^{2n} h, h\rangle \to 0.$$

As $\mathcal{D}$ is dense in L^2 and $E_1 - E_r$ is bounded on $L^2(\mu)$, we obtain $E_1 = E_r$, which means that $\|P^2\|_2 \leq r^2 < 1$, a contradiction with our assumption that $\|P^2\|_2 = \|P\|_2^2 = 1$.

(b) Instead of $\mathcal{D}$ given above, take

$$\mathcal{D}' := \{h \in L^2(\mu);\ |h| \leq C_N \cdot (Q_N f \wedge Q_N g), \mu - a.e. \text{ for some } N \geq 1,\ C_N > 0\}$$

where $Q_N = I + P + \cdots + P^N$. By the essential irreducibility, $\mathcal{D}'$ is dense in $L^2(\mu)$. Moreover for any r strictly greater than the r.h.s. r_0 of (5.3),

$$\lim_{n\to\infty} r^{-n} \langle P^n h, h\rangle = 0,\ \forall h \in \mathcal{D}'.$$

The remained proof is the same as that of part (a), so omitted. □

With exactly the same proof of part (a) above, we also have the following result for not-necessarily nonnegative operator, which will be crucial for our investigation of spectral gap (a similar result is proved by the author in [44], Proposition 2.9 in the continuous time case):

Lemma 5.4. *Let P be a symmetric bounded operator on $L^2(\mu)$. Assume that there is a dense subset $\mathcal{D} \subset L^2(\mu)$ and some constant $r_0 > 0$ such that $\forall f \in \mathcal{D}$, $\exists C(f) > 0$:*

$$|\langle P^n f, f\rangle| \le C(f) r_0^n, \ \forall n \ge 1,$$

then $r_{sp}(P) \le r_0$.

The following result answers completely Question 4 in the Introduction in the symmetric case, but leaves it open in the non-symmetric case.

Theorem 5.5. *Let P be a nonnegative kernel satisfying (A1), symmetric and bounded on $L^2(\mu)$ (μ being a probability measure on E).*

(a) For each measurable subset $A \subset E$,

$$\begin{aligned} r_{sp}\left(1_{A^c} P 1_{A^c}|_{L^2(\mu)}\right) &= \inf\left\{r;\ \exists 1 \le u \in L^2(\mu),\ 1_{A^c} Pu \le ru,\ \mu - a.s.\right\} \\ &= \inf\left\{r;\ \exists u \in \mathcal{B}_\mu^+(A^c),\ 1_{A^c} P 1_{A^c} u \le ru,\ \mu - a.s.\right\} \end{aligned} \tag{5.4}$$

where $\mathcal{B}_\mu^+(A^c) := \{u : A^c \to [0, +\infty]$ measurable$;\ 0 < u < \infty,\ \mu - a.e.\}$. In particular

$$\begin{aligned} r_{ess}(P|_{L^2(\mu)}) &= \inf_{K \subset\subset E} \inf_{1 \le u \in L^2(\mu)} esssup_{x \in K^c} \frac{Pu(x)}{u(x)} \\ &= \inf_{K \subset\subset E} \inf_{u \in \mathcal{B}_\mu^+(K^c)} esssup_{x \in K^c} \frac{P(1_{K^c} u)(x)}{u(x)} \\ &\le \inf_{U \in \mathcal{A}(P)} r_{ess}(P|_{b_U \mathcal{B}}) \end{aligned} \tag{5.5}$$

and the last inequality becomes equality in one of the following cases:

(a.i) E is countable (discrete) and μ charges every point of E;

(a.ii) $P(x, dy) = p(x, y)\mu(dy)$ where $p(x, y)$ is $\mathcal{B} \times \mathcal{B}$-measurable, and $\int_E p(x, y)^2 \mu(dy) < +\infty,\ \forall x \in E$.

(b) If P is moreover Markov, then for any $A \in \mathcal{B}$,

$$r_{sp}\left(1_{A^c} P 1_{A^c}|_{L^2(\mu)}\right) = esssup_{x \in A^c} \limsup_{n \to \infty} \left[\mathbb{P}_x(\sigma_A = n)\right]^{1/n}. \tag{5.6}$$

$$\begin{aligned} r_{ess}(P|_{L^2(\mu)}) &= \inf\left\{r > 0;\ \exists K \subset\subset E,\ \mathbb{E}^x \left(\frac{1}{r}\right)^{\sigma_K} \right. \\ &\qquad \left. < +\infty,\ \mu - a.e.\ x\right\} \\ &= \inf_{K \subset\subset E} esssup_{x \in K^c} \limsup_{n \to \infty} \left[\mathbb{P}_x(\sigma_K = n)\right]^{1/n}. \end{aligned} \tag{5.7}$$

Proof. (a) Denote the three terms of (5.4) by r_0, r_1, r_2 respectively. Obviously $r_1 \geq r_2$. We show at first $r_2 \geq r_0 = r_{sp}\left(1_{A^c} P 1_{A^c}|_{L^2(\mu)}\right)$.

For any $r > r_2$, there is some $u \in \mathcal{B}_\mu^+(A^c)$ such that $1_{A^c} P 1_{A^c} u \leq ru,\ \mu - a.s.$. Now put $v := u/(1+u)^2$. We have

$$(1_{A^c} P 1_{A^c})^n\, v \leq (1_{A^c} P 1_{A^c})^n\, u \leq r^n u,\ \mu - a.e.$$

Consequently

$$\langle v, (1_{A^c} P 1_{A^c})^n\, v\rangle_\mu \leq r^n \langle v, u\rangle_\mu \leq r^n.$$

Thus by Lemma 5.3, $r_{sp}\left(1_{A^c} P 1_{A^c}|_{L^2(\mu)}\right) \leq r$, then $\leq r_2$ for $r > r_2$ is arbitrary.

It remains to show that $r_0 \geq r_1$ by a similar argument in $(a) \Rightarrow (c)$ of Theorem 3.10. For any $r > r_0 = r_{sp}\left(1_{A^c} P 1_{A^c}|_{L^2(\mu)}\right)$, let

$$h := 1_{A^c} + \sum_{n=1}^{\infty} r^{-n}\,(1_{A^c} P 1_{A^c})^n\, 1 \in L^2(\mu).$$

Then $h \in L^2(\mu)$ and $1_{A^c} Ph \leq rh$. Now for any $\varepsilon > 0$, consider $u(x) = 1 + Lh(x)$ where $L \geq \|P\|_2/\varepsilon$, we have $\mu - a.s.$,

$$\frac{1_{A^c} Pu}{u} \leq 1_{A^c} \frac{\|P\|_2 + L 1_{A^c} Ph}{u} \leq 1_{A^c} \frac{\|P\|_2 + Lrh}{1 + L \cdot h} \leq r + \varepsilon$$

which entails $r_1 \leq r + \varepsilon$. As $r > r_0$ and $\varepsilon > 0$ are arbitrary, we obtain the desired claim "$r_1 \leq r_0$". So (5.4) is established.

The first and second equalities in (5.5) follow immediately from (5.4) and Theorem 5.1, and the third inequality in (5.5) by Remark (3.v). Theorem 3.11(b) implies that the third inequality in (5.5) becomes equality in the case (a.i), for $esssup_{K^c} = \sup_{K^c}$. In case (a.ii), let $r > r_{ess}(P|_{L^2(\mu)})$ be arbitrary. Then by Theorem 5.1, there is some compact $K \subset\subset E$ such that $r > r_{sp}(1_{K^c} P 1_{K^c}|_{L^2(\mu)})$. Put as above

$$h := 1_{K^c} + \sum_{n=1}^{\infty} r^{-n}\,(1_{K^c} P 1_{K^c})^n\, 1 \in L^2(\mu).$$

By assumption in (a.ii), $Ph(x) < +\infty$ for every $x \in E$. Then $h(x) \leq 1_{K^c}(x) + Ph(x)/r < +\infty, \forall x \in E$. Now letting $u(x) := 1 + Lh(x)$, for any $\varepsilon > 0$, we have as in the proof of $r_0 \geq r_1$ above that for L large enough,

$$1_{K^c} \frac{Pu(x)}{u(x)} \leq r + \varepsilon.$$

This entails by Theorem 3.11(b) that $\inf_{U \in \mathcal{A}(P)} r_{ess}(P|_{b_U \mathcal{B}}) \leq r + \varepsilon$, the desired inverse inequality.

(b) For (5.6), let r be any number strictly greater than the r.h.s. of (5.6), put

$$u(x) := \mathbb{E}^x \left(\frac{1}{r}\right)^{\sigma_A}.$$

Then μ-a.s., $0 < u(x) < \infty$, i.e., $u \in \mathcal{B}_\mu^+(A^c)$, and $(1_{A^c}P1_{A^c})u \le ru$ on A^c. Hence by part (a), $r_{sp}\left(1_{A^c}P1_{A^c}|_{L^2(\mu)}\right) \le r$. That proves "$\le$" in equality (5.6).

Inversely for any $r > r_{sp}\left(1_{A^c}P1_{A^c}|_{L^2(\mu)}\right)$, by part (a), there is some $u \in L^2(\mu)$ such that $u(x) \ge 1$ on E and $(1_{A^c}P1_{A^c})u \le ru$, $\mu - a.e.$ on A^c. Thus for $\mu - a.e.\ x \in A^c$,

$$\mathbb{P}_x(\sigma_A > n) \le \mathbb{E}^x 1_{[\sigma_A > n]}u(X_n) = (1_{A^c}P1_{A^c})^n u(x) \le r^n u(x)$$

where it follows that r is not smaller than the r.h.s. of (5.6). In other words (5.6) is established.

The second equality in (5.7) is elementary. By Theorem 5.1, the first equality in (5.7) follows from (5.6). □

Remarks (5.iv). Whether $\inf_{U\in\mathcal{A}(P)} r_{ess}(P|_{b_U\mathcal{B}}) < 1$ does imply $r_{ess}(P|_{L^2(\mu)}) < 1$ for a nonsymmetric Markov kernel P with invariant measure μ is a very interesting open question (see however Chen [3] for a confirmative answer for several important situations). Of course it is again more difficult to extend Theorem 5.5 to the non-symmetric case.

Remarks (5.v). Let P be a nonsymmetric Markov kernel with invariant measure μ. Then PP^* is Markov, symmetric in $L^2(\mu)$. By following the general spectral theory ([38], Vol.IV), one can show that $r_{ess}(P|_{L^2(\mu)}) \le \sqrt{r_{ess}(PP^*|_{L^2(\mu)})}$. And one can apply Theorem 5.5 to estimate $r_{ess}(PP^*|_{L^2(\mu)})$. But the range of this approach is limited: in general μ is unknown in practice and then the adjoint operator P^*, depending on μ, is difficult to calculate.

L. Miclo [32], by exploring this approach, has investigated the logarithmic Sobolev inequality of P.

6. Relations with the large deviation principles

In this section we assume that $P(x, dy)$ is a Markov kernel (i.e., $P \ge 0$ and P1=1). We apply the previous results to large deviations.

6.1. An extension of a result of de Acosta

The purpose of this paragraph is to give applications of the two parameters β_τ and β_w to large deviations, without assumption (A1). Let

$$L_n(\omega) := \frac{1}{n}\sum_{k=1}^n \delta_{X_n(\omega)},\ n \ge 1$$

be the empirical measures of our Markov process $(X_n)_{n\ge1}$. It is a random element in $M_1(E)$ equipped with the $\sigma-$field $\sigma(\nu \to \nu(f);\ f \in b\mathcal{B})$ (which is smaller than the σ-field generated by the τ-open subsets). A subset in $M_1(E)$ is called *measurable*, if it is an element in the last $\sigma-$field.

The Donsker-Varadhan entropy $J : M_1(E) \to [0, +\infty]$ is given by

$$J(\nu) := \sup_{1 \le u \in b\mathcal{B}} \int_E \log \frac{u}{Pu} d\nu. \tag{6.1}$$

de Acosta [8] showed that if $\beta_\tau(P^N) = 0$ (resp. P is Feller and $\beta_w(P^N) = 0$) for some $N \ge 1$, then the empirical occupation measures (L_n) satisfies a uniform large deviation upper bound with good rate function given by J, on $M_1(E)$ w.r.t. the τ-topology (resp. "w"-topology). The following is an extension of his result, which supports our conjecture after Theorem 3.5.

Theorem 6.1. *Let P be a Markov kernel and $(\Omega, (X_n)_{n\ge0}, (\mathbb{P}_x)_{x\in E})$ be the associated Markov process.*

(a) If

$$r_\tau(P) := \lim_{n\to\infty} \left[\beta_\tau(P^n)\right]^{1/n} = \inf_{n\ge1} \left[\beta_\tau(P^n)\right]^{1/n} = 0, \tag{6.2}$$

then (L_n) satisfies the uniform good upper bound of large deviation on $(M_1(E), \tau)$, more precisely,

(τ-GRF) J is a Good Rate Function on $(M_1(E), \tau)$, i.e., $[J \le L]$ is τ-compact for each $L \ge 0$;

(uniform τ-ULD) (uniform Upper bound of Large Deviation) For each closed measurable set $F \subset (M_1(E), \tau)$,

$$\limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in E} \mathbb{P}_x(L_n \in F) \le -\inf_{\nu\in F} J(\nu).$$

If moreover P is irreducible w.r.t. some maximal probability measure μ (i.e., $\mu \ll \sum_{k=1}^\infty 2^{-n-1} P^n(x, \cdot)$ for any $x \in E$ (μ-irreducibility) and $\mu P \ll \mu$ (maximality)), such that

$$J(\nu) < +\infty \Longrightarrow \nu \ll \mu; \tag{6.3}$$

then

(uniform τ-LLD) (uniform Lower bound of Large Deviation) For each open measurable set $G \subset (M_1(E), \tau)$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \inf_{x\in E} \mathbb{P}_x(L_n \in G) \ge -\inf_{\nu\in G} J(\nu).$$

When those three properties hold, we say that (L_n) satisfies the uniform large deviation principle w.r.t. τ-topology (uniform τ-LDP in short).

(b) Assume that P is Feller. If

$$r_w(P) := \lim_{n\to\infty} \left[\beta_w(P^n)\right]^{1/n} = \inf_{n\ge1} \left[\beta_w(P^n)\right]^{1/n} = 0, \tag{6.4}$$

then (L_n) satisfies the uniform good upper bound of large deviation on $(M_1(E), w)$, i.e., satisfying (w-GRF) and (uniform w-ULD) which are defined as above with the τ-topology replaced by the weak convergence topology "w".

Its proof is postponed to §10.

Remarks (6.i). The study of large deviations of Markov processes was openned by Donsker-Varadhan [16], see Deuschel-Stroock [13] and Dembo-Zeitouni [11] for historical comments and references.

The existence and uniqueness of invariant probability measure μ and the μ-irreducibility, together with (6.3) are all necessary to the LDP of (L_n) w.r.t. τ-topology, uniform or pointwise, see [47]. To see it more clearly let us consider a two-points Markov chain with $P(1,1) = P(1,2) = 1/2$ and $P(2,2) = 1$. Of course $r_{ess}(P) = 0$. It is obvious that

$$\mathbb{P}_{x=2}(L_n = \delta_0) = 0$$

But $J(\delta_0) = \log 2$, the lower bound of large deviation with rate function J above is not true.

Condition (6.3) was isolated by de Acosta [9] and Jain [21] for the lower bound of large deviations with rate function given by the Donsker-Varadhan entropy J. In the general irreducible case (or even μ-essentially irreducible case), the rate function governing the lower bound is given by the modified Donsker-Varadhan entropy

$$J_\mu(\nu) := \begin{cases} J(\nu) & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise.} \end{cases}$$

See [9], [21] for the irreducible case and [46] for the μ-essentially irreducible case. Condition (6.3) just means $J_\mu = J$. A simple sufficient condition isolated in [9], [21] for (6.3) is: for some $N \geq 1$, $P^N(x, dy) \ll \mu(dy)$ for every $x \in E$.

Remarks (6.ii). As the $\tau-$topology is much stronger than "w", then the uniform τ-LDP is stronger than that w.r.t. "w".

Now we turn to a particular case where (6.2) and (6.4) can be shown to be equivalent to the uniform τ-LDP.

6.2. Equivalence between LDP and $r_{ess}(P) = 0$

Consider the following assumptions:

(A2) P is Feller and P^N is strongly Feller for some $N \geq 1$;

(A3) P is topologically transitive.

Here the topological transitivity means that for each $x \in E$ and for each non-empty open subset $G \subset E$, there is some $n \geq 0$ such that $P^n(x, G) > 0$.

