On metastability

Laurent Miclo*

Institut de Mathématiques de Toulouse, UMR 5219,
Toulouse School of Economics, UMR 5314
CNRS and University of Toulouse

Abstract

Consider finite state space irreducible and absorbing Markov processes. A general spectral criterion is provided for the absorbing time to be close to an exponential random variable, whatever the starting point. When exiting points are added to the state space, our criterion also insures that the exit time and position are almost independent. Since this is valid for any exiting extension of the state space, it corresponds to an instance of the metastability phenomenon. Simple examples at small temperature suggest that this new spectral criterion is quite sharp. But the main interest of the underlying quantitative approach, based on Poisson equations, is that it does not rely on a small parameter such as temperature, nor on reversibility.

Keywords: metastability, finite absorbing Markov process, exit time, exit position, spectral quantitative criterion, first Dirichlet eigenvalue, quasi-stationary distribution, Poisson equation, eigentime identity, small temperature.

MSC2010: primary: 60J27, secondary: 60E15, 15A42, 60J35, 60J45

*LM acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010.

1 Introduction

The **metastability** phenomenon occurs when a system relatively quickly reaches an apparent equilibrium, before this stochastic balance vanishes in a somewhat unpredictable way. This behavior can be found in various domains, such as physics, chemistry, biochemistry, neuroscience, population dynamics, economics, politics or even (personal?) history. The simplest mathematical model is based on absorbing finite Markov processes, when a quasi-stationary distribution is (almost) attained before the final absorption. The goal of this paper is to give a surprising simple spectral characterization of metastability in this Markovian context.

More precisely, consider a **sub-Markovian generator** $L \coloneqq (L(x,y))_{x,y\in V}$ on a finite state space V which contains at least two points (otherwise Theorem 1, Theorem 3 and Corollary 4 are trivially true with $\Sigma^\dagger = 0$). It is a matrix whose off-diagonal entries are non-negative and whose row sums are non-positive. We assume that L is **irreducible**, in the sense that for any $x \neq y \in V$, there exists a path $(x_k)_{k\in[\![0,l]\!]}$ (where $[\![0,l]\!] \coloneqq \{0,1,\ldots,l\}$) going from x to y: $x_0 = x$, $x_l = y$ and $L(x_k, x_{k+1}) > 0$ for all $k \in [\![0, l-1]\!]$. For any $x \in V$, $X_x \coloneqq (X_x(t))_{t\in[0,\tau_x)}$ will stand for an associated Markov process starting from x, up to its vanishing time τ_x.

Consider Λ the multiset of the eigenvalues of $-L$ counted with their algebraic multiplicities. By Perron-Frobenius' theorem, Λ contains an eigenvalue $\lambda_0 \geqslant 0$ which is strictly smaller than the real parts of all the other elements of Λ. It is sometimes called the **first Dirichlet eigenvalue** or the **exponential survival rate** of L, see for instance the book [3] of Collet, Martínez and San Martín. In particular, the algebraic multiplicity of λ_0 is 1. To avoid a trivial statement below, we assume that L is strictly sub-Markovian, in the sense that $\lambda_0 > 0$.

Perron-Frobenius' theorem also insures the existence and uniqueness of a probability ν on V, called the **quasi-stationary distribution**, such that

$$\nu L \quad = \quad -\lambda_0 \nu \tag{1}$$

For more details about the eigenmeasure ν, which gives positive weights to all the elements of V, we refer again to the book of Collet, Martínez and San Martín [3].

To state our main result, we need to introduce another spectral quantity. Consider δV the set of **interior exit points**:

$$\delta V \quad \coloneqq \quad \left\{ \omega \in V \,:\, \sum_{y\in V} L(\omega, y) < 0 \right\}$$

When a Markov process X_x, $x \in V$, associated to L visits a point of δV, there is a positive probability that it vanishes at its next attempt to jump. Let us transform this Markov process into an ergodic one, by requiring that instead of vanishing, a new position is chosen according to ν. It amounts to replace the sub-Markov generator L by the Markov generator $\widetilde{L}$ defined by

$$\forall\, x \neq y \in V, \qquad \widetilde{L}(x,y) \quad \coloneqq \quad L(x,y) + \left| \sum_{z\in V} L(x,z) \right| \nu(y) \tag{2}$$

(the entries of $\widetilde{L}$ on the diagonal are deduced from the fact that the row sums vanish).