Lemma 6.2. *Assume (A2) and (A3). Fix some $x_0 \in E$ and define $\mu(A) := R(x_0, A) := \sum_{k=0}^{\infty} 2^{-k-1} P^{N+k}(x_0, A)$ for any $A \in \mathcal{B}$. Then for any $x \in E$, $R(x, \cdot)$ is equivalent to $R(x_0, \cdot)$. In particular P is μ-irreducible and $J = J_\mu$, i.e., (6.3) holds.*

Proof. It is enough to show that $P^n(x, dy) \ll R(x_0, dy)$ for all $x \in E$ and $n \geq N$. Indeed, that implies $R(x, dy) \ll R(x_0, dy)$ and also $R(x_0, dy) \ll R(x, dy)$ by changing role of x and x_0. And (6.3) follows by the last sentence in Remarks (6.i).

For the claim above, it suffices to establish that $P^n(x, A) = 0$, $\forall x \in E$ for any $A \in \mathcal{B}$ verifying $R(x_0, A) = 0$. To that end assume in contrary that $P^n 1_A$ is not identically zero over E. Since P^n $(n \geq N)$ is strongly Feller by (A2), $P^n 1_A$ is continuous. Hence by (A3),

$$R(x_0, A) \geq 2^{N-n} \sum_{k=0}^{\infty} 2^{-k-1} \left(P^k P^n 1_A \right)(x_0) > 0$$

a contraction with the assumption $R(x_0, A) = 0$, proving the desired claim. □

Theorem 6.3. *Given a Markov kernel P satisfying (A2) and (A3). These properties are equivalent:*

(a) (L_n) *satisfies the uniform* τ*-LDP.*
(b) (uniform hyper-exponential recurrence) For any $r > 0$*, there is some compact* $K \subset\subset E$ *such that*

$$\sup_{x \in E} \mathbb{E}^x \left(\frac{1}{r} \right)^{\sigma_K} < +\infty.$$

(c) $r_{ess}(P|_{b\mathcal{B}}) = 0$.
(d) $\inf_{K \subset\subset E} \inf_{1 \leq u \in b\mathcal{B}} \sup_{x \in K^c} \frac{Pu(x)}{u(x)} = 0$.
(e) $r_w(P) = 0$.
(f) $r_\tau(P) = 0$.

Proof. In [47], we have proved the equivalence between (a) and (b) under (A2)+(A3). By Lemma 3.12(a) and (3.11) in Theorem 3.10, we have the equivalence between (b) and (c). The equivalence between (c) and (d) follows from Theorem 3.10.

The equivalence between (c) and (e) follows from Theorem 3.5, as well as $(c) \Rightarrow (f)$. Finally $(f) \Rightarrow (a)$ follows from Theorem 6.1. □

Theorem 6.4. *Given a Markov kernel P satisfying (A2) and (A3). These properties are equivalent:*

(a) $\mathbb{P}_x(L_n \in \cdot)$ *satisfies uniformly for initial states in the compacts, the large deviation principle of Donsker-Varadhan (called locally uniform LDP simply) on* $(M_1(E), \tau)$ *. More precisely in the statement of Theorem 6.1(a), (*τ*-GJF) is true, and the upper bound (resp. the lower bound) of large deviation hold true with* $\sup_{x \in E}$ *(resp.* $\inf_{x \in E}$*) replaced by* $\sup_{x \in K}$ *(resp.* $\inf_{x \in K}$*) for each* $K \subset\subset E$.
(b) (hyper-exponential recurrence) For any $r > 0$*, there is some compact* $K \subset\subset E$ *such that for any compact* $K' \subset\subset E$*,*

$$\sup_{x \in K'} \mathbb{E}^x \left(\frac{1}{r} \right)^{\tau_K} < +\infty.$$

(c) $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 0$*;*
(d) (Donsker-Varadhan's condition) $\inf_{K\subset\subset E}\inf_{u\in\mathcal{A}(P)}\sup_{x\notin K}\frac{Pu(x)}{u(x)} = 0.$

If E is locally compact, then they are equivalent to

(e) (L_n) *satisfies the locally uniform LDP w.r.t. the weak convergence topology (defined similarly as in (a)).*

Proof. We have proved in [47] the equivalence between (a) and (b) (and (e) in the locally compact case). By Corollary 3.13, parts (b) and (c) here are equivalent. The equivalence between (c) and (d) follows from Theorem 3.11(b). □

Remarks (6.iii). Under the assumption that E is locally compact, $P(x,dy) = p(x,y)\mu(dy)$ such that $p(x,\cdot) > 0$, $\mu - a.s.$ and $x \to p(x,\cdot)$ is continuous from E to $L^1(E,\mu)$ where μ is the invariant probability measure of P (which are stronger than (A2)+(A3)), Donsker and Varadhan proved in their pioneering work [16] that condition (d) is enough to the locally uniform LDP w.r.t. the weak convergence topology in part (e). It is a happy thing that their sufficient condition found 26 years ago is indeed necessary, showing the deepness of their work.

7. Applications to estimates of geometric convergence rate

In this section two applications are given: 1) estimate of eigenvalues of a symmetric Markov kernel on a countable space by means of the degrees of geometric recurrence; 2) geometric convergence in the metric of Wasserstein implies that in $b_U\mathcal{B}$ and in $L^p(\mu)$.

7.1. Estimate of eigenvalues of a symmetric Markov kernel by means of the degree of geometric recurrence

In this paragraph we suppose that E is at most countable, and P is an irreducible Markov kernel on E such that its invariant measure μ is a probability and charges all points of E, and P is symmetric on $L^2(\mu)$. In that case (A2) and (A3) are satisfied. The spectrum of $P|_{L^2(\mu)}$ is contained in $[-1,1]$. Assume that $r_{ess}(P|_{L^2(\mu)}) < 1$.

Let $(\lambda_j^+)_{j=0,1,\cdots,m^+}$ ($m^+ \in \mathbb{N}\cup\{\infty\}$) be the sequence in the non-increasing order of the eigenvalues counted up to multiplicity of $P|_{L^2(\mu)}$ above

$$\lambda_{ess}^+(P) := \sup\{\lambda;\ \lambda\in\sigma_{ess}(P|_{L^2(\mu)})\}\vee 0;$$

and $(\lambda^-)_{j=0,1,\cdots,m^-}$ ($m^- \in \mathbb{N}\cup\{\infty\}$) be the sequence in the non-decreasing order of the eigenvalues counted up to multiplicity of $P|_{L^2(\mu)}$ below

$$\lambda_{ess}^-(P) := \inf\{\lambda;\ \lambda\in\sigma_{ess}(P|_{L^2(\mu)})\}\wedge 0.$$

We adopt the following convention: if the sequence $(\lambda_j^+)_{0\le j\le m^+}$ (resp. $(\lambda_j^-)_{0\le j\le m^-}$) is finite, then we set $\lambda_j^+ := \lambda_{ess}^+(P|_{L^2(\mu)})$ (resp. $\lambda_j^- := \lambda_{ess}^-(P|_{L^2(\mu)})$) for $j > m^+$ (resp. $j > m^-$).

Using the famous min-max principle, we can control $\lambda_m^\pm$ by means of the degrees of geometric recurrence in the following way:

Proposition 7.1. *Let E, P, μ, $(\lambda_m^{\pm})$ be as above such that $r_{ess}(P|_{L^2(\mu)}) < 1$. For each $1 \le m \le \#(E)$(the cardinal of E), define the $m-$th degree of geometric recurrence by*

$$d_m := \inf_{K\subset E:\ \#(K)=m} \sup_{x\in K^c} \limsup_{n\to\infty} (\mathbb{P}_x(\sigma_K > n))^{1/n} \tag{7.1}$$

Then with the convention above, $|\lambda_m^{\pm}| \le d_m$ for all $m \in [1, \#(E)] \bigcap \mathbb{N}$.

Proof. We prove it only for E infinite. By the min-max principle ([38], Theorem XIII.1), for each $m \ge 1$,

$$\mu_m^+ := \inf_{f_1,\cdots,f_m\in L^2(\mu)} \sup_{f:\ \|f\|_2=1, f\perp\{f_1,\cdots,f_m\}} \langle f, Pf\rangle_\mu$$

is either the m-th eigenvalue (counted with multiplicity) of P above $\sup\{\lambda;\ \lambda \in \sigma_{ess}(P|_{L^2(\mu)})\}$, or coincides with the last quantity. Then by our convention about the definition of λ_j^+, for any finite subset $K = \{x_k|\ 1 \le k \le m\}$ of m-elements, setting $f_k = 1_{x_k}$, $k = 1, \cdots, m$, we have

$$\lambda_m^+ \le \sup_{f:\ \|f\|_2=1, f|_K=0} \langle f, Pf\rangle_\mu \le r_{sp}\left(1_{K^c} P 1_{K^c}|_{L^2(\mu)}\right).$$

But by (5.6) in Theorem 5.5,

$$r_{sp}\left(1_{K^c} P 1_{K^c}|_{L^2(\mu)}\right) = \sup_{x\in K^c} \limsup_{n\to\infty} (\mathbb{P}_x(\sigma_K > n))^{1/n}.$$

Since the subset K of m-elements is arbitrary, we obtain $\lambda_m^+ \le d_m$.

Similarly

$$\mu_m^- := \inf_{f_1,\cdots,f_m\in L^2(\mu)} \sup_{f:\ \|f\|_2=1, f\perp\{f_1,\cdots,f_m\}} \langle f, -Pf\rangle_\mu$$

is either the m-th eigenvalue (counted with multiplicity) of $-P$ above $\sup\{\lambda;\ \lambda \in \sigma_{ess}(-P|_{L^2(\mu)})\}$, or coincides with the last quantity. With the same argument as above, for each subset K of m-elements,

$$-\lambda_m^- = \mu_m^- \le \sup_{f:\ \|f\|_2=1, f|_K=0} \langle f, -Pf\rangle_\mu \le r_{sp}\left(1_{K^c} P 1_{K^c}|_{L^2(\mu)}\right)$$

and we can conclude in the same way. □

Remarks (7.i). An omitted point in the proposition above is the control of λ_0^-: it can not, in fact, be estimated by means of the degrees of geometric recurrence. For example let $P(1, 2) = P(2, 1) = 1$ be a Markov kernel on $E = \{1, 2\}$. The two eigenvalues of P are 1 and -1, and $r_{ess}(P) = 0$ (its essential spectrum is empty). Then $\lambda_0^- = -1$ but $d_m = 0$ for all $m \ge 1$.

To remedy this defective point, one can apply the above result to P^2 which is nonnegative definite.

7.2. Exponential convergence rate in $b_u\mathcal{B}$ and $L^p(\mu)$ by means of that in the metric of Wasserstein

Recall at first the L^p-Wasserstein distance between two probability measures μ and ν on the Polish space E with a compatible metric ρ:

$$W_p^\rho(\mu,\nu) = \inf_{(X,Y)} \|\rho(X,Y)\|_p \ (p \in [1,+\infty])$$

where the infimum is taken over all couples of E-valued random variables (X, Y) such that the law of X, Y are respectively μ and ν (the law of the couple (X, Y) on $E \times E$ is called a *coupling* of μ and ν).

Coupling technique was used by Nummelin and Tuominen [35], [36], Lindvall [28] (see also [31], Chap.15 and 16) etc to prove the geometric ergodicity by means of the Kendall renewal theorem. By using both coupling method and the Wasserstein distance (which depends on an ingenious construction of the metric ρ), Chen [1], [2], Chen and Wang [4], [5] etc obtained sharp estimates of the spectral gap in L^2 for the continuous time symmetric Markov semigroups. The following result, largely inspired by the last works, extends e.g. Theorems 6.1 and 6.2 in [1] or Theorem 1.7 in [4].

Proposition 7.2. *Let E be moreover locally compact and ρ a metric compatible with the topology of E, and P a Markov kernel on E satisfying (A2). Assume that there are constants $p \in [1,+\infty)$, $R \geq 1$, $r_0 \in (0,1)$ such that for all $x, y \in E$,*

$$W_p^\rho(P^n(x,\cdot), P^n(y,\cdot)) \leq \rho(x,y) R r_0^n, \ \forall n \geq 1. \tag{7.2}$$

If $u(x) := 1 + \rho(x, x_0)^p$ ($x_0 \in E$ is fixed) is $P^n(x_0,\cdot)$-integrable for all n, then $u \in \mathcal{A}(P)$ and

$$r_{ess}(P|_{b_u\mathcal{B}}) \leq (r_0)^p, \tag{7.3}$$

and there is a unique invariant probability measure μ such that $\mu(u) < +\infty$. Moreover

$$r_{exp}(P|_{b_u\mathcal{B}}) := \lim_{n\to\infty} \left(\|P^n - \mu(\cdot)\|_{b_u\mathcal{B}}\right)^{1/n} \leq r_0, \tag{7.4}$$

$$r_{exp}(P|_{L^p(\mu)}) := \lim_{n\to\infty} \left(\|P^n - \mu(\cdot)\|_{L^p(\mu)}\right)^{1/n} \leq \max\{r_0, r_{ess}(P|_{L^p(\mu)})\}. \tag{7.5}$$

Remarks (7.ii). In the context of this proposition, if P is moreover symmetric w.r.t. μ, then

$$r_{exp}(P|_{L^2(\mu)}) \leq r_0$$

by Theorem 5.5, (7.5) and (7.3).

Remarks (7.iii). Indeed (7.5) holds under (7.2) and the existence of invariant probability measure μ such that $\rho(x, x_0) \in L^p(\mu)$ (uniqueness of μ is guaranteed by (7.2)), without (A2). This follows from the proof below.

An open question is to remove $r_{ess}(P|_{L^p(\mu)})$ in (7.5) in the non-symmetric case.

Remarks (7.iv). Compared with the geometric convergence (7.2) in the metric of Wasserstein, (7.4) implies that for any $r > r_0$, there is constant $C \geq 1$ such that

$$\|P^n f - \mu(f)\|_u \leq C\|f\|_u r^n, \ \forall n \geq 1$$

for all $f \in b_u\mathcal{B}$. Similar geometric convergence can be derived from (7.2) only for f Lipchitzian. This gain is largely due to the assumption (A2) as seen for the following example: let $P(x,\cdot) = \delta_{\theta(x)}$ on $E = \mathbb{R}^d$ where $\theta x = rx$ with $r \in (0,1)$. It satisfies (7.2) but does not verify (7.4) (indeed $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|b_u\mathcal{B}) = 1$ by Theorem 4.6).

Proof. We prove (7.3) by two steps.

1) Let $n \geq 1$ be arbitrary but fixed. For any $\varepsilon > 0$ and $x, y \in E$, there are two E-valued random variables $X_n(x)$, $X_n(y)$ of laws $P^n(x,\cdot)$, $P^n(y,\cdot)$ respectively, defined on a same probability space $(\Omega, \mathcal{F}, \mathbb{P})$), such that

$$\mathbb{E}\rho(X_n(x), X_n(y))^p \leq [\rho(x,y)R(r_0+\varepsilon)^n]^p.$$

Using $(a+b)^p \leq (1+\varepsilon)^{p-1}a^p + ((1+\varepsilon)/\varepsilon)^{p-1}b^p$ $(\forall a, b \geq 0)$, we get from the estimation above (with $y = x_0$) that for $u(x) = 1 + \rho(x,x_0)^p$,

$$\begin{aligned} P^n u(x) &\leq 1 + \left(\frac{1+\varepsilon}{\varepsilon}\right)^{p-1} \mathbb{E}\rho(X_n(x_0), x_0)^p + (1+\varepsilon)^{p-1}\mathbb{E}\rho(X_n(x), X_n(x_0))^p \\ &\leq \left(\frac{1+\varepsilon}{\varepsilon}\right)^{p-1} \int_E u(x)P^n(x_0,dx) + (1+\varepsilon)^{p-1}[\rho(x,x_0)R(r_0+\varepsilon)^n]^p. \end{aligned}$$

The first term at the last line above is a finite constant. Hence $u \in \mathcal{A}(P^n)$. Moreover letting $K_L := \{x;\ \rho(x,x_0) \leq L\}$ which is compact, we have for L sufficiently large,

$$1_{K_L^c}(x)\frac{P^n u(x)}{u(x)} \leq \varepsilon + (1+\varepsilon)^{p-1}R^p(r_0+\varepsilon)^{np}.$$

Since $|a^p - b^p| \leq p|a-b|(a^{p-1} + b^{p-1})$ $(a, b \geq 0)$, we have

$$\begin{aligned} &|P^n u(x) - P^n u(y)| \\ &\leq p\mathbb{E}\rho(X_n(x), X_n(y))\left(\rho(X_n(x),x_0)^{p-1} + \rho(X_n(y),x_0)^{p-1}\right) \\ &\leq p\|\rho(X_n(x), X_n(y))\|_p \cdot \left(\|\rho(X_n(x),x_0)\|_p^{p-1} + \|\rho(X_n(y),x_0)\|_p^{p-1}\right) \end{aligned}$$

and $\|\rho(X_n(x),x_0)\|_p$ is locally bounded (a consequence of $u \in \mathcal{A}(P)$), $P^n u$ is continuous (for all $n \geq 1$). Hence by Remarks (3.vi), $P^u(x,dy) := [u(y)/u(x)]P(x,dy)$ satisfies (A2). We can thus apply Theorem 3.11(a) to conclude that

$$r_{ess}(P|_{b_u\mathcal{B}}) = \left(r_{ess}(P^n|_{b_u\mathcal{B}})\right)^{1/n} \leq \left(\varepsilon + (1+\varepsilon)^{p-1}R^p(r_0+\varepsilon)^{np}\right)^{1/n}.$$

Letting at first $\varepsilon \to 0+$ and next $n \to \infty$, we obtain (7.3).

This implies that P has an invariant probability measure μ such that $\mu(u) < +\infty$, by Proposition 4.2(b) (applied to P^u) or Theorem 4.6. By condition (7.2), invariant probability measure of P is unique.

2) Note that $r_{exp}(P|_{b_u\mathcal{B}}) \geq r_{ess}(P|_{b_u\mathcal{B}})$. If $r_{exp}(P|_{b_u\mathcal{B}}) = r_{ess}(P|_{b_u\mathcal{B}})$, then (7.4) follows by (7.3). Assume then $r_{exp}(P|_{b_u\mathcal{B}}) > r_{ess}(P|_{b_u\mathcal{B}})$.

Fix $r \in (r_{ess}(P|_{b_u\mathcal{B}}), r_{exp}(P|_{b_u\mathcal{B}}))$ such that $\Gamma_r := \{\lambda \in \mathbb{C} : \ |\lambda| = r\}$ does not contain spectral point of $P|_{b_u\mathcal{B}}$. Consider the sum Π_r of the eigen-projections E_{λ_j} where $\{\lambda_j; j\} = \{\lambda \in \sigma(P|_{b_u\mathcal{B}}); \ |\lambda| > r\}$.

By $r_{exp}(P|_{b_u\mathcal{B}}) > r > r_{ess}(P|_{b_u\mathcal{B}})$, the dimension of the range of Π_r, denoted by $m + 1$, is finite and not less than 2, i.e., $m \in [1, +\infty)$. Now by (2.3) and Proposition 4.2 (applied to $P^u f := (1/u)P(uf)$), there are

- a basis $\{f_k\}_{0\leq k\leq m}$ of the range of Π_r where $f_0 = 1$ and,
- a basis $(\nu_k)_{0\leq k\leq m}$ of the range of Π_r^* composed of complex valued measures with $\int_E u d\nu_k$ well defined, where $\nu_0 = \mu$,

such that $\langle \nu_k, f_l\rangle = \delta_{kl}$ (i.e., 1 if $k = l$ and 0 otherwise) for all $0 \leq k, l \leq m$ and

$$\Pi_r f = \sum_{k=0}^{m} \langle \nu_k, f\rangle f_k = \mu(f) + \sum_{k=1}^{m} \langle \nu_k, f\rangle f_k, \ \forall f \in b_u\mathcal{B}.$$

A crucial consequence of the fact above is: there exists some $h : E \to \mathbb{R}$ ρ-Lipchitzian such that $\Pi_r h - \mu(h) \neq 0$ (otherwise for each $k = 1, \cdots, m$, $\langle \nu_k, f\rangle = 0$ for any ρ-Lipchitzian function f, which would imply $\nu_k = 0$, a contradiction with the fact that "$m \geq 1$").

Now we can conclude easily: on one hand we have by (2.3)

$$\lim_{n\to\infty} \frac{\|P^n(\Pi_r h - \mu(h))\|_u}{r^n} = +\infty,$$

and on the other hand, by condition (7.2), for all $n \geq 1$, $x, y \in E$,

$$|P^n h(x) - P^n h(y)| \leq \|h\|_{Lip} \cdot W_1^{\rho}(P^n(x, \cdot), P^n(y, \cdot)) \leq \|h\|_{Lip} \rho(x, y) R r_0^n,$$

(where $\|h\|_{Lip} := \sup_{x\neq y} |h(x) - h(y)|/\rho(x, y)$) which implies that

$$\limsup_{n\to\infty} \frac{\|P^n(\Pi_r h - \mu(h))\|_u}{r_0^n} \leq \|\Pi_r\|_u \cdot \limsup_{n\to\infty} \sup_{x\in E} \frac{|P^n h(x) - \int_E P^n h(y) d\mu(y)|}{[1 + \rho(x, x_0)^p] r_0^n} < +\infty$$

where $\|\Pi_r\|_u$ is the operator norm in $b_u\mathcal{B}$. Combining those two facts we have $\lim_{n\to\infty} r_0^n/r^n = +\infty$, i.e., $r < r_0$. Since there is a sequence of such ($r = r_m$) increasing to $r_{exp}(P|_{b_u\mathcal{B}})$, we get hence $r_{exp}(P|_{b_u\mathcal{B}}) \leq r_0$.

The proof of (7.5) is similar to Step 2 above (**Note:** we have no longer control (7.3) about $r_{ess}(P|_{L^p(\mu)})$, that explains why $r_{ess}(P|_{L^p(\mu)})$ appears in the r.h.s. of

(7.5)). Indeed we may assume that $r_{exp}(P|_{L^p(\mu)}) > r_{ess}(P|_{L^p(\mu)})$ (trivial otherwise). For r strictly between $r_{ess}(P|_{L^p(\mu)})$ and $r_{exp}((P|_{L^p(\mu)})$ such that Γ_r does not contain spectral point of $P|_{L^p(\mu)}$, define Π_r in the same way as above. Now

$$\Pi_r f = \mu(f) + \sum_{k=1}^{m} \langle g_k, f\rangle_\mu f_k$$

where $(f_k)_{1\le k\le m}$ (resp. $(g_k)_{1\le k\le m}$) is a basis of the range of $\Pi_r - \mu$ (resp. of $(\Pi_r)^* - \mu$ acting on the dual $L^{p'}(\mu)$ of $L^p(\mu)$), and $m \ge 1$. Now for the argument in Step 2 above works, it suffices to notice that every ρ-Lipchitzian function h belongs to $L^p(\mu)$ (for $\rho(x, x_0)^p \in L^1(\mu(dx))$). □

Coupling method produces the geometric convergence of type (7.2), as shown by the following simple observation which should be well known.

Proposition 7.3. *Let P be a Markov kernel on E with a compatible metric ρ, and $p \in [1, +\infty)$. Assume that there is a coupling of P, i.e., a Markov kernel Q on $E \times E$ verifying: for all $A, B \in \mathcal{B}$ and $(x, y) \in E \times E$,*

(i) $Q((x, y), A \times E) = P(x, A)$ *and* $Q((x, y), E \times A) = P(y, A)$
(ii) $Q((x, x), A \times B) = P(x, A \bigcap B)$,

such that for some $r_0 \in (0, 1)$,

$$Q\rho^p(x, y) \le (r_0)^p \rho(x, y)^p, \ \forall (x, y) \in E \times E, \ x \ne y. \tag{7.6}$$

Then (7.2) is satisfied with the same r_0 and $R = 1$.

Proof. Let $(Z_n(x, y) = (X_n(x), Y_n(y)))_{n\ge 0}$ be the Markov chain defined on $(\Omega = (E \times E)^{\mathbb{N}}, \mathbb{P})$ with initial point (x, y) and transition kernel Q (that determines the law $\mathbb{P}$). Then $(X_n(x)), (Y_n(y))$ are respectively a Markov chain with the same transition kernel P, but with initial point x, y respectively (by condition (i) on Q).

Note that (7.6) is indeed valid for $x = y$ too, by condition (ii) on Q. Thus

$$\mathbb{E}\rho(X_n(x), Y_n(y))^p = Q^n \rho^p(x, y) \le (r_0)^{np} \rho(x, y)^p$$

where (7.2) follows. □

Remarks (7.v). The geometric convergence of type (7.2) can be naturally derived from Lyapunov exposant in the point of view of random dynamical systems, too. See examples in the next section.

8. Applications: several concrete models

8.1. Forward recurrence time chain

This model (see [31], Chap.2) is given as: $E = \mathbb{N}^*$ and the Markov transition kernel

$$P(1, j) = p_j, \ j \ge 1 \text{(regeneration distribution)}, \ P(k, k-1) = 1, \ \forall k \ge 2.$$

It is known ([31], Chap.16) that P is geometrically ergodic iff $\sum_j r^{-j} p_j < +\infty$ for some $r < 1$. Assume then

$$r_0 := \inf\{r > 0;\ \sum_j r^{-j} p_j < +\infty\} = \limsup_{j\to\infty} (p_j)^{1/j} < 1. \tag{8.1}$$

Let us show that

$$r_{ess}(P|_{b_{u_r}\mathcal{B}}) \le r, \forall r > r_0$$

where $u_r(x) := (1/r)^x$. It follows simply from

$$Pu_r(x) = 1_{x=1} \sum_{j\ge 1} r^{-j} p_j + 1_{x>1} r u_r(x)$$

and Theorem 3.11(a).

On the other hand, notice that for each $N \ge 2$,

$$\mathbb{E}^{x=1}\left(\frac{1}{r}\right)^{\tau_{[1,N]}} = \frac{1}{r}\sum_{j=1}^{N} p_j + \sum_{j>N} p_j r^{-(j-N)}$$

which is infinite if $0 \le r < r_0$. Hence by Corollary 3.13, we have

$$r_0 \le \inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \le r_{ess}(P|_{b_{u_r}\mathcal{B}}) \le r,\ \forall r > r_0. \tag{8.2}$$

Thus for this model, $r_0 = \inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}})$. Assume moreover that the number of $j :\ p_j > 0$ is infinite, the locally uniform LDP holds for (L_n) iff $r_0 = 0$ (by Theorem 6.4).

Remarks (8.i). This simple example serves as a counter-example to remarks already done in §3. Indeed, without any condition on $(P(1, j) = p_j)$, we have

1. For $u_r(x) := (1/r)^x$ where $r > 0$ is arbitrary, we see that $P(1_{[2,+\infty)}u_r)(x) = r u_r(x)$ for all $x \ge 2$. One might apply Theorem 3.11(a.ii) to derive $r_{ess}(P|_{b_{u_r}\mathcal{B}}) = 0$, a false fact once if $r_0 > 0$. Where is the question? The answer resides in the condition that $u_r \in \mathcal{A}(P)$, required for applying Theorem 3.11(a.ii).
2. For any $N \ge 1$ and $x \in \mathbb{N}^*$, we have $\mathbb{P}_x(\sigma_{[1,N]} > n) = 0$ for all n large enough. Hence the r.h.s. of (3.22) in Lemma 3.12 with σ_K in place of τ_K equals to zero. In other words, equality (3.22) in Lemma 3.12 with σ_K in place of τ_K is in general false.

8.2. (Reflected) Random Walk on $\mathbb{R}^+$ or $\mathbb{N}$.

In this model, $E = \mathbb{R}^+$ or $\mathbb{N}$, and our Markov process, given $X_0 = x \in E$, is defined recursively by

$$X_{n+1}(x) = (X_n(x) + W_n)^+$$

where $(W_n)_{n\ge 1}$ is a sequence of i.i.d.r.v. valued in $\mathbb{R}$ or $\mathbb{Z}$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Write always $W = W_1$. Assume that

$$\mathbb{E}W < 0 \text{ and the distribution of } W \text{ is absolutely continuous if } E = \mathbb{R}^+. \quad (8.3)$$

It entails that its transition kernel $P(x, dy)$ is strongly Feller (indeed $x \to P(x, \cdot)$ is continuous from E to $(M_1(E), \|\cdot\|_{var})$). Under that condition and the irreducibility, the chain is positive recurrent (this positive recurrence implies also $\mathbb{E}W < 0$).

Consider the log-Laplace transform $\Lambda_W(\lambda) := \log \mathbb{E}e^{\lambda W}$ valued in $(-\infty, +\infty]$.

Proposition 8.1. *For the reflected random walk described above, assume (8.3).*

(a) we have

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = \exp(-\Lambda_W^*(0)) = \inf_{a\ge 0} \mathbb{E}e^{aW}. \quad (8.4)$$

where $\Lambda_W^*(z) := \sup\{\lambda z - \Lambda_W(\lambda);\ \lambda \in \mathbb{R}\}$ *is the Fenchel-Legendre transformation of* Λ_W. *In particular,*

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) < 1 \iff \exists a > 0 :\ \mathbb{E}e^{aW} < +\infty;$$
$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 0 \iff \mathbb{P}(W > 0) = 0. \quad (8.5)$$

(b) If $a > 0$ *is such that* $\mathbb{E}e^{aW} < 1$, *then for* $u_a(x) := e^{ax}$,

$$r_{exp}(P|_{b_{u_a}\mathcal{B}}) \le \mathbb{E}e^{aW} \quad (8.6)$$

and the equality holds for $a = a_0$ *if* $\mathbb{E}e^{a_0 W} = \inf_{a\ge 0} \mathbb{E}e^{aW} < 1$.

Proof. We shall prove it only in the case where $E = \mathbb{R}^+$.

(a) Its proof is divided into three points.

1) We first prove (8.5) by assuming that (8.4) is true.

Since $\Lambda_W(\lambda)$ is convex on $\mathbb{R}$, we always have

$$\Lambda_W(\lambda) \ge \lambda \mathbb{E}W.$$

Since $\mathbb{E}W < 0$, that implies $\Lambda_W^*(0) = -\inf_{\lambda\ge 0} \Lambda_W(\lambda)$ and $\Lambda^*(\mathbb{E}W) = 0$. Consequently $\Lambda_W^*(0) > 0$ iff $\exists \lambda > 0 :\ \mathbb{E}e^{\lambda W} < 1$ iff $\exists a > 0 :\ \mathbb{E}e^{aW} < \infty$ (by "$\mathbb{E}W < 0$"), the first claim in (8.5).

Let us show $\Lambda_W^*(0) = +\infty$ iff $\mathbb{P}(W > 0) = 0$. Indeed $\Lambda_W^*(0) = +\infty$ iff $\inf_{a\ge 0} \mathbb{E}e^{aW} = 0$. If $P(W > 0) = 0 = \mathbb{P}(W \ge 0)$ (by the absolute continuity), then $0 \le \inf_{a\ge 0} \mathbb{E}e^{aW} \le \lim_{a\to+\infty} \mathbb{E}e^{aW} = 0$.

Inversely note that if $\inf_{a\ge 0} \mathbb{E}e^{aW} = 0$, that infimum can not be attained at any finite $a \in \mathbb{R}^+$. Then by the convexity of $a \to \mathbb{E}e^{aW}$, $0 = \inf_{a\ge 0} \mathbb{E}e^{aW} = \lim_{a\to+\infty} \mathbb{E}e^{aW}$ which implies $e^{aW} \to 0$ in probability as $a \to +\infty$. Thus $\mathbb{P}(W \ge 0) = 0$.

2) Let us show

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \le \exp(-\Lambda_W^*(0)).$$

The inequality above holds trivially when $\Lambda_W^*(0) = 0$ or equivalently $A = 0$, where $A := \sup\{a \geq 0; \mathbb{E}e^{aW} < +\infty\}$ (by Step 1). Assume then $A > 0$. In that case, for each $0 < a < A$, $u_a(x) := e^{ax}$ verifies

$$P(1_{(0,+\infty)}u_a)(x) = \mathbb{E}e^{a(x+W)}1_{[x+W>0]} \leq u_a(x)\mathbb{E}e^{aW},$$
$$P(u_a)(x) = \mathbb{E}e^{a(x+W)^+} \leq u_a(x)\left(\mathbb{E}e^{aW} + 1\right).$$

The second control above implies $u_a \in \mathcal{A}(P)$ as well as $u_a^{1+\varepsilon} \in \mathcal{A}(P)$ (condition (3.14)). The first inequality above is the checked drift condition in Theorem 3.11(a.ii). We obtain consequently by Theorem 3.11(a),

$$r_{ess}(P_{b_{u_a}\mathcal{B}}) \leq \mathbb{E}e^{aW}. \tag{8.7}$$

Since $a \to \mathbb{E}e^{aW}$ is convex on $[0, A]$, $\inf_{0<a<A} \mathbb{E}e^{aW} = \inf_{0\leq a\leq A} \mathbb{E}e^{aW} = \exp(-\Lambda_W^*(0))$, we obtain the desired inequality.

3) We show now the inverse inequality $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \geq \exp(-\Lambda_W^*(0))$. Assume without loss of generality that $\Lambda_W^*(0) < +\infty$ or equivalently $\mathbb{P}(W > 0) > 0$. For any $N \geq 0$ and each $x > N$ fixed, letting $S_n := W_1 + \cdots + W_n$, we have

$$\begin{aligned}\mathbb{P}_x(\tau_{[0,N]} > n) &= \mathbb{P}(x + S_k > N;\ \forall k = 1, \cdots, n)\\ &\geq \mathbb{P}\left(S_k > 0,\ \forall k = 1, \cdots, n\right)\end{aligned}$$

for all $n \geq 1$. By Lemma 8.2 below, and noting that Λ_W^* is non-decreasing on $[\mathbb{E}W, +\infty)$, we obtain for all $x > N$,

$$\limsup_{n\to\infty} \left[\mathbb{P}_x(\tau_{[0,N]} > n)\right]^{1/n} \geq e^{-\Lambda_W^*(0)}.$$

Now by Corollary 3.13, we get

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \geq \exp(-\Lambda_W^*(0))$$

the desired inverse inequality.