For $x \neq y \in V$, $\widetilde{L}(x,y)$ is different from $L(x,y)$ if and only if $x \in \delta V$, in which case the vanishing rate $|\sum_{y\in V} L(x,y)|$ is dispatched into the jump rates $|\sum_{y\in V} L(x,y)|\nu(y)$. For $\omega \in \delta V$, denote

$$V_\omega^\dagger \quad \coloneqq \quad V \backslash \{\omega\}$$

which is non-empty, due to the hypothesis $\mathrm{card}(V) \geqslant 2$. Endow $V_\omega^\dagger$ with the sub-Markovian generator $L_\omega^\dagger \coloneqq (L_\omega^\dagger(x,y))_{x,y\in V_\omega^\dagger} \coloneqq (\widetilde{L}(x,y))_{x,y\in V_\omega^\dagger}$.

Consider $\Lambda_\omega^\dagger$ the multiset of the eigenvalues of $-L_\omega^\dagger$ counted with their algebraic multiplicities. Since there is no reason for $L_\omega^\dagger$ to be reversible (even when L is assumed to be reversible), a priori the elements of $\Lambda_\omega^\dagger$ are complex numbers whose real part is positive, by strict sub-Markovianity of $L_\omega^\dagger$, namely

$$\Lambda_\omega^\dagger \quad \subset \quad \{z \in \mathbb{C} \,:\, \Re(z) > 0\}$$

Since the entries of the matrix $L_\omega^\dagger$ are real-valued, the set $\Lambda_\omega^\dagger$ is stable by complex conjugation, so that the following quantity is positive

$$\Sigma_\omega^\dagger \quad \coloneqq \quad \sum_{\lambda \in \Lambda_\omega^\dagger} \frac{1}{\lambda} \tag{3}$$

Consider the probability ζ defined on δV by

$$\forall\ \omega \in \delta V, \qquad \zeta(\omega) \quad \coloneqq \quad \frac{\left|\sum_{y\in V} L(\omega, y)\right| \nu(\omega)}{\sum_{w\in\delta V} \left|\sum_{z\in V} L(w, z)\right| \nu(w)} \tag{4}$$

Finally introduce

$$\Sigma^\dagger \quad \coloneqq \quad \sum_{\omega\in\delta V} \Sigma_\omega^\dagger\, \zeta(\omega) \tag{5}$$

The interest of this quantity comes from the following surprisingly simple bound about metastability:

Theorem 1 *We have*

$$\sup_{x\in V} \sup_{t\geqslant 0} |\mathbb{P}[\tau_x > t] - \exp(-\lambda_0 t)| \quad \leqslant \quad 4\lambda_0 \Sigma^\dagger$$

The interpretation of this result is as follows. For any $\omega \in \delta V$, the quantity $\Sigma_\omega^\dagger$ measures how difficult it is to reach the interior exit boundary point ω for the underlying process. Then $\Sigma^\dagger$ stands an average over all the $\omega \in \delta V$: it measures the difficulty of "internal mixing". The quantity $1/\lambda_0$ quantifies the difficulty of getting out of the state space. Thus the above result states that when it is easier to mix than to exit, a metastability phenomenon occurs for the exit time (and the exit position according to the following bounds) and this principle can be quantified in a very clear and spectral manner.

With respect to the informal definition of metastability given at the beginning of this introduction, this theorem does not deal with the fact that an apparent equilibrium has been relatively quickly reached, but only with its vanishing in an unpredictable way (due to the memoryless property of the exponential distribution). In the present setting, the apparent equilibrium corresponds to the quasi-stationary distribution. To quantify the fact it has almost been attained well before the process vanishes, we can introduce **conditioned strong quasi-stationary times**: starting from $x \in V$, they are stopping times $\varsigma_x \leqslant \tau_x$ (with respect to the filtration generated by X_x and independent noise) such that conditioned by ς_x on $\{\varsigma_x < \tau_x\}$, the law of X_{ς_x} is the quasi-distribution ν. Taking into account that on $\{\varsigma_x < \tau_x\}$, X_{ς_x} is independent from ς_x and distributed according to ν implies that $\tau_x - \varsigma_x$ is conditionally distributed according to an exponential random variable of parameter λ_0. In particular, if $\lambda_0\Sigma^\dagger$ is very small, due to Theorem 1, ς_x will have to be negligible with respect to τ_x on $\{\varsigma_x < \tau_x\}$. Thus we would have a spectral characterization through the quantity $\lambda_0\Sigma^\dagger$ of the full metastability phenomenon if the following result was true:

Conjecture 2 For any $x \in V$, there exists a conditioned strong quasi-stationary time ς_x such that

$$\sup_{x\in V} \mathbb{P}[\varsigma_x = \tau_x] \quad \leqslant \quad C\lambda_0\Sigma^\dagger$$

where $C > 0$ is a universal constant.

□

When $\mathbb{P}[\varsigma_x = \tau_x] = 0$, ς_x is called a **strong quasi-stationary times**. Such times where constructed in [7] for birth and death processes starting from the non-absorbing boundary of their finite segment state spaces. For more general approaches that can be used for conditioned strong quasi-stationary times, see Fill [8] and [16]. Conditioned strong quasi-stationary times were formally introduced in Manzo and Scoppola [13] in the context of metastability. We will not investigate further the notion of conditioned strong quasi-stationary times in this paper. Instead we will study the behavior of the exit position distribution. The fact that the latter can be almost independent from the starting point, for *any* absorbing extension of L, as explained below, should be equivalent to the metastability of L as mentioned above Conjecture 2. Proposition 23 and Remark 29, valid in the small temperature framework of Section 5, are strong hints in this direction.

The bound of Theorem 1 extends into a similar result for the exit time and position couple. Denote $\bar{V} := V \sqcup \partial V$, where ∂V is a non-empty set not intersecting V. Be careful about the distinction: δV consists of *internal* boundary points, while the elements of ∂V will be *external* boundary points (even if it would be sufficient to choose a set ∂V in bijection with δV, each internal boundary point leading to exactly one external boundary point). Consider a Markov generator $\bar{L} := (\bar{L}(x,y))_{x,y\in\bar{V}}$ on $\bar{V}$ which is an **absorbing extension** of L:

$$\forall\ x,y\in\bar{V},\qquad \bar{L}(x,y) \quad=\quad \begin{cases} L(x,y) & \text{, when } x,y\in V\\ 0 & \text{, when } x\in\partial V\end{cases}$$

The weights $(\bar{L}(x',y'))_{x'\in V,\, y'\in\partial V}$ enable, for any $x\in V$, to extend X_x into a Markov process $\bar{X}_x := (\bar{X}_x(t))_{t\geqslant 0}$ taking values in $\bar{V}$ in the following way: the value $\bar{X}_x(\tau_x) = y$ is chosen with the probability measure proportional to $(\bar{L}(X_x(\tau_x-),y))_{y\in\partial V}$, and afterward we take $\bar{X}_x(t) = X_x(\tau_x)$ for all $t \geqslant \tau_x$.

Consider the probability measure μ defined on ∂V by

$$\forall\ y\in\partial V,\qquad \mu(y) \quad:=\quad \frac{1}{Z}\sum_{x\in V}\nu(x)\bar{L}(x,y) \tag{6}$$

where Z is the normalizing constant:

$$Z \quad:=\quad \sum_{x\in V,\, y\in\partial V}\nu(x)\bar{L}(x,y)$$

which is positive, due to the sub-Markov assumption. Up to removing from ∂V the points $y\in\partial V$ such that $\mu(y) = 0$, we can assume that μ gives a positive weight to all points of ∂V.

Recalling the definitions (3) and (5) we introduce another probability χ on ∂V:

$$\forall\ \omega\in\partial V,\qquad \chi(\omega) \quad:= \tfrac{1}{Z\Sigma^\dagger}\textstyle\sum_{x\in\delta V}\Sigma_x^\dagger\nu(x)\bar{L}(x,\omega)$$

(one would have noted that for any $x\in\delta V$, $\left|\sum_{y\in V}L(x,y)\right| = \sum_{\omega\in\partial V}\bar{L}(x,\omega)$, so that $Z\Sigma^\dagger$ is indeed the normalizing constant in the above formula). Since the quantities $\Sigma_x^\dagger$ are positive on δV and that the support of μ is ∂V, we get that the support of χ is also ∂V.