(b) Construct a coupling Markov kernel $Q((x, y); \cdot)$ of P by

$$Q((x, y) \in A \times B) := \mathbb{P}((X_1(x), X_1(y)) \in A \times B),\ \forall A, B \in \mathcal{B}$$

where $X_1(x)$ (recalling it) is $x + W$ for all $x \in E$. Put $\rho(x, y) = |e^{ay} - e^{ax}|$ which defines a compatible metric on $E = \mathbb{R}^+$, where $a > 0$ is such that $\mathbb{E}e^{aW} < 1$. Noting that

$$0 \leq e^{a(y+w)^+} - e^{a(x+w)^+} \leq (e^{ay} - e^{ax})e^{aw},\ \forall w \in \mathbb{R}, x < y$$

(easy if one divides its verification into three cases: $w \leq -y$ or $w \geq -x$ or $w \in (-y, -x)$), we have for $0 \leq x < y$,

$$(Q\rho)(x, y) = \mathbb{E}e^{a(y+W)^+} - \mathbb{E}e^{a(x+W)^+} \leq \rho(x, y)\mathbb{E}e^{aW}$$

and same for $x > y \geq 0$. In other words condition (7.6) in Proposition 7.3 is verified. Hence (8.6) follows by Propositions 7.3 and 7.2. □

The following lemma, used in the proof of part (a) above, reinforces the classical Cramer theorem.

Lemma 8.2. *Let* $(W_n)_{n\geq 0}$ *be a sequence of real i.i.d.r.v. defined on* $(\Omega, \mathcal{F}, \mathbb{P})$ *with* $\mathbb{P}(W > 0) > 0$*, and* $S_n := \sum_{k=1}^n W_k$*. Then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\,(S_k > 0, \forall k = 1, \cdots, n) = -\inf_{x\geq 0} \Lambda_W^*(x). \tag{8.8}$$

Proof. **1)** The "$\leq$" is known. Indeed, by the Cramer theorem without moment condition in [11], we have

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\,(S_k \geq 0, \forall k = 1, \cdots, n)\\ &\leq \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \geq 0\right) \leq -\inf_{x\geq 0} \Lambda_W^*(x).\end{aligned}$$

2) We prove now

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\,(S_k > 0, \forall k = 1, \cdots, n) \geq -\inf_{x>0} \Lambda_W^*(x).$$

In Sanov theorem's approach of Cramer's theorem developed in [45], we have proved that for any open $G \subset \mathbb{R}$,

$$\inf_{x\in G} \Lambda_W^*(x) = \inf\left\{h(\nu, \mu);\ \int_{\mathbb{R}} |x| d\nu(x) < +\infty,\ \ \int_{\mathbb{R}} x d\nu(x) \in G\right\} \tag{8.9}$$

where $\mu = \mathcal{L}(W)$ (the law of $W = W_1$), and for any probability measure ν,

$$h(\nu; \mu) := \begin{cases} \int_{\mathbb{R}} \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu, & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise} \end{cases}$$

is the relative entropy. Thus it is sufficient to show that

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\,(S_k > 0, \forall k = 1, \cdots, n) \geq -h(\nu; \mu) \tag{8.10}$$

for any probability measure ν such that $\int_{\mathbb{R}} |x| d\nu(x) < +\infty$ and $\int_{\mathbb{R}} x d\nu(x) > 0$ and $h(\nu; \mu) < +\infty$. The following method is standard in the theory of large deviation for treating lower bound.

Without loss of generality assume that $(\Omega = \mathbb{R}^{\mathbb{N}^*}, \mathbb{P} = \mu^{\otimes \mathbb{N}^*})$ and $(W_n(\omega) = \omega_n)$ is the system of coordinates. Let $\mathbb{Q} = \nu^{\mathbb{N}^*}$. On $\mathcal{F}_n = \sigma(W_k;\ 1 \leq k \leq n)$, we have

$$\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_n} = \exp\left(\sum_{k=1}^n \log \frac{d\nu}{d\mu}(W_k)\right).$$

Putting

$$A_n := [S_k > 0, \forall k = 1, \cdots, n],$$
$$B_{n,\varepsilon} := \left[\sum_{k=1}^{n} \log \frac{d\nu}{d\mu}(W_k) \le n(h(\nu;\mu) + \varepsilon)\right],$$

where $\varepsilon > 0$ is arbitrary (independent of n), we get thus

$$\mathbb{P}(A_n) \ge \int_\Omega \exp\left(-\sum_{k=1}^{n} \log \frac{d\nu}{d\mu}(W_k)\right) 1_{A_n} d\mathbb{Q} \ge e^{-n(h(\nu;\mu)+\varepsilon)} \mathbb{Q}(A_n \bigcap B_{n,\varepsilon}).$$

Consequently for (8.10), it remains to show that $\mathbb{Q}(B_{n,\varepsilon}) \to 1$ and $\lim_{n\to\infty} \mathbb{Q}(A_n) > 0$. The first is a consequence of the law of large number. The second is a well known fact in random walks (for $\mathbb{E}^{\mathbb{Q}} W > 0$), see e.g. [6], §5.4, Theorem 2.

3) It remains to establish $\inf_{x\ge 0} \Lambda_W^*(x) \ge \inf_{x>0} \Lambda_W^*(x)$. It is obviously true if $\Lambda_W^*(0) = +\infty$. Assume then $\Lambda_W^*(0) < +\infty$. Note that condition $P(W > 0) > 0$ implies that $\Lambda^*(a) < +\infty$ for some $a > 0$ by (8.9) (taking $\nu(dx) = 1_{(0,L)}(x)\mu(dx)/\mu((0,L))$ with $\mu((0,L)) > 0$). Then $\Lambda_W^* < +\infty$ on $[0,a]$ (by convexity), and consequently Λ_W^* is continuous on $[0,a]$ (by the convexity and lower semi-continuity of Λ_W^*). The proof is completed. □

Let us present three explicit examples:

Example 8.3. Let $E = \mathbb{N}$, $\mathbb{P}(W = 1) = p = 1 - \mathbb{P}(W = -1) =: 1 - q$ with $q > p$. By a simple calculus, $\inf_{a\ge 0} \mathbb{E}e^{aW} = 2\sqrt{pq}$, and it is attainted at $a_0 = \log\sqrt{q/p}$. Thus by Proposition 8.1 and (8.7),

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = r_{ess}(P|_{b_{u_0}\mathcal{B}}) = r_{exp}(P|_{b_{u_0}\mathcal{B}}) = 2\sqrt{pq}$$

where $u_0(x) = (\sqrt{q/p})^x$ for all $x \in E = \mathbb{N}$. This model is moreover symmetric w.r.t. its unique invariant probability measure μ given by $\mu(k) = (p/q)^k c$. Thus

$$r_{ess}(P|_{L^2(\mu)}) = r_{exp}(P|_{L^2(\mu)}) = 2\sqrt{pq},$$

by Remarks (7.ii) (for $r_{exp}(P|_{L^2(\mu)})$) and Theorem 5.5 (for $r_{ess}(P|_{L^2(\mu)})$).

For this example $\lim_{N\to\infty}(\mu([N, +\infty)))^{1/N} = p/q$, then when $p < 1/2$ is sufficiently close to $1/2$, the last quantity p/q is close to $2\sqrt{pq} = \inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}})$. This shows that the lower bound (3.28) of $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}})$ by means of the concentration of invariant measure μ in Proposition 3.14 (with $d(x,y) := |x - y|$ on $\mathbb{N}$) becomes almost sharp.

Example 8.4. Let $E = \mathbb{N}$, and $\mathbb{P}(W = 1) = p$, $\mathbb{P}(W = 0) = r$, $\mathbb{P}(W = -1) = q$ where $0 < p < q < 1$ and $p + r + q = 1$.

We have that $\inf_{a\ge 0} \mathbb{E}e^{aW} = 2\sqrt{pq} + r$, and it is attainted at $a_0 = \log\sqrt{q/p}$. Thus by Proposition 8.1 and (8.7),

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = r_{ess}(P|_{b_{u_0}\mathcal{B}}) = r_{exp}(P|_{b_{u_0}\mathcal{B}}) = 2\sqrt{pq} + r$$

where $u_0(x) = (\sqrt{q/p})^x$ for all $x \in E = \mathbb{N}$. P is reversible w.r.t. the same geometric law μ given in the precedent example. Thus for the same reason,

$$r_{ess}(P|_{L^2(\mu)}) = r_{exp}(P|_{L^2(\mu)}) = 2\sqrt{pq} + r.$$

Example 8.5. In the above model, let $E = \mathbb{R}^+$, and the law of W is $\mathcal{N}(m, \sigma^2)$ with $m < 0$, $\sigma^2 > 0$. Then $\inf_{a\geq 0} \mathbb{E}e^{aW} = \exp\left(-m^2/2\sigma^2\right)$, and it is attained at $a_0 = -m/\sigma^2$. Thus by Proposition 8.1 and (8.7),

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = r_{ess}(P|_{b_{u_0}\mathcal{B}}) = r_{exp}(P|_{b_{u_0}\mathcal{B}}) = \exp\left(-\frac{m^2}{2\sigma^2}\right)$$

where $u_0(x) = \exp\left(-mx/\sigma^2\right)$ for all $x \in E = \mathbb{R}^+$.

8.3. Linear iteration under random perturbation.

We now study a Markov chain with values in $\mathbb{R}^d$ (taken from [31], Chap.4), given by

$$X_0(x) = x,\ \ X_{n+1}(x) = AX_n(x) + BW_{n+1},\ \forall n \geq 0$$

where $(W_n)_{n\in\mathbb{Z}}$ is a sequence of i.i.d.r.v. valued in $\mathbb{R}^m$, A is a matrix $d \times d$ and B a matrix of $d \times m (d, m \geq 1)$. We assume that (taken from [31], Chap.4, Prop.4.4.3)

$$[A^{d-1}B|A^{d-2}B|\cdots|AB|B] \text{ is of (full) rank } d \tag{8.11}$$

$$\text{the law of } W \text{ is absolutely continuous on } \mathbb{R}^m. \tag{8.12}$$

Here $[A_1|A_2|\cdots|A_d]$ where A_i is of form $d \times n_i$, denotes the matrix $d \times (n_1 + \cdots + n_d)$ whose i-th line is the union in order of the i-th lines of $A_1, \cdots, A_d$. They entail (A2) for $N = d$. Indeed P is obviously Feller. Let us show that P^N is strongly Feller for $N = d$. Notice that

$$P^N(x, dy) = \text{law of } A^N x + A^{N-1}BW_1 + \cdots + ABW_{N-1} + BW_N.$$

Write $Y_N := \sum_{j=1}^N A^{N-j}BW_j$. The idea is to compare it with the Gaussian model well studied in [31], Chap.4, Prop.4.4.3..

Let $(\tilde{W}_j)$ be a sequence of i.i.d.r.v. valued in $\mathbb{R}^k$ with the standard Gaussian law $\mathcal{N}(0, I)$ and $\tilde{Y}_N := \sum_{j=1}^N A^{N-j}B\widetilde{W}_j$. It is easy to check that the variance matrix of $\tilde{Y}_N$ is

$$\sum_{j=1}^N A^{N-j}BB^t(A^{N-j})^t = [A^{N-1}B|\cdots|AB|B]\cdot[A^{N-1}B|\cdots|AB|B]^t$$

(A^t denotes the transposition of A). It is non-degenerate once if $N \geq d$ by (8.11). Hence the law of $\tilde{Y}_N$ is equivalent to the Lebesgue measure dy on $\mathbb{R}^d$.

In further by assumption (8.12), the law of Y_N is absolutely continuous w.r.t. the law of $\tilde{Y}_N$, then w.r.t. dy too.

Consequently for $N \geq d$, $P^N(x, dy) = h_{Y_N}(y - A^N x)dy$ where h_{Y_N} is the probability density of Y_N. But it is well known that $x \to f(\cdot - A^N x)$ is continuous from $\mathbb{R}^d$ to $L^1(dy)$ for any $f \in L^1(dy)$. Thus P^N is strongly Feller.

From the expression of $P^N(x, dy)$ above, we see that under (8.13) below, as $N \to \infty$, $P^N(x, \cdot)$ ($\forall x$) converges weakly to the law μ of

$$\sum_{j=0}^{\infty} A^j B W_{-j}.$$

(this series converges in law, then a.s.). In other words μ is the unique invariant measure of this Markov chain, and when μ is equivalent to dx, (A3) is verified.

To estimate $r_{ess}(P)$, we begin with a simple observation. If

$$\|A\|_2 := \sup\{|Ax|; |x| \leq 1\} < 1, \ \mathbb{E}|W| < +\infty,$$

where $|\cdot|$ denotes the Euclidean norm of $\mathbb{R}^d$, then for $u_1(x) := |x| + 1$, we see that

$$Pu_1(x) \leq \|A\|_2 u_1(x) + \|B\|_2 \mathbb{E}|W|$$

where follows $r_{ess}(P|_{b_{u_1}\mathcal{B}}) \leq \|A\|_2$. But this is too rough.

Proposition 8.6. *Assume (8.11), (8.12), and*

$$r_{sp}(A) := \sup\{|\lambda|;\ \lambda \text{ is an eigenvalue of } A\} < 1. \tag{8.13}$$

(a) Given $1 \leq p < +\infty$. *If* $\mathbb{E}|W|^p < +\infty$, *then for* $u_p(x) := 1 + |x|^p$,

$$r_{ess}(P|_{b_{u_p}\mathcal{B}}) \leq (r_{sp}(A))^p \ \text{ and } \ r_{exp}(P|_{b_{u_p}\mathcal{B}}) = r_{sp}(A). \tag{8.14}$$

(b) If $\mathbb{E}|W|^p < +\infty$ *for all* $p \in [1, +\infty)$, *then*

$$\inf_{u \in \mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 0.$$

In particular when the density function of $\sum_{j=0}^{\infty} A^j B W_{-j}$ *is moreover* $dx-$ *a.e.-positive on* $\mathbb{R}^k$, *then the empirical occupation measures* (L_n) *satisfies the local* τ*-LDP (stated in Theorem 6.4(a)).*

Proof. (a) Note that for any initial point $x, y \in \mathbb{R}^d$,

$$X_n(x) = A^N x + \sum_{j=1}^{N} A^{N-j} B W_j = A^N x + Y_N, \ X_n(x) - X_n(y) = A^N(x - y).$$

Then $|X_n(x) - X_n(y)| = |A^N(x - y)| \leq \|A^N\|_2 |x - y|$, which implies (7.2) w.r.t. the Euclidean metric $\rho(x, y) = |x - y|$. Since $\mathbb{E}|W|^p < +\infty$, we can apply Proposition 7.2 and get

$$r_{exp}(P|_{b_{u_p}\mathcal{B}}) \leq \limsup_{N \to \infty} (\|A^N\|_2)^{1/N} = r_{sp}(A)$$
$$r_{ess}(P|_{b_{u_p}\mathcal{B}}) \leq \limsup_{N \to \infty} (\|A^N\|_2)^{p/N} = (r_{sp}(A))^p.$$

Now for (8.14), it remains to prove that $r_{exp}(P|_{b_{u_p}\mathcal{B}}) \geq r_{sp}(A)$. This follows from

$$r_{exp}(P|_{b_{u_p}\mathcal{B}}) \geq \limsup_{n\to\infty} \sup_{|x|,|y|\leq 1} |\mathbb{E}X_n(x) - \mathbb{E}X_n(y)|^{1/n}$$
$$= \limsup_{n\to\infty} \sup_{|x|,|y|\leq 1} |A^n(x-y)|^{1/n} = r_{sp}(A).$$

(b) Letting $p \to \infty$ in the first inequality in (8.14), we get $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u}\mathcal{B}) = 0$. Finally the local τ-LDP follows by Theorem 6.4. □

8.4. Autoregressive Model.

It is given by

$$Y_n := a_1 Y_{n-1} + a_2 Y_{n-2} + \cdots + a_d Y_{n-d} + W_n$$

where $d \geq 1$, $a_j \in \mathbb{R}$ and $(W_n)_{n\in\mathbb{Z}}$ is a sequence of real i.i.d.r.v. Then $X_n := (Y_n, \cdots, Y_{n-d+1})^t$ is $\mathbb{R}^d$-valued Markov chain. Moreover it satisfies

$$X_{n+1} = AX_n + BW_n$$

where $A = (a_{ij})_{n\times n}$ with $a_{1j} = a_j$, $a_{j1} = 1$ for $j \geq 2$ and $a_{ij} = 0$ otherwise, and $B = (1, 0, \cdots, 0)^t$. In other words, it is a particular case of the previous model.

In Meyn and Tweedie [31], Prop.4.4.2, it is shown that A, B satisfies the full rank condition (8.11). And the eigenvalues of A are just a_1 and 0 (the algebraic multiplicity of 0 is $d-1$).

Thus *Proposition 8.6 applies for* (X_n) *once if the law of* $W = W_1$ *is absolutely continuous w.r.t.* dx *and*

$$r_{sp}(A) = |a_1| < 1,\ \mathbb{E}|W|^p < +\infty, \quad \text{for some or for all } \ p \geq 1.$$

8.5. A random non-linear dynamical system

Now we study the following non-linear model in $\mathbb{R}^d$ ($d \geq 1$):

$$X_0(x) = x \in \mathbb{R}^d,\ X_{n+1}(x) = F(X_n(x), W_{n+1}),\ n \geq 0, \tag{8.15}$$

where

(8.15a) The noise $(W_n)_{n\in\mathbb{Z}}$ is a sequence of $\mathbb{R}^m$ ($m \geq 1$)-valued i.i.d.r.v. defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that the law of $W = W_1$ is absolutely continuous w.r.t. the Lebesgue measure dw on $\mathbb{R}^m$;
(8.15b) $F(x, w) \in C^1(\mathbb{R}^d \times \mathbb{R}^k)$.

It is the so called nonlinear state space model in [31], Chap.7, where properties such as irreducibility and periodicity are characterized from the point of view of topological dynamical system.

Following [31], set $F_1(x, w_1) := F(x, w_1)$ and for $k \geq 2$,

$$F_k(x; w_1, \cdots, w_k) = F\left(F_{k-1}(x, w_1, \cdots, w_{k-1}), w_k\right) : \ \mathbb{R}^d \times (\mathbb{R}^m)^k \to \mathbb{R}^d.$$

The transition probability kernel of this Markov chain is given by

$$P^n f(x) = \mathbb{E} f(X_n(x)) = \mathbb{E} f\left(F_n(x; W_1, \cdots, W_n)\right).$$

It is obviously Feller.

Lemma 8.7. *If for some $N \geq 1$,*

$$\left(D_{w_1} F_N | \cdots | D_{w_{N-1}} F_N | D_{w_N} F_N\right) \text{ is of full rank } d \text{ over } \mathbb{R}^d \times (\mathbb{R}^m)^N \quad (8.16)$$

where $D_{w_k} F_N$ is the differential w.r.t. the variable w_k (represented as a matrix $d \times m$), then $P^N(x, dy) = p_N(x, y)dy$ and it is strongly Feller.

When $F(x, w) = Ax + Bw$, we re-find the model in the preceding paragraph and $\left(D_{w_1} F_N | \cdots | D_{w_N} F_N\right) = \left(A^{N-1}B | \cdots | AB | B\right)$. In other words (8.16) is exactly an extension of (8.11). But condition (8.16) here is stronger than [31], Chap.7, (CM3), p156 (because we want the strong Feller property which is stronger than "T-chain" checked there).

Proof. For each $x \in \mathbb{R}^d$, $X_N(x) = F_N(x; W_1, \cdots, W_N)$. As the law of $(W_1, \cdots, W_N)$, being the product measure of the law of W, is absolutely continuous w.r.t. the Lebesgue measure on $(\mathbb{R}^m)^N$, then the law of $X_N(x) = F_N(x; W_1, \cdots, W_N)$ is absolutely continuous w.r.t. the Lebesgue measure dy on $\mathbb{R}^d$ by the full rank assumption (8.16). Denote this density by $y \to p_N(x, y)$.

If $x_n \to x$, we also have $p_N(x_n, y)$ converges in measure dy to $p(x, y)$. Since $\int_{\mathbb{R}^d} p_N(x_n, y)dy = \int_{\mathbb{R}^d} p_N(x, y)dy = 1$, we have indeed $p_N(x_n, y) \to p_N(x, y)$ in $L^1(dy)$. Hence $x \to P^N(x, dy) = p_N(x, y)dy$ is continuous from $\mathbb{R}^d$ to $(M_b(E), \|\cdot\|_{var})$, which is stronger than the strong Feller property of P^N. □

We are now ready to prove

Proposition 8.8. *For the model (8.15), assume (8.16).*

(a) Assume that for some $p \in [1, +\infty)$,

$$\mathbb{E}\left(\sup_{x \in \mathbb{R}^d} |D_x F(x; W)|^p + |F(0, W)|^p\right) < +\infty, \quad (8.17)$$

where $|A| := \|A\|_2$ for a matrix A. If for some $k \geq 1$ and some $L > 0$,

$$\mathbb{E} \sup_{|x| \geq L} |D_x F_k(x; W_1, \cdots, W_k)|^p < 1, \quad (8.18)$$

then $u_p(x) := 1 + |x|^p \in \mathcal{A}(P)$ and

$$r_{ess}(P|_{b_{u_p}\mathcal{B}}) \leq \left(\mathbb{E} \sup_{|x| \geq L} |D_x F_k(x; W_1, \cdots, W_k)|^p\right)^{1/k} < 1. \quad (8.19)$$

(b) If $F(0, W) \in \bigcap_{1\le p<+\infty} L^p(\mathbb{P})$, $\sup_{x\in\mathbb{R}^d} |D_x F(x, W)| \in L^\infty(\Omega, \mathbb{P})$ *and for some* $L > 0, k \ge 1$,

$$\| \sup_{|x|\ge L} |D_x F_k(x; W_1, \cdots, W_k)\|_{L^\infty(\Omega,\mathbb{P})} < 1 \tag{8.20}$$

then

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 0. \tag{8.21}$$

Proof. (a) Write simply $\Delta_k(x) := D_x F_k(x, W_1, \cdots, W_k) = D_x X_k(x)$. It is easy to see that $u_p \in \mathcal{A}(p)$ under (8.17). Since $|X_1(x) - X_1(y)| \le |x - y| \sup_{z\in\mathbb{R}^d} |D_z F(z, W)|$, we see by (8.17) that when $|x - y| \to 0$,

$$\|X_1(x) - X_1(y)\|_p \le |x - y| \| \sup_{z\in\mathbb{R}^d} |D_z F(z, W)| \|_p \to 0.$$

Hence Pu_p is continuous. By Remarks (3.vi), P^{u_p} satisfies (A1) as P.
Noting that for $|x| \ge L$, $|X_k(x) - X_k(Lx/|x|)| \le \sup_{|z|\ge L} |\Delta_k(z)| \cdot |x|$, we have for any $\varepsilon > 0$ (using $(a+b)^p \le (1+\varepsilon)^{p-1}a^p + [(1+\varepsilon)/\varepsilon]^{p-1}b^p$ for $a, b \ge 0$),

$$\begin{aligned} P^k u_p(x) &= 1 + \mathbb{E}|X_k(x)|^p \le 1 + \mathbb{E}\left(\sup_{|z|\ge L} |\Delta_k(z)| \cdot |x| + \sup_{|y|\le L} |X_k(y)| \right)^p \\ &\le 1 + (1+\varepsilon)^{p-1}\mathbb{E}\left(\sup_{|z|\ge L} |\Delta_k(z)| \right)^p \cdot u_p(x) \\ &\quad + \left(\frac{1+\varepsilon}{\varepsilon}\right)^{p-1} \mathbb{E} \sup_{|y|\le L} |X_k(y)|^p. \end{aligned}$$

Hence by Theorem 3.11(a) applied to P^k,

$$r_{ess}(P^k|_{b_{u_p}\mathcal{B}}) \le (1+\varepsilon)^{p-1}\mathbb{E}\left(\sup_{|z|\ge L} |\Delta_k(z)| \right)^p.$$

Using the fact that $r_{ess}(P^k|_{b_{u_p}\mathcal{B}}) = r_{ess}(P|_{b_{u_p}\mathcal{B}})^k$ and since $\varepsilon > 0$ is arbitrary, we so obtain (8.19).

(b) By (8.19), we have for any $p \in [1, +\infty)$,

$$r_{ess}(P|_{b_{u_p}\mathcal{B}}) \le \left(\| \sup_{|x|\ge L} |D_x F_k(x; W_1, \cdots, W_k)| \|_{L^\infty(\Omega,\mathbb{P})} \right)^{p/k}$$

where (8.21) follows by assumption (8.20) and by letting $p \to +\infty$. □

Remarks (8.iii). In the linear model treated previously, $F(x, w) = Ax + BW$, we have $D_x F_k(x, W_1, \cdots, W_k) = A^k$ which does not depend on (W_j). Hence conditions (8.18) and (8.20) are identical to (8.13), and part (b) extends Proposition 8.6(b).

Remarks (8.iv). Let us compare this result with Meyn and Tweedie [31], Theorem 16.2.7. Assume that the density h_W of W, is lower semi-continuous. Let

$$A_+(x) = \{F_k(x, w_1, \cdots, w_k);\ k \ge 1,\ (w_1, \cdots, w_k) \in O_w^k\}$$

where $O_w := \{h_W > 0\}$. Suppose that there is a unique closed set $M \subset \mathbb{R}^d$ such that

$$\overline{A^+(x)} = M,\ \forall x \in M \tag{8.22}$$

(the so called M-irreducibility, which is equivalent to the μ- irreducibility of (X_n), see [31], Theorem 7.2.6). With all assumptions above and (CM3) in [31] mentioned before (instead of the stronger (8.16) here), Theorem 16.2.7 in [31] says that when M is compact (i.e., the Lagrange stability there), then $P|_M$ is Doeblin recurrent (under our strong condition (8.16), $(P|_M)^k$ becomes compact for k sufficiently large by Lemma 8.7 and Proposition 3.2(g)).

The main gains in our proposition are

- removedness of the irreducibility assumption;
- the compactness of M is substituted by (8.18) or (8.20);
- an explicit estimate of $r_{ess}(P)$;
- in the irreducible (8.22) and aperiodic case, our result says that the geometrical ergodicity holds for every initial state x in the whole space $\mathbb{R}^d$, not restricted to M. The last point is interesting because *a priori* M is unknown (but I believe that this is known to specialists).

Corollary 8.9. *Suppose that the conditions in Proposition 8.8(b) are satisfied, and the density function h_W of W is lower semi-continuous. Assume the M-irreducibility. Then restricted our process (X_n) to M, the empirical measures (L_n) satisfies the locally uniform LDP on $(M_1(M), \tau)$ with rate function $J^M(\nu) = J(\nu),\ \forall \nu \in M_1(M)$.*

Proof. Restricted to M, $P|_M$ satisfies (A3) too. Then the locally uniform LDP on $(M_1(M), \tau)$ follows by Theorem 6.4 and Proposition 8.8(b). □

We now present a direct corollary of Theorem 3.5 to the uniform ergodicity.

Corollary 8.10. *For the model (8.15), assume (8.16). If for some $k \ge 1$,*

$$\int_0^{+\infty} \Gamma_k(r)dr < +\infty,\ \mathbb{P} - a.s. \tag{8.23}$$

where $\Gamma_k(r) := \sup_{|x|=r} |F_k(x, W_1, \cdots, W_k)|$, then $r_{ess}(P|_{b\mathcal{B}}) = 0$. In particular if (X_n) is M-irreducible, then it is Doeblin recurrent, and restricted to M, (X_n) satisfies the uniform τ-LDP.

Proof. For any $L > 0$, if $|x| > L$, then

$$|X_k(x) - X_k(Lx/|x|)| \le \int_L^{|x|} \Gamma_k(r)dr$$

where it follows

$$\sup_{|x|>L} |X_k(x)| \le \sup_{|z|\le L} |X_k(z)| + \int_L^{+\infty} \Gamma_k(r)dr.$$

Consequently

$$\sup_{x\in\mathbb{R}^d} P^k(x; B(0,2N)^c) \le \mathbb{P}\left(\sup_{|z|\le L} |X_k(z)| \ge N\right) + \mathbb{P}\left(\int_L^{+\infty} \Gamma_k(r)dr \ge N\right) \longrightarrow 0$$

as N goes to infinity. In other words $\beta_w(P^k) = 0$. Then the desired result follows from Theorem 3.5.

Finally restricted to the irreducible set M, (A3) is satisfied. Thus the uniform τ-LDP follows from Theorem 6.3. □

We turn now to estimate the geometric convergence rate of this model. Using the notations of the proof of Proposition 8.8, we have

$$|X_n(x) - X_n(y)| \le \sup_{z\in\mathbb{R}^d} |\Delta_n|(z)|x-y|.$$

Thus $\|X_n(x) - X_n(y)\|_p \le \|\sup_{z\in\mathbb{R}^d} |\Delta_n|(z)\|_p \cdot |x-y|$. Consequently for and the usual Euclidean metric ρ on $E = \mathbb{R}^d$, we have

$$W_p\left(P^n(x,\cdot), P^n(y,\cdot)\right) \le \mathbb{E}\|\sup_{z\in\mathbb{R}^d} |\Delta_n|(z)\|_p \cdot |x-y|,\ \forall n\ge 1,\ x,y\in\mathbb{R}^d. \tag{8.24}$$

In further noting that

$$\begin{aligned}\Delta_{n+m}(x; W_1,\cdots,W_{n+m}) &= D_x F_m(X_n(x); W_{n+1},\cdots,W_{n+m}) \\ &= \Delta_m(X_n(x); W_{n+1},\cdots,W_{n+m}) \\ &\quad \cdot\Delta_n(x; W_1,\cdots,W_n)\end{aligned}$$

we have by independence

$$\|\sup_{z\in\mathbb{R}^d} |\Delta_{n+m}|(z)\|_p \le \|\sup_{z\in\mathbb{R}^d} |\Delta_n|(z)\|_p \cdot \|\sup_{z\in\mathbb{R}^d} |\Delta_m|(z)\|_p$$

i.e., this quantity is sub-multiplicative. By Proposition 7.2, we obtain thus

Proposition 8.11. *For the model (8.15), if moreover* $\mathbb{E}|F(0,W)|^p < +\infty$ *for some* $p\in[1,+\infty)$, *then* $u_p(x) := 1 + |x|^p \in \mathcal{A}(P)$ *and*

$$r_{exp}(P|_{b_{u_p}\mathcal{B}}) \le \inf_{n\ge 1}\left(\|\sup_{x\in\mathbb{R}^d} |D_x X_n(x)|\|_p\right)^{1/n}. \tag{8.25}$$

The quantity at the r.h.s. of (8.25) is the largest Lyapunov exposant.

8.6. A bordline case

Consider now the random perturbed dynamical system

$$X_0(x) := 0,\ X_{n+1}(x) = f(X_n(x)) + W_{n+1},\ n \geq 0, \tag{8.26}$$

where

(8.26a) $(W_n)_{n\geq 0}$ is a sequence of i.i.d.r.v. valued in $\mathbb{R}^d$, such that the law of $W = W_1$ is absolutely continuous w.r.t. the Lebesgue measure and $\mathbb{E}W = 0$;
(8.26b) $f : \mathbb{R}^d \to \mathbb{R}^d$ is measurable and locally bounded.

It is a particular case of the preceding model with $F(x, w) = f(x) + w$ (except that "$f \in C^1$" is not imposed here). Once if $f \in C^1$ and

$$\sup_{|x|\geq L} |D_x f(x)| < 1, \quad \text{for some } L > 0, \tag{8.27}$$

Condition (8.18) and (8.20) are satisfied, and $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 0$ by Proposition 8.8(b). We now treat a bordline case where $f(x) \sim x$ for large $|x|$.

Proposition 8.12. *For the model (8.26), assume that* $d = 1$. *Set*

$$v^{++} := \limsup_{x\to+\infty}(f(x) - x);\quad v^{+-} := \liminf_{x\to+\infty}(f(x) - x);$$
$$v^{-+} := \limsup_{x\to-\infty}(f(x) - x);\quad v^{--} := \liminf_{x\to-\infty}(f(x) - x)$$

which take values in $[-\infty, +\infty]$. *Let* Λ_W^* *be as in Proposition 8.1.*

(a) If $v^{+-} > -\infty$ *and* $v^{-+} < +\infty$,

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \geq \max\left\{e^{-\inf_{z>-v^{+-}} \Lambda_W^*(z)};\ e^{-\inf_{z<-v^{-+}} \Lambda_W^*(z)}\right\}. \tag{8.28}$$

In particular if $v^{+-} \geq 0$ *or* $v^{-+} \leq 0$, *then* $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 1$.

(b) Assume moreover that $-1_{x>0}f(x), 1_{x<0}f(x) \leq C$, *i.e., upper bounded. If* $-v^{++}, v^{--} > 0$ *(may be infinite), then*

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \leq \exp\left(-\Lambda_W^*(-v^{++}-) \wedge \Lambda_W^*(-v^{--}+)\right) \tag{8.29}$$

where $\Lambda_W^*(a+) := \lim_{z\to a+} \Lambda_W^*(z)$ *and* $\Lambda_W^*(a-) := \lim_{z\to a-} \Lambda_W^*(z)$.