The distribution of the exit couple satisfies:

Theorem 3 *We have*

$$\sup_{x\in V,\, y\in\partial V,\, t\geqslant 0}\left|\frac{\mathbb{P}[\tau_x\leqslant t,\, X_x(\tau_x) = y] - (1-\exp(-\lambda_0 t))\mu(y)}{\chi(y)}\right| \quad\leqslant\quad 12\lambda_0\Sigma^\dagger$$

In practice, V will be a subset of a larger state space and ∂V will be the set of nearest (outward) neighbors of V in this bigger space. Theorem 3 and the assumption that $\lambda_0\Sigma^\dagger$ is small will then enable to replace V by a single point in order to reduce the state space, leading to a controlled clustering procedure for Markov processes.

Despite numerous investigations of metastability, see e.g. the book of Bovier and den Hollander [1] or the recent paper of Di Gesù, Lelièvre, Le Peutrec and Nectoux [4], as well as the references therein (even if these two works are mainly dealing with continuous frameworks), both bounds of Theorem 1 and 3 seem to be new. They are in fact generalizations of some estimates of [15], which was restricted to the reversible and small temperature setting, and without spectral interpretation of the bounds.

For $x \in V$, let μ_x be the distribution of the exit position, namely the law of $X_x(\tau_x)$. It becomes closer and closer to μ, as $\lambda_0\Sigma^\dagger$ goes to zero, as an immediate consequence of Theorem 3, by letting t go to infinity and summing over $y \in V$ the bound

$$|\mathbb{P}[\tau_x \leqslant t,\, X_x(\tau_x) = y] - (1 - \exp(-\lambda_0 t))\mu(y)| \quad \leqslant \quad 12\lambda_0\Sigma^\dagger\chi(y)$$

Recall that the total variation norm between two probability measures γ', γ on the same finite space V is given by

$$\|\gamma' - \gamma\|_{\text{tv}} \quad \coloneqq \quad \sum_{y\in V} |\gamma(y) - \gamma'(y)|$$

Corollary 4 *We have*

$$\sup_{x\in V} \|\mu_x - \mu\|_{\text{tv}} \quad \leqslant \quad 12\lambda_0\Sigma^\dagger$$

Typically, the above results are to be applied to families of absorbed Markov processes $(L^{(n)})_{n\in\mathbb{N}}$ on respective state spaces $(V^{(n)},)_{n\in\mathbb{N}}$ and metastability will occur if

$$\lim_{n\to\infty} \lambda_0^{(n)}\Sigma^{(n)\dagger} \quad = \quad 0$$

Then for large n, the exit time is close to an exponential random variable and the exit time and position are almost independent.

This behavior is radically opposite to the cut-off phenomenon, see for instance the review by Diaconis [5]. Or at least to its sub-Markovian version, where the absorbing times are investigated instead of the more classical mixing or strong stationary times (of course there are relations between these absorbed and ergodic versions, see for instance Diaconis and Fill [6]). A strong stationary time is a finite stopping time σ such that σ is independent from the stopped position which is furthermore distributed according to the stationary distribution. In the cut-off phenomenon for strong stationary times, they become asymptotically deterministic, while in metastability, the absorbing times become asymptotically totally impredictable exponential times.

The metastability phenomenon is illustrated in Section 5 by very simple examples on two-point or three-point state spaces at small temperature. It provides a hint of the sharpness of Corollary 4 and of the results of next section, while discussing that of Theorem 1. The situation of generalized Metropolis algorithms will be treated in a future manuscript, including an investigation of quasi-invariant probability measures at small temperature (which requires some care, see Lemma 30 at the end of the present paper). The traditional Metropolis algorithms (where an additional reversibility assumption is made) could be treated with the help of the computations of [15], which served as a distant model for the present paper. Nevertheless, our motivation here is to go beyond such small temperature settings and to propose a general spectral criterion for metastability for irreducible finite sub-Markovian processes, in particular the reversibility is now completely removed, due to the introduction of the important quantity $\Sigma^\dagger$, as shown by Theorems 1 and 3.

The plan of the paper is as follows. The next section present some estimates on the solutions of Poisson equations, which are at the heart of our approach. Sections 3 and 4 respectively deal with the proofs of Theorems 1 and 3. Section 5 is devoted to the explicit treatment of the generic two-point state space case at small temperature, as well as of some three-point state space examples, which despite their apparent simplicity, already displays important features of more general cases.

2 Poisson equation

The main ingredient in the proofs of Theorems 1 and 3 is an estimate on the solutions of some Poisson equations. Let us present them in a general finite framework.

Let $\widetilde{L} \coloneqq (\widetilde{L}(x,y))_{x,y\in V}$ be an irreducible Markov generator on a non-empty finite state space V. Denote by $\widetilde{\pi}$ its unique invariant measure and let us fix a point $\omega \in V$. Let φ be the unique function on V solution to the **Poisson equation**

$$\left\{ \begin{array}{rcl} \widetilde{L}[\varphi] &=& \mathbb{1}_{\{\omega\}} - \widetilde{\pi}(\omega) \\ \widetilde{\pi}[\varphi] &=& 0 \end{array} \right. \tag{7}$$

Our purpose in this section is to give some bounds on $\|\varphi\|_\infty$. To do so, we need to introduce the following objects, similarly to the introduction, except we will not put ω in index of $V^\dagger$ and $L^\dagger$, because ω is fixed in Theorem 5 below. Consider $V^\dagger \coloneqq V\backslash\{\omega\}$ and $L^\dagger$ the absorbing sub-Markov generator $(L^\dagger(x,y))_{x,y\in V^\dagger} \coloneqq (\widetilde{L}(x,y))_{x,y\in V^\dagger}$. Let $\Lambda^\dagger$ be the multi-set consisting of the spectrum of $-L^\dagger$ with its algebraic multiplicities. Note that by irreducibility of $\widetilde{L}$, $0 \notin \Lambda^\dagger$, which enables us to introduce

$$\Sigma^\dagger \quad \coloneqq \quad \sum_{\lambda\in\Lambda^\dagger} \frac{1}{\lambda}$$

The main result of this section is:

Theorem 5 *We have*

$$\|\varphi\|_\infty \quad \leqslant \quad 2\widetilde{\pi}(\omega)\Sigma^\dagger$$

The proof of this result will require two steps presented in the next subsections, first a rough bound that will next be refined.

2.1 A rough estimate

Consider $\widetilde{\Lambda}$ the multiset consisting of the spectrum of $-\widetilde{L}$ with its algebraic multiplicities. By Markovianity and irreducibility of $\widetilde{L}$, $0 \in \widetilde{\Lambda}$ with multiplicity 1. Denote $\widetilde{\Lambda}_* \coloneqq \widetilde{\Lambda}\backslash\{0\}$ and

$$\widetilde{\Sigma}_* \quad \coloneqq \quad \sum_{\lambda\in\widetilde{\Lambda}_*} \frac{1}{\lambda}$$

More generally than (7), we consider for any $x \in V$, the solution φ^x of the Poisson equation

$$\left\{ \begin{array}{rcl} \widetilde{L}[\varphi^x] &=& \mathbb{1}_{\{x\}} - \widetilde{\pi}(x) \\ \widetilde{\pi}[\varphi^x] &=& 0 \end{array} \right. \tag{8}$$

The interest of these objects is:

Proposition 6 *We have*

$$\|\varphi\|_\infty \;\leqslant\; \max_{x\in V}\|\varphi^x\|_\infty \;\leqslant\; \widetilde\Sigma_*$$

in particular the last r.h.s. is real (this can also be seen from the complex conjugation stability of $\widetilde\Lambda_*$) and positive (as soon as V is not reduced to a singleton).

Proof

For $y\in V$, consider $\widetilde X_y \coloneqq (\widetilde X_y(t))_{t\geqslant 0}$ be a Markov process starting from y and whose generator is $\widetilde L$. For $x,y\in V$, define the absorption time

$$\widetilde\tau_y^x \;\coloneqq\; \inf\{t\geqslant 0\,:\, \widetilde X_y(t)=x\}$$