(c) If $v^{++} = v^{+-} \in (-\infty, 0)$ *and* $v^{-+} = v^{--} \in (0, +\infty)$, *then*

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = \exp\left(-\Lambda_W^*(-v^{++}) \wedge \Lambda_W^*(-v^{--})\right). \tag{8.30}$$

We leave the verification of the following elementary remarks to the reader.

Remarks (8.v). By part (a) of this proposition, we have $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 1$ in each of the following cases:

(i) $v^{+-} \geq 0$ or $v^{-+} \leq 0$;
(ii) $v^{+-} < 0$ and $v^{-+} > 0$, but 0 is not an interior point of $\{a : \mathbb{E}e^{aW} < +\infty\}$.

In particular if $0 \notin \{a;\ \mathbb{E}e^{aW} < +\infty\}^0$, then $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 1$.

Remarks (8.vi). In case (c) above, intuitively $-v^{++}$ (resp. v^{--}) is the mean asymptotic velocity of $(X_n(x))$ to the direction of the origin when $x \to +\infty$ (resp. $x \to -\infty$). In case (c), $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) < 1$ iff $\mathbb{E}e^{aW} < +\infty$ for a belonging in some neighborhood of 0.

In case (b) of this proposition, $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) < 1$ if $\mathbb{E}e^{aW} < +\infty$ for a belonging in some neighborhood of 0.

Remarks (8.vii). In case of part (b) of this proposition, if 0 is an interior point of $\{a : \mathbb{E}e^{aW} < +\infty\}$, we have $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 0$ in each of the following cases:

(i) $v^{++} = -\infty$ and $v^{--} = +\infty$;
(ii) $v^{++} \in (-\infty, 0)$ and $v^{--} = +\infty$, and $\mathbb{P}(W > -v^{++}) = 0$ (see the proof in Step 4) below).
(iii) $v^{++} = +\infty$ and $v^{--} \in (0, +\infty)$ and $\mathbb{P}(W < -v^{--}) = 0$;
(iv) $v^{++} \in (-\infty, 0)$, $v^{--} \in (0, +\infty)$ and $\mathbb{P}(W > -v^{++}) = \mathbb{P}(W < -v^{--}) = 0$.

From the remarks above, we see that the exponential integrability of W plays an essential role in the estimate of $r_{ess}(P)$ in the actual critical case: that is completely different from the situation of Proposition 8.8.

Proof of Proposition 8.12. First of all notice that (A1) is satisfied by this model (but not (A2) for f may be non-continuous).

1) Our trick is to use comparison technique. Given constant $c \in \mathbb{R}$, consider a new Markov chain $X^c(x) = (X_n^c(x))_{n\geq 0}$ given by

$$X_0^c(x) = x,\ X_{n+1}^c(x) = X_n^c(x) + c + W_{n+1},\ \forall n \geq 0. \tag{8.31}$$

Denote by $(X_n(x))$ our chain given by (8.26) with $X_0(x) = x$. We shall prove the following simple facts.

(1) If $f(x) \leq x + c$ for all $x \geq N > 0$, then for all $x > N$,

$$X_n(x) \leq X_n^c(x),\ \forall 0 \leq n < \sigma_{(-\infty,N]}(X(x)).$$

(2) If $f(x) \geq x + c$ for all $x \geq N > 0$, then for all $x > N$,

$$X_n(x) \geq X_n^c(x),\ \forall 0 \leq n < \sigma_{(-\infty,N]}(X^c(x)).$$

(3) If $f(x) \leq x + c$ for all $x \leq -N < 0$, then for all $x < -N$,

$$X_n(x) \leq X_n^c(x),\ \forall 0 \leq n < \sigma_{[-N,+\infty)}(X^c(x)).$$

(4) If $f(x) \geq x + c$ for all $x \leq -N < 0$, then for all $x < -N$,

$$X_n(x) \geq X_n^c(x),\ \forall 0 \leq n < \sigma_{[-N,\infty)}(X(x)).$$

Here we give only the proof of (1). Indeed it is obviously true for $n = 0$. By recurrence assume that $X_n(x) \le X_n^c(x)$ where $0 \le n < \sigma_{(-\infty,N]}(X(x))$. Then $X_n(x) > N$ and we have

$$\begin{aligned} X_{n+1}(x) &= f(X_n(x)) + W_{n+1} \le X_n(x) + c + W_{n+1} \\ &\le X_n^c(x) + c + W_{n+1} = X_{n+1}^c(x), \end{aligned}$$

the desired result.

2) We show now part (a). For any $c < v^{+-}$, there is some $N > 0$ such that $f(x) > x + c$ for all $x \ge N$. By point (2) above, we have for all $x > N$,

$$\mathbb{P}\left(\sigma_{(-\infty,N]}(X(x)) > n\right) \ge \mathbb{P}\left(\sigma_{(-\infty,N]}(X^c(x)) > n\right).$$

Note that $X_n^c(x) = x + \sum_{k=1}^n (W_k + c)$ and

$$\left[\sigma_{(-\infty,N]}(X^c(x)) > n\right] \supset \left[\sum_{k=1}^m (W_k + c) \ge 0,\ \forall m = 1, \dots, n\right].$$

By Lemma 8.2 and its proof (2) (without condition $P(W > 0) > 0$ there), we have

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\sigma_{(-\infty,N]}(X^c(x)) > n\right) \ge -\inf_{z>0} \Lambda_{W+c}^*(z) = -\inf_{z>0} \Lambda_W^*(z-c).$$

Hence for any $c < v^{+-}$, we have

$$\begin{aligned} \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\sigma_{[-N,N]}(X(x)) > n\right) &\ge \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\sigma_{(-\infty,N]}(X(x)) > n\right) \\ &\ge -\inf_{z>-c} \Lambda_W^*(z) \end{aligned}$$

where it follows by Corollary 3.13,

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \ge \exp\left(-\inf_{z>-c} \Lambda_W^*(z)\right).$$

Letting $c \uparrow v^{+-}$, we obtain

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \ge \exp\left(-\inf_{z>-v^{+-}} \Lambda_W^*(z)\right).$$

Similarly using point (3), we get

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \ge \exp\left(-\inf_{z<-v^{-+}} \Lambda_W^*(z)\right).$$

Hence (8.28) is proved. For the last claim, we prove it only in the case where $v^{+-} \ge 0$.

- If $v^{+-} > 0$, then $-\inf_{z>-v^{+-}} \Lambda_W^*(z) \ge -\Lambda_W^*(0) = 0$ (for $\mathbb{E}W = 0$).
- If $v^{+-} = 0$, then by the proof (3) of Lemma 8.2 and the fact that $\mathbb{P}(W > 0) > 0$, we have $-\inf_{z>-v^{+-}} \Lambda_W^*(z) = -\inf_{z\ge 0} \Lambda_W^*(z) \ge -\Lambda_W^*(0) = 0$.

3) We prove now part (b). At first if 0 is not an interior point of $[\Lambda_W < +\infty]$, then the r.h.s. of (8.29) is 1 (verification left to the reader), i.e., (8.29) holds true automatically in that case. Assume then 0 is an interior point of $[\Lambda_W < +\infty]$.

Let $u_a(x) := e^{ax}$ where $a > 0$ verifies $\mathbb{E}e^{aW} < +\infty$. For any $0 > c_1 > v^{++}$, there is some $N_1 > 0$ such that for all $x > N_1$, $f(x) < x + c_1$. We have

$$Pu_a(x) = \mathbb{E}e^{a(f(x)+W)} \le 1_{[x>N_1]}u_a(x)e^{ac_1+\Lambda_W(a)} + 1_{[x\le N_1]}\mathbb{E}e^{a(f(x)+W)}.$$

Similarly let $u_{-b}(x) := e^{-bx}$ where $b > 0$ verifies $\mathbb{E}e^{-bW} < +\infty$. For any $0 < c_2 < v^{--}$, there is some $N_2 > 0$ such that for all $x < -N_2$, $f(x) > x + c_2$. We have

$$\begin{aligned} Pu_{-b}(x) &= \mathbb{E}e^{-b(f(x)+W)} \\ &\le 1_{[x<-N_2]}u_{-b}(x)e^{-bc_2+\Lambda_W(-b)} + 1_{[x\ge -N_2]}\mathbb{E}e^{-b(f(x)+W)}. \end{aligned}$$

By our extra assumption on $f(x)$ and our choices of a, b,

$$1_{[x\le N_1]}\mathbb{E}e^{a(f(x)+W)} + 1_{[x\ge -N_2]}\mathbb{E}e^{-b(f(x)+W)} \le L$$

(L depends on C, a, b, N_1, N_2 etc). Consequently

$$\begin{aligned} P(u_a(x) + u_b(x)) \le\ & 1_{x\notin[-N_2,N_1]}(u_a(x) + u_b(x)) \\ & \times \max\{e^{ac_1+\Lambda_W(a)},\ e^{-bc_2+\Lambda_W(-b)}\} + L \end{aligned}$$

where it follows by Theorem 3.11 that

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \le \max\{e^{ac_1+\Lambda_W(a)},\ e^{-bc_2+\Lambda_W(-b)}\}$$

for any $0 < a \in [\Lambda_W < +\infty]$, any $0 > -b \in [\Lambda_W < +\infty]$. Taking the infimum of the r.h.s. above over all such a, b, we obtain (by noting that $-c_1, c_2 > 0 = \mathbb{E}W$)

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \le \max\{e^{-\Lambda_W^*(-c_1)},\ e^{-\Lambda_W^*(-c_2)}\}.$$

Letting $c_1 \downarrow v^{++}$ and $c_2 \uparrow v^{--}$, we obtain the desired result.

4) It remains to show part (c). Since $v^{++} = v^{+-} < 0$ and $v^{-+} = v^{--} > 0$ (and finite), we see that $-1_{x>0}f(x)$, $1_{x<0}f(x)$ are upper bounded. By part (b) and the lower semi-continuity of Λ_W^*, we have

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) \le \max\{e^{-\Lambda_W^*(-v^{++})},\ e^{-\Lambda_W^*(-v^{--})}\}.$$

Thus by part (a), for (8.30), we have only to show that

$$\Lambda_W^*(-v^{++}) = \inf_{z>0} \Lambda_W^*(-v^{++} + z),\quad \Lambda_W^*(-v^{--}) = \inf_{z>0} \Lambda_W^*(-v^{--} - z).$$

Here only the first will be proved. For it notice that if $\mathbb{P}(W > -v^{++}) > 0$, this fact is proved in part (3) of the proof of Lemma 8.2 (for $\Lambda_W^*(-v^{++}+z) = \Lambda_{W+v^{++}}^*(z)$). Assume then $\mathbb{P}(W > -v^{++}) = 0$.

In that case $\inf_{z>0} \Lambda_W^*(-v^{++}+z) = +\infty$ by (8.9). Assume by absurd that $\Lambda_W^*(-v^{++}) < +\infty$. By the continuity of Λ_W^* on $[0, -v^{++}]$ and (8.9), there exists a sequence of probability measures (ν_n) on $\mathbb{R}$ such that $\int_{\mathbb{R}} x d\nu_n(x) \to -v^{++}$ and $h(\nu_n; \mu) \to \Lambda_W^*(-v^{++})$ (where μ, recalling it, is the law of W). Hence $\nu_n \ll \mu$ and $\{d\nu_n/d\mu\}$ is μ-uniformly integrable. Therefore there is a sub-sequence (n_k) such that $\{d\nu_{n_k}/d\mu\}$ converges to $h \in L^1(\mathbb{R}, \mu)$ in the topology $\sigma(L^1(\mu), L^\infty)$ (h must be a probability density function). Noting that $\mathbb{P}(W > -v^{++}) = 0$ (assumption), we have

$$\begin{aligned}\int_{\mathbb{R}} xh(x)d\mu(x) &= \int_{(-\infty,-v^{++}]} xh(x)d\mu(x)\\ &= \inf_{L>0} \int_{(-\infty,-v^{++}]} [x \vee (-L)]h(x)d\mu(x)\\ &= \inf_{L>0} \lim_{k\to\infty} \int_{(-\infty,-v^{++}]} [x \vee (-L)]d\nu_{n_k}(x)\\ &\geq \limsup_{k\to\infty} \int_{(-\infty,-v^{++}]} x d\nu_{n_k}(x) = -v^{++}.\end{aligned}$$

That is impossible by the absolute continuity of the law μ of W and $\mu([-v^{++}, +\infty)) = \mathbb{P}(W > -v^{++}) = 0$. In other words we have shown $\Lambda_W^*(-v^{++}) = +\infty$ as desired. The proof is completed. □

We now present an example to illustrate the results obtained previously.

Example 8.13. In the model (8.26) with $d = 1$, let $f(x) := (a|x|^\alpha + b)sgn(x)$ for $|x| \geq 1$ and $f \in C^1(\mathbb{R})$, where $a > 0$ and b, α are real constants. Assume that $\mathbb{E}e^{\delta|W|} < +\infty$ for some $\delta > 0$.

(i) If $\alpha < 0$, then $r_{ess}(P|_{b\mathcal{B}}) = 0$ by Corollary 8.10.
(ii) If $\alpha = 0$, then $\beta_w(P) = 0$, then $r_{ess}(P|_{b\mathcal{B}}) = 0$ by Theorem 3.5.
(iii) If $\alpha > 0$, then $\lim_{|x|\to\infty} |X_n(x)| = +\infty$ and consequently $r_{ess}(P|_{b\mathcal{B}}) = 1$ by Corollary 3.3.
(iv) If $\alpha < 1$, or if $\alpha = 1$ but $a < 1$, then $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 0$ by Proposition 8.8(b).
(v) If $\alpha = 1$ and $a = 1$, then

$$\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = \begin{cases} 1, & \text{if } b \geq 0 \\ \exp\left(-\Lambda_W^*(-b) \wedge \Lambda_W^*(b)\right), & \text{if } b < 0 \end{cases}$$

where the first case and the second follow respectively by part (a) and (c) of Proposition 8.12.
(vi) If $\alpha = 1$ and $a > 1$ or if $\alpha > 1$, then $\inf_{u\in\mathcal{A}(P)} r_{ess}(P|_{b_u\mathcal{B}}) = 1$ by Proposition 8.12(a).

Assume moreover that the density function of the law of W is dx-a.e. positive, (A3) is verified in such case. Then by the table above, the uniform τ-LDP holds iff $\alpha \leq 0$ by Theorem 6.3; and the local τ-LDP holds iff $\alpha < 1$, or $\alpha = 1$ but $b < 0$ with $\Lambda_W^*(-b) \wedge \Lambda_W^*(b) = +\infty$ (which is equivalen to $\mathbb{P}(W \notin [b, -b]) = 0$), by Theorem 6.4.

9. Proof of results of Sections 3 and 4

9.1. Proof of Proposition 3.2.

In this paragraph, $\beta(P) := \beta(P|_{b\mathcal{B}})$, the measure of non-compactness given in (2.5), $\|f\| = \sup_{x\in E}|f(x)|$, and $\|P\|$ denotes the norm of P on $b\mathcal{B}$.

Proof of Part (a). • $\beta_w(P) \le 2\beta(P)$: By the definition (2.5) of $\beta(P|_{b\mathcal{B}})$, for any $r > 2\beta(P|_{b\mathcal{B}})$, there exist a finite number of compacts $K_1, \cdots, K_N$ such that

$$\min_{1\le j\le N} \sup_{K\subset\subset E} \|P1_{K^c} - P1_{K_j^c}\| \le r.$$

Setting $C = \bigcup_j K_j$ which is again compact, we get from the above relation and the nonnegativeness of P,

$$\sup_{K:\, C\subset K\subset\subset E} \|P1_{C^c} - P1_{K^c}\| \le r.$$

Hence for any $x \in E$,

$$P(x, C^c) \le r + \inf_{K:\, C\subset K\subset\subset E} P(x, K^c) = r$$

by the tightness of the measure $P(x, dy)$. As that inequality holds for all $x \in E$, we get $\beta_w(P) \le \|P1_{C^c}\| \le r$, which is desired inequality for $r > 2\beta(P|_{b\mathcal{B}})$ is arbitrary.

We turn now to show (3.3). For any sequence $(f_n)_{n\ge 0} \subset C_b(E)$ decreasing pointwise to zero over E, with $\|f_0\| \le 1$ and for any $K \subset\subset E$, we have by Dini's monotone convergence theorem that f_n converges to zero uniformly over K. Consequently

$$\lim_{n\to\infty} \|Pf_n\| \le \lim_{n\to\infty} \|P(1_K f_n)\| + \lim_{n\to\infty} \|P(1_{K^c} f_n)\| \le \|P1_{K^c}\|.$$

As (f_n) and $K \subset\subset E$ are arbitrary, the desired inequality follows.

We prove the inverse inequality for (3.3) in the locally compact case. Take a sequence of compacts (K_n) such that $K_n \uparrow E$ and $K_n \subset (K_{n+1})^0$ (interior). We may construct continuous $f_n : E \to [0, 1]$ such that $f_n|_{K_n} = 0$ and $f_n(x) = 1$ for x belonging to (?) the closed set $E\backslash(K_{n+1})^0$. Such sequence (f_n) is necessarily decreasing to 0 pointwise over E. Thus

$$\beta_w(P) = \lim_{n\to\infty} \sup_{x\in E} P(x, K_{n+1}^c) \le \lim_{n\to\infty} \|Pf_n\|. \qquad \square$$

Proof of Part (b). Obviously $\beta_\tau(P) \le \sup_{(f_n)} \lim_{n\to\infty} \|Pf_n\|$. For the inverse inequality, let $(f_n)_{n\ge 0}$ be an arbitrary sequence in $b\mathcal{B}$ converging pointwise to zero over E with $\sup_n \|f_n\| \le 1$. For any $\varepsilon > 0$, let $A_n = \bigcup_{k\ge n}[|f_k| > \varepsilon]$. Then (A_n) decreases to $\emptyset$. Consequently

$$\lim_{n\to\infty} \|Pf_n\| \le \lim_{n\to\infty} \|P(1_{A_n})\| + \varepsilon\|P\| \le \beta_\tau(P) + \varepsilon\|P\|$$

where the desired result follows. $\square$

Proof of Part (c). We show now $\beta_\tau(P) \le 2\beta(P|_{b\mathcal{B}})$. Let (A_n) be an arbitrary sequence in $\mathcal{B}$ decreasing to $\emptyset$. By the definition (2.5) of $\beta(P|_{b\mathcal{B}})$, for any $r > 2\beta(P|_{b\mathcal{B}})$, there exists some $N \ge 1$ such that

$$\min_{0\le j\le N} \sup_{n\ge 0} \|P1_{A_n} - P1_{A_j}\| \le r.$$

But by the nonnegativeness of P we obtain

$$\sup_{n\ge N} \|P1_{A_N} - P1_{A_n}\| \le r.$$

It implies that for any $x \in E$, $P1_{A_N}(x) \le r + \lim_{n\to\infty} P1_{A_n}(x) = r$ (by monotone convergence). Thus

$$\lim_{n\to\infty} \|P1_{A_n}\| \le \|P1_{A_N}\| \le r$$

where the desired inequality follows. □

Proof of Part (d). They are obvious. □

Proof of Part (e). The case β_τ: For any sequence $(A_n)_{n\ge 0} \subset \mathcal{B}$ decreasing to $\emptyset$, $f_n := P1_{A_n}$ decreases pointwise to zero over E. For any $r > \beta_\tau(P)$, there exists some $N \ge 1$ such that $\|f_N\| \le r$. Set $g_n := f_{N+n}/r \in b\mathcal{B}$, which decreases to zero with $g_0 \le 1$. We have by part (b),

$$\lim_{n\to\infty} \|QP1_{A_n}\| \le r \lim_{n\to\infty} \|Qg_n\| \le r \cdot \beta_\tau(Q)$$

where follows $\beta_\tau(QP) \le \beta_\tau(Q)\beta_\tau(P)$.

The case β_w: For any $K \subset\subset E$, since $\beta_w(1_K P) = 0$ (assumption), by Prokorov's tightness criterion, for any $\varepsilon > 0$, there is some $K_1 \subset\subset E$ such that

$$\|1_K P 1_{K_1^c}\| < \varepsilon.$$

Now for any compact $K_2 \supset K_1$,

$$\beta_w(QP) \le \|Q1_K P1_{K_2^c}\| + \|Q1_{K^c} P1_{K_2^c}\| \le \|Q\|\varepsilon + \|Q1_{K^c}\| \cdot \|P1_{K_2^c}\|.$$

Taking the infimum at first for all compact $K_2 \supset K_1$ and next over all compact K, we obtain

$$\beta_w(QP) \le \|Q\|\varepsilon + \beta_w(Q) \cdot \beta_w(P).$$

As $\varepsilon > 0$ is arbitrary, the desired inequality follows. □

Proof of Part (f). For any sequence $(A_n)_{n\ge 0} \subset \mathcal{B}$ decreasing to $\emptyset$, $f_n := P1_{A_n}$ decreases pointwise to zero over E, and that convergence is uniform on each $K \subset\subset E$ by the assumption $\beta_\tau(1_K P) = 0$. Consequently

$$\lim_{n\to\infty} \|(QP)1_{A_n}\| = \lim_{n\to\infty} \|Q1_{K^c} P1_{A_n}\| \le \|Q1_{K^c}\| \cdot \beta_\tau(P).$$

As (A_n) and $K \subset\subset E$ are arbitrary, part (f) follows. □

Proof of Part (g). By Lemma 9.1 below, for any $K \subset\subset E$, $\beta(Q1_K P) = 0$. Hence

$$\beta(QP) \le \beta(Q1_K P) + \beta(Q1_{K^c} P) = \beta(Q1_{K^c} P) \le \|Q1_{K^c}\| \cdot \|P\|$$

where it follows $\beta(QP) \le \beta_w(Q)\|P\|$. But $\beta_\tau(Q) = 0$ implies $\beta_w(Q) = 0$, by Lemma 3.1. Thus $\beta(QP) = 0$, the desired result. □

In the proof of part (g) of Proposition 3.2 above, we have used

Lemma 9.1. *If $\beta_\tau(P) = 0 = \beta_\tau(Q)$, then QP is a compact operator on $b\mathcal{B}$.*

Proof. **1)** *Claim: if $\beta_\tau(P) = 0$, then P is a Dunford-Pettis operator (see [30], Def. 3.7.6., p219), i.e., if $f_n \in b\mathcal{B}$ weakly converges to zero, then $\|Pf_n\| \to 0$.*

Indeed, for such sequence (f_n), $\sup_n \|f_n\| \le C < +\infty$, and $f_n(x) \to 0$ for every $x \in E$ (for δ_x belongs to the topological dual space $(b\mathcal{B})'$ of $b\mathcal{B}$). Thus by Proposition 3.2(a) and our assumption, $\|Pf_n\| \to 0$.

2) By 1) and [30], Prop. 3.7.11.(i), p221, the image $P(B(0, 1))$ of the unit ball $B(0, 1)$ in $b\mathcal{B}$ is weakly compact in $b\mathcal{B}$. In other words, P is weakly compact on $b\mathcal{B}$, and so does Q. But $b\mathcal{B}$ is a so called (abstract) *M-space*, by Schaefer [40] (Chap. II, §9, Corollary 1, p.128), product QP of two weakly compact operators on $b\mathcal{B}$ is compact. □

I believe that "$\beta_w(P) \le \beta_\tau(P)$" holds on a Polish space.

9.2. Proof of Theorem 3.5

Proof of Part (b). By Proposition 3.2(a) and (c),

$$\max\{\beta_\tau(P^n), \beta_w(P^n)\} \le 2\beta(P^n|_{b\mathcal{B}}).$$

Then by Nussbaum's formula (2.7),

$$\max\{\limsup_{n\to\infty}[\beta_\tau(P^n)]^{1/n},\ \limsup_{n\to\infty}\left[\beta_w(P^n)\right]^{1/n}\} \le r_{ess}(P|_{b\mathcal{B}}).$$

Moreover by Proposition 3.2(e), $\beta_\tau(P^{m+n}) \le \beta_\tau(P^m)\beta_\tau(P^n)$ for all $m, n \in \mathbb{N}$. Thus

$$\limsup_{n\to\infty}[\beta_\tau(P^n)]^{1/n} = \lim_{n\to\infty}[\beta_\tau(P^n)]^{1/n} = \inf_{n\ge 1}[\beta_\tau(P^n)]^{1/n}.$$

Hence the first inequality in (3.6) is proved. To prove the second inequality in (3.6), noting that $1_{K^c} P^n 1_{K^c} \ge (1_{K^c} P 1_{K^c})^n$, we have

$$\begin{aligned}\beta_w(P^n) = \inf_{K\subset\subset E} \|P^n 1_{K^c}\| &\ge \inf_{K\subset\subset E} \|1_{K^c} P^n 1_{K^c}\| \\ &\ge \inf_{K\subset\subset E} \|(1_{K^c} P 1_{K^c})^n\|. \end{aligned} \tag{9.1}$$

Thus by Gelfand's formula of spectral radius,

$$\limsup_{n\to\infty}[\beta_w(P^n)]^{1/n} \ge \inf_{n\ge 1}\inf_{K\subset\subset E} \|(1_{K^c} P 1_{K^c})^n\|^{1/n} = \inf_{K\subset\subset E} r_{sp}\,(1_{K^c} P 1_{K^c})$$

which is the desired second inequality in (3.6). □

Proof of Part (a). At first by Proposition 3.2(e) for β_w and by (A1), $\beta_w(P^{m+n}) \le \beta_w(P^m)\beta_w(P^n)$ for all $m, n \in \mathbb{N}$. Then

$$r_w(P) := \lim_{n\to\infty} [\beta_w(P^n)]^{1/n} = \inf_{n\ge 1} [\beta_w(P^n)]^{1/n}.$$

Let us show $r_{ess}(P|_{b\mathcal{B}}) = r_w(P)$. By part (b) proved above, $r_{ess}(P|_{b\mathcal{B}}) \ge r_w(P)$. To prove the crucial inverse inequality, notice that for each $K \subset\subset E$, $1_K P^{2N}$ is a compact operator on $b\mathcal{B}$ by Proposition 3.2(g). Hence from $P^{2N} = 1_K P^{2N} + 1_{K^c} P^{2N}$ and the invariance of the essential spectrum by compact perturbation (well known, see Lemma 2.1), we have

$$\sigma_{ess}\left(P^{2N}|_{b\mathcal{B}}\right) = \sigma_{ess}\left(1_{K^c} P^{2N}|_{b\mathcal{B}}\right), \ \forall K \subset\subset E.$$

Consequently

$$\begin{aligned} r_{ess}\,(P|_{b\mathcal{B}}) &= \left[r_{ess}\left(P^{2N}|_{b\mathcal{B}}\right)\right]^{1/2N} \le \inf_{K\subset\subset E}\left[r_{sp}\left(1_{K^c}P^{2N}|_{b\mathcal{B}}\right)\right]^{1/2N} \\ &= \inf_{K\subset\subset E}\inf_{k\ge 1} \|\left(1_{K^c}P^{2N}\right)^k\|^{1/2Nk} = \inf_{k\ge 1}\inf_{K\subset\subset E}\|\left(1_{K^c}P^{2N}\right)^k\|^{1/2Nk} \\ &\le \limsup_{k\to\infty}\left(\|P^{2N}\|\cdot\beta_w(P^{2N(k-1)})\right)^{1/2Nk} \end{aligned}$$

where the last inequality follows from $\left(1_{K^c}P^{2N}\right)^k 1 \le P^{2N(k-1)}(1_{K^c}P^{2N}1)$ for $k \ge 2$ and the definition of β_w. Thus the last inequality above implies

$$r_{ess}\,(P|_{b\mathcal{B}}) \le \limsup_{n\to\infty}\left(\beta_w(P^n)\right)^{1/n} = r_w(P)$$

the desired result. Hence the first equality in (3.5) is proved.

It remains to show: $\limsup_{n\to\infty}(\beta_w(P^n))^{1/n} = \inf_{K\subset\subset E} r_{sp}(1_{K^c}P1_{K^c})$. The "$\ge$" is already proved in part (b). For the inverse inequality, notice that for each $K \subset\subset E$,

$$P^n = (1_K P + 1_{K^c}P)^n = (1_{K^c}P)^n + \sum_{j\in\{0,1\}^n:\, j_k=0 \text{ for some } k} A_{j_1}\cdots A_{j_n}$$

where $A_0 := 1_K P$ and $A_1 := 1_{K^c}P$. Hence By Proposition 3.2(e) for β_w and by (A1), $\beta_w(A_{j_1}\cdots A_{j_n}) = 0$ if $j_k = 0$ for some $1 \le k \le n$. Consequently

$$\beta_w(P^n) = \beta_w\left((1_{K^c}P)^n\right) = \beta_w\left((1_{K^c}P1_{K^c})^{n-1}\cdot P\right) \le \|(1_{K^c}P1_{K^c})^{n-1}\|\cdot\|P\|$$

where it follows

$$\limsup_{n\to\infty}\left(\beta_w(P^n)\right)^{1/n} \le \limsup_{n\to\infty}\left(\|(1_{K^c}P1_{K^c})^{n-1}\|\cdot\|P\|\right)^{1/n} = r_{sp}(1_{K^c}P1_{K^c})$$

for all $K \subset\subset E$. So the last equality in (3.5) is shown. □

9.3. Proof of Theorem 4.6

The classical proof of this theorem in the “$U = 1$” case given in Revuz [39] is based on the one-to-one correspondence between bounded space-time harmonic functions and the tail σ-field of the associated Markov chain. This correspondence is no longer available because of non-boundedness of functions in $b_U\mathcal{B}$. Our proof below is based on representation of finite Markov matrix.

At first by Proposition 4.5, $r_{sp}(P|_{b_U\mathcal{B}}) = 1$. We divide the proof of (4.2) into three steps.

Step 1. By Proposition 2.3, there is some minimal integer $d \ge 1$ such that the peripherical spectrum $\{\lambda \in \sigma(P^d);\ |\lambda| = 1\}$ is reduced to the singleton $\{1\}$. Let Π be the eigenprojection of $Q := P^d$ associated with 1. Since $1 \notin \sigma_{ess}(P^N)$ by assumption, $L := dimRange(\Pi) < +\infty$. Let us show that $P^d\Pi = \Pi$, i.e., the algebraic and geometric multiplicities of the eigenvalue 1 of $Q = P^d$ coincide. To this end consider the Laurent series of $(\lambda I - Q)^{-1}$ near 1,

$$(\lambda - Q)^{-1} = A_{-m}(\lambda - 1)^{-m} + \cdots + A_{-1}(\lambda - 1)^{-1} + \sum_{k=0}^{\infty} A_k(\lambda - 1)^k$$

where $A_{-1} = \Pi$ and $A_{-k-1} = (Q - I)^k\Pi$ ([50], Chap. VIII, §8). We should prove that $m = 1$. Assume in contrary that $A_{-m} \neq 0$ for $m \ge 2$ in the development above. Hence there would be some continuous functional ϕ on the complexification of $b_U\mathcal{B}$ such that $\psi := (A_{-m})^*\phi \neq 0$. Since $(Q^* - I)\psi = 0$, by Proposition 4.2(a) (applied to $Q^U f = (1/U)Q(Uf)$ on $b\mathcal{B}$), ψ is a nonzero measure ν.

By the Markov property of $Q = P^d$, for each $\lambda > 1$,

$$(\lambda - 1)(\lambda - Q)^{-1} = (\lambda - 1)\sum_{k=0}^{\infty} \lambda^{-k-1}Q^k$$

is again Markov. Hence $A_{-m}f = \lim_{\lambda\to 1+}(\lambda - 1)^m(\lambda - Q)^{-1}f = 0$ for any bounded $f \ge 0$. Consequently for any $0 \le f \in b\mathcal{B}$,

$$\langle \nu, f\rangle = \langle \psi, f\rangle = \langle \phi, A_{-m}f\rangle = 0.$$

Hence $\nu = 0$ (for ν is a measure), a contradiction.

Step 2. Since $P^d\Pi = \Pi$ (proved in the Step above) and $r_{sp}(P^d - \Pi) < 1$, there are $C > 0$ and $r \in (0, 1)$ such that

$$\|P^{nd} - \Pi\|_{b_U\mathcal{B}\to b_U\mathcal{B}} \le Cr^n,\ \forall n \ge 0. \tag{9.2}$$

As Q^n is nonnegative, Π is nonnegative on $b_U\mathcal{B}$. Hence there is a basis of $Range(\Pi)$ constituted of nonnegative nonzero functions $\{f_1, \dots, f_L\} \subset b_U\mathcal{B}$ such that for any $0 \le f \in b_U\mathcal{B}$, $C_k(f) \ge 0,\ k = 1, \dots, L$ where $C_k(f),\ k = 1, \dots, L$ are determined by

$$\Pi f = \sum_{k=1}^{L} C_k(f) f_k.$$

As $\Pi 1 = \lim_{n\to\infty} Q^n 1 = 1$, we may choose (f_k) in such a way that $\sum_{k=1}^L f_k(x) = 1$ for any $x \in E$.

Now note that $Pf_l = P\Pi f_l = \Pi P f_l \in Range(\Pi)$, then

$$Pf_l = \sum_{k=1}^L f_k a_{kl},\ a_{kl} \geq 0.$$

Moreover from

$$1 = P\sum_{l=1}^L f_l = \sum_{k=1}^L f_k \left(\sum_{l=1}^L a_{kl}\right)$$

and the linear independence of (f_k), $\sum_{l=1}^L a_{kl} = 1$ for each k. In other words $A = (A_{kl})$ is a Markov matrix, similar to P restricted to $Range(\Pi)$. Hence $d \geq 1$ is the minimal positive integer such that the only eigenvalue of A^d is 1.

Such Markov matrix A can be represented in the following way: there are $d_j \geq 1$ $(j = 1, \ldots, k)$ vectors $z^{i,j} = (z_1^{i,j}, \cdots, z_L^{i,j})^t$ (^{t} means the transposition), $i = 1, \ldots, d_j$ such that

(i) $(z^{i,j};\ 1 \leq i \leq d_j,\ 1 \leq j \leq k)$ is a basis of $\mathbb{R}^L$;
(ii) $z_l^{i,j} \geq 0$, $\sum_{i,j} z_l^{i,j} = 1$ for each $l = 1, \ldots, L$;
(iii) for all i, j, $Az^{i,j} = z^{i-1,j}$ (for each j fixed, $i - l$ is calculated modulo d_j);
(iv) d is the least common multiple of d_j, $j = 1, \ldots, k$.

Now set

$$U_{i,j} = \sum_{l=1}^L z_l^{i,j} f_l,\ 1 \leq i \leq d_j,\ 1 \leq j \leq k$$

which constitute a new basis of $Range(\Pi)$. We have

$$PU_{i,j} = \sum_{l=1}^L z_l^{i,j} \sum_{k=1}^L f_k a_{kl} = \sum_{k=1}^L f_k z_k^{i-1,j} = U_{i-1,j}.$$

In particular, $P^d U_{i,j} = U_{i,j}$. Moreover $\sum_{i,j} U_{i,j} = 1$ over E.

Step 3. Now for any $f \in b_U\mathcal{B}$,

$$\Pi f = \sum_{i,j} C_{i,j}(f) U_{i,j},$$

where $\phi_{i,j}(f) = C_{i,j}(f)$ defines a nonnegative continuous functional on $b_U\mathcal{B}$. Since $\langle \phi_{i,j}, P^d f\rangle = \langle \phi_{i,j}, \Pi P^d f\rangle = \langle \phi_{i,j}, f\rangle$, applying Proposition 4.2 to $Q^U f := (1/U)Q(Uf)$ acting on $b\mathcal{B}$, we see that $\phi_{i,j}$ must be a nonnegative measure $\mu_{i,j}$ such that $\mu_{i,j}(U) < +\infty$ and $\mu_{i,j}P^d = \mu_{i,j}$. Moreover by

$$1 = \Pi 1 = \sum_{i,j} \mu_{i,j}(1) U_{i,j},\ \sum_{i,j} U_{i,j} = 1$$

we get $\mu_{i,j}(1) = 1$, and $E_{i,j} = [U_{i,j} = 1]$, i, j are pairwise disjoint. By $\mu_{i,j}(U_{i,j}) = 1$, $\mu_{i,j}$ is supported on $E_{i,j}$. In summary we have proved

$$\Pi = \sum_{j=1}^{k}\sum_{i=1}^{d_j} U_{i,j} \otimes \mu_{i,j}, \ \ P^l\Pi = \sum_{j=1}^{k}\sum_{i=1}^{d_j} U_{i-l,j} \otimes \mu_{i,j}.$$

Consequently by (9.2), for any $l = 0, 1, \dots, d-1$,

$$\|P^{nd+l} - P^l\Pi\|_{b_U\mathcal{B}\to b_U\mathcal{B}} \le \|P^{nd} - \Pi\|_{b_U\mathcal{B}\to b_U\mathcal{B}} \cdot \max_{0\le l\le d-1} \|P^l\|_{b_U\mathcal{B}\to b_U\mathcal{B}}$$

where (4.2) follows.

The proof of the last claim is the same as in Revuz [39], Chap.6, Theorem 3.7, so omitted. □

10. Proof of Theorem 6.1

Part (a). 1) Upper Bound. For any $V \in b\mathcal{B}$, consider $P_V(x, dy) := e^{V(x)}P(x, dy)$ and the Feynman-Kac formula says that

$$(P_V)^n f(x) := \mathbb{E}^x f(X_n) \exp\left(\sum_{k=0}^{n-1} V(X_k)\right).$$

As in [13] and [46] (Appendix B), introduce the uniform (upper) Cramer functional

$$\begin{aligned}\Lambda(V) \ &:= \limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in E} \mathbb{E}^x \exp\left(\sum_{k=0}^{n-1} V(X_k)\right)\\ &= \limsup_{n\to\infty} \frac{1}{n} \log \|(P_V)^n\|\\ &= \log r_{sp}\left(P_V|_{b\mathcal{B}}\right).\end{aligned} \tag{10.1}$$

By a generalized Ellis-Gärtner theorem, Theorem B.5 in [46], for the good upper bound of large deviations in part (a) w.r.t. the τ-topology, it suffices to establish the following monotone continuity of Λ on $b\mathcal{B}$:

(MC) For any $(V_n)_{n\in\mathbb{N}} \subset b\mathcal{B}$ such that $V_n(x) \downarrow 0$ for all $x \in E$, then $\Lambda(V_n) \to 0$.

To relate that property with the parameter β_τ, we note the following simple inequality

$$\Lambda(V) \le \frac{1}{N} \log \|P^N(e^{NV})\|, \ \forall N \ge 1. \tag{10.2}$$

Indeed, write $\sum_{k=0}^{nN-1} V(X_k) = \sum_{j=0}^{N-1} S_j$ where $S_j := \sum_{k=0}^{n-1} V(X_{kN} + j)$. We have by Jensen's inequality

$$\begin{aligned}\log \mathbb{E}^x \exp\left(\sum_{k=0}^{nN-1} V(X_k)\right) &= \log \mathbb{E}^x \exp\left(\frac{1}{N}\sum_{j=0}^{N-1} NS_j\right)\\ &\le \frac{1}{N}\sum_{j=0}^{N-1} \log \mathbb{E}^x \exp\left(NS_j\right)\\ &\le \frac{1}{N}\sum_{j=0}^{N-1} \log e^{N\|V\|}\left[\sup_{x\in E} P^N(e^{NV})(x)\right]^{n-1}\end{aligned}$$

where it follows

$$\begin{aligned}\Lambda(V) &= \limsup_{n\to\infty} \frac{1}{nN} \log \sup_{x\in E} \mathbb{E}^x \exp\left(\sum_{k=0}^{nN-1} V(X_k)\right)\\ &\le \frac{1}{N} \log \|P^N(e^{NV})\|,\end{aligned}$$

the desired inequality (10.2).

Having it let us verify (MC). Fix $N \ge 1$. For any sequence (V_n) specified in (MC), $f_n := e^{NV_n} - 1 \le e^{N\|V_0\|}$ decreases pointwise to 0 over E. Thus by Proposition 3.2(b),

$$\lim_{n\to\infty} \|P^N(e^{NV_n})\| \le 1 + e^{N\|V_0\|}\beta_\tau(P^N).$$

Consequently by (10.2) we have

$$\begin{aligned}\lim_{n\to\infty} \Lambda(V_n) &\le \lim_{n\to\infty} \frac{1}{N} \log \|P^N(e^{NV_n})\|\\ &\le \frac{1}{N} \log\left(1 + e^{N\|V_0\|}\beta_\tau(P^N)\right) \le \frac{1}{N} e^{N\|V_0\|}\beta_\tau(P^N)\end{aligned}$$

where the desired property (MC) follows by letting $N \to +\infty$ and by assumption $r_\tau(P) = 0$.

2) Lower Bound. By the argument in the proof of Theorem 5.1 in [46] and our assumption (6.3), it is enough to show that for any $A \in \mathcal{B}$ with $\mu(A) > 0$, there is some $N \ge 1$ such that

$$\inf_{x\in E} \mathbb{P}_x(L_N(A) > 0) > 0.$$

Indeed by the upper bound in (a) and the μ-irreducibility, there is a unique invariant probability measure $\alpha \sim \mu$, and for any $\delta \in (0, \alpha(A))$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in E} \mathbb{P}_x(|L_N(A) - \alpha(A)| > \delta) \le -\inf\{J(\nu);\ |\nu(A) - \alpha(A)| \ge \delta\}$$

and the last infimum must be attained by some ν_0 with $|\nu_0(A) - \alpha(A)| \geq \delta$. Such ν_0 is different from α, hence $J(\nu_0) > 0$ (otherwise ν_0 is an invariant probability measure, then coincides with α by the uniqueness of invariant probability measure). In other words we have derived

$$\limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in E} \mathbb{P}_x(|L_N(A) - \alpha(A)| > \delta) \leq -J(\nu_0) < 0,$$

where the desired property above follows immediately.

Part (b). For the good upper bound of large deviations w.r.t. the "w"-topology, by Theorem II.3.3 in [45], it is enough to establish the monotone continuity of Λ on $C_b(E)$, i.e., $\Lambda(V_n) \to 0$ for any sequence $(V_n) \subset C_b(E)$ decreasing pointwise to zero over E. Its verification is completely parallel to that of (MC) in part (a), so omitted. □

Acknowledgements. This work has been reported in different seminars and conferences in Universities of Orléans, Dijon, Versailles, Hong-Kong, Peking Normal etc. during 2002. I thank P. Maheux, S. Fang, E. Rio, K.L. Yang and M.F. Chen for their kind invitations and hospitalities. I am grateful to F.Y. Wang and M. Hino for stimulating discussions on the subject at the very beginning of this work, to M.F. Chen and L. Miclo for their conscientious comments on the first version, and to the referee for correction of typing mistakes.

References

1. Chen, M.F.: Optimal Markovian couplings and applications. Mathematica Sinica, New Series **10** (3), 260–275 (1994)
2. Chen, M.F.: Estimation of spectral gap for Markov chains. Mathematica Sinica, New Series **12** (4), 337–360 (1996)
3. Chen, M.F.: Equivalence of exponential ergodicity and L^2-exponential convergence for Markov chains. Stoch. Proc. Appl. **87**, 281–297 (2000)
4. Chen, M.F., Wang, F.Y.: Applications of coupling method to estimate of the first eigenvalue on manifolds. Science in China Serie A **23** (11), 1130–1140 (1993)
5. Chen, M.F., Wang, F.Y.: Estimation of spectral gap for elliptic operators. Trans. Am. Math. Soc. **349** (3), 1239–1267 (1997)
6. Chow, Y.S., Teicher, H.: Probability Theory. Springer-Verlag, 1978
7. Cyon, H.L., Froese, R.G., Kirsch, W., Simon, B.: Schrödinger operators with Application to Quantum Mechanics and Global Goemetry. Text and Monographs in Physics, Springer-Verlag, 1987
8. de Acosta, A.: Upper bounds for large deviations of dependent random vectors. Z. Wahrsch. verw. Geb. **69**, 551–565 (1985)
9. de Acosta, A.: Large deviations for vector valued additive functionals of a Markov process: lower bound. Ann. Probab. **16**, 925–960 (1988)
10. Dellacherie, D., Meyer, P.A.: Probabilités et Potentiels. Chap. I-IV, Hermann, 1976
11. Dembo, A., Zeitouni, O.: Large deviations Techniques and Applications. Second Edition. Springer, New York, 1998
12. de Pagter, B., Schep, A.R.: Measures of non-compactness of operators in Banach lattices. J. Funct. Anal. **78**, 31–55 (1988)
13. Deuschel, J.D., Stroock, D.W.: Large deviations. Pure and Appl. Math. Academic Press **137**, (1989)

14. Doeblin, W.: Sur les propriétés asymptotiques du mouvement régis par certain types de chaîne simples. Bull. Math. Soc. Roum. Sci. **39** (1), 57–115; (2), 3–61 (1937)
15. Doeblin, W.: Eléments d'une théorie générale des chaines simples constantes de Markov. Ann. Sci. de l'Ecole Normale Sup. **57** (III), 61–111 (1940)
16. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time, I-IV. Comm. Pur. Appl. Math. **28**, 1–47 and 279–301 (1975); **29**, 389–461 (1976); **36**, 183–212 (1983)
17. Doob, J.L.: Stochastic Processes. John Wiley & Sons, New York, 1953
18. Gong, F.Z., Wu, L.M.: Spectral gap of positive operators and its applications. Preprint main results are announced in CRAS t.331, Série I, 983–988 (2000)
19. Grillo, G.: On Persson's theorem in local Dirichlet spaces. Z. Anal. Anwendungen **17** (2), 329–338 (1998)
20. Horowitz, S.: Transition probabilities and contractions of L^∞. Z.W. **24**, 263–274 (1972)
21. Jain, N.C.: Large deviations for additive functionals of Markov processes. Ann. Probab. **17** (3), 1073–1098 (1990)
22. Kato, T.: Perturbation Theory For Linear Operators. 2nd ed. (2nd corrected printing), Springer-Berlin, 1984
23. Kemeny, J.G., Snell, J.L., Knapp, A.W.: Denumerable Markov chains. GTM 40, Springer-Verlag, 1976
24. Kendall, D.G.: Unitary dilations of Markov transition operators and the corresponding integral representation for transition probability matrices. In: Probability and Statistics, U. Grenander, (ed.), Almqvist and Wiksell, Stockholm, 1959, pp. 139–161
25. Ledoux, M.: Concentration of measure and logarithmic Sobolev inequalities. Séminaire de Probab. XXXIII, LNM **1709**, 120–216, Springer, (1999)
26. Ligget, T.M.: Interacting Particle Systems. Springer-Verlag, 1985
27. Ligget, T.M.: Exponential L_2 convergence of attractive reversible nearest particle systems. Ann. Probab. **17** (2), 403–432 (1989)
28. Lindvall, T.: Lectures on the coupling method. John Wiley & Sons, New York, 1992
29. Malyshev, V.A., Men'sikov, M.V.: Ergodicity, continuity and analyticity of countable Markov chains. Trans. Moscow Math. Soc. **1** 1–48 (1982)
30. Meyer-Nieberg, P.: Banach Lattices, Springer-Verlag, 1991
31. Meyn, S.P., Tweedie, R.L.: Markov chains and stochastic stability. Springer-Verlag, 1993
32. Miclo, L.: Relation entre isopérimétrie et trou spectral pour les chaî nes de Markov finies. Probab. Th. Rel. Fields **114**, 431–485 (1999)
33. Nagel, R.: One-Parameter Semigroups of Positive Operators. Lect. Notes in Math. **1184**, Springer-Verlag, Berlin, 1986
34. Nummelin, E.: General Irreducible Markov Chains and Non-Negative operators. Cambridge Tracts in Math., **83**, Cambridge University Press, 1984
35. Nummelin, E., Tuominen, P.: Geometric ergodicity of Harris recurrent Markov chains with applications to renewel theory. Stoch. Proc. Appl. **12**, 187–202 (1982)
36. Nummelin, E., Tuominen, P.: The rate of convergence in Orey's theorem for Harris recurrent Markov chains with applications to renewel theory. Stoch. Proc. Appl. **15**, 295–311 (1983)
37. Nussbaum, R.D.: The radius of the essential spectrum. Duke Math. J. **38**, 473–478 (1970)
38. Reed, M., Simon, B.: Methods of Modern Mathematical Physics, IV. Analysis of Operators, Academic Press, 1975
39. Revuz, D.: Markov Chains. North-Holland, 1976
40. Schaefer, H.H.: Banach Lattices and Positive Operators. Springer-Verlag, 1974

41. Wang, F.Y.: Functional inequalities for empty essential spectrum. J. Funct. Anal. **170**, 219–245 (2000)
42. Wang, F.Y.: Functional inequalities, semigroup properties and spectrum estimates. Infinite Dimensional Anal., Quantum probability and Rel. Fields. **3** (2), 263–295 (2000)
43. Weis, L.: On the computation of some quantities in the theory of Fredholm operators. In: Proceeding 12th Winer School on Abstract Anal., Suppl. Rend. Circ. Matem. Palermo II **5**, 1984
44. Wu, L.M.: Moderate deviations of dependent random variables related to CLT, Ann. Probab. **23**, 420–445 (1995)
45. Wu, L.M.: An introduction to large deviations (in chinese). In: Several Topics in Stochastic Analysis J.A. Yan, S. Peng, S. Fang, L. Wu, (eds.), Press of Sciences of China, Beijing, pp. 225–336
46. Wu, L.M.: Uniformly integrable operators and large deviations for Markov processes. J. Funct. Anal. **172**, 301–376 (2000)
47. Wu, L.M.: Some notes on large deviations of Markov processes. Acta Math. Sinica. English Series **16**, 369–394 (2000)
48. Wu, L.M.: Large and moderate deviations and exponential convergence for stochastic damping Hamilton systems. Stoch. Proc. Appl. **91** (2), 205–238 (2001)
49. Wu, L.M.: Uniform positive improvingness, tail norm condition and spectral gap. Preprint, 2002
50. Yosida, K.: Functional Analysis. Third Edition, Springer-Verlag, 1971