Applying the martingale problem to the function φ^x up to the time $\widetilde\tau_y^x\wedge t$, with given $t\geqslant 0$, we get

$$\begin{aligned}\varphi^x(\widetilde X_x(t\wedge\widetilde\tau_y^x)) &= \varphi^x(y)+\int_0^{t\wedge\widetilde\tau_y^x}\widetilde L[\varphi^x](\widetilde X_x(s))\,ds+M_{t\wedge\widetilde\tau_y^x}\\ &= \varphi^x(x)-\nu(x)(t\wedge\widetilde\tau_y^x)+M_{t\wedge\widetilde\tau_y^x}\end{aligned}$$

where $(M_t)_{t\geqslant 0}$ is a martingale. Taking expectations, we deduce

$$\mathbb{E}[\varphi^x(\widetilde X_x(t\wedge\widetilde\tau_y^x))] \;=\; \varphi^x(y)-\nu(x)\mathbb{E}[t\wedge\widetilde\tau_y^x]$$

Since V is finite and $\widetilde L$ is irreducible, φ^x is bounded and $\widetilde\tau_y^x$ is a.s. finite, so we can let t go to infinity in the above formula and obtain

$$\varphi^x(x) \;=\; \varphi^x(y)-\nu(x)\mathbb{E}[\widetilde\tau_y^x]$$

In particular for any $x,y\in V$, we have $\varphi^x(x)\leqslant\varphi^x(y)$. Since $\nu[\varphi^x]=0$, it follows that

$$\alpha_x \;\coloneqq\; -\varphi^x(x)\;\geqslant\; 0$$

For $x\in V$ fixed, integrating the relations

$$\forall\ y\in V,\qquad \varphi^x(y) \;=\; -\alpha_x+\nu(x)\mathbb{E}[\widetilde\tau_y^x] \tag{9}$$

with respect to $\nu(y)$, we get

$$\begin{aligned}0 &= \nu[\varphi^x]\\ &= -\alpha_x+\sum_{y\in V}\mathbb{E}[\widetilde\tau_y^x]\nu(x)\nu(y)\end{aligned}$$

namely

$$\alpha_x \;=\; \sum_{y\in V}\mathbb{E}[\widetilde\tau_y^x]\nu(x)\nu(y) \tag{10}$$

The eigentime identity (for a simple proof see e.g. [17]) asserts that

$$\forall\ y\in V,\qquad \sum_{x\in V}\mathbb{E}[\widetilde\tau_y^x]\nu(x) \;=\; \widetilde\Sigma_*$$

thus summing with respect to $x\in V$, we obtain

$$\sum_{x\in V}\alpha_x \;=\; \widetilde\Sigma_*$$

and in particular

$$\forall\, x \in V, \qquad \alpha_x \;\leqslant\; \widetilde{\Sigma}_* \tag{11}$$

Coming back to (9), we deduce

$$\forall\, x \in V, \qquad \|\varphi^x\|_\infty \;\leqslant\; \alpha_x \vee \max_{y\in V} \nu(x)\mathbb{E}[\widetilde{\tau}_y^x]$$

According to the eigentime identity, we have

$$\begin{aligned}\max_{y\in V} \nu(x)\mathbb{E}[\widetilde{\tau}_y^x] &\leqslant \max_{y\in V} \sum_{x\in V} \nu(x)\mathbb{E}[\widetilde{\tau}_y^x]\\ &= \widetilde{\Sigma}_*\end{aligned}$$

and it remains to take into account (11) to deduce the desired bound.

■

2.2 A refined estimate

To prove Theorem 5, we consider an extension $\widehat{V} \coloneqq V \sqcup \{\bar{\omega}\}$, where $\bar{\omega} \notin V$, endowed with the irreducible generator $\widehat{L} \coloneqq (\widehat{L}(x,y))_{x,y\in V}$ defined by

$$\forall\, x \neq y \in \widehat{V}, \qquad \widehat{L}(x,y) \;\coloneqq\; \begin{cases} \widetilde{L}(x,y) & \text{, if } x,y \in V\\ a & \text{, if } x = \omega \text{ and } y = \bar{\omega}\\ \widetilde{\pi}(\omega) & \text{, if } x = \bar{\omega} \text{ and } y = \omega\\ 0 & \text{, otherwise}\end{cases} \tag{12}$$

The invariant measure $\widehat{\pi}$ associated to $\widehat{L}$ is given by

Lemma 7 *We have*

$$\widehat{\pi} \;=\; \frac{\widetilde{\pi} + a\delta_{\bar{\omega}}}{1+a}$$

Proof

Denote $\mu \coloneqq \widetilde{\pi} + a\delta_{\bar{\omega}}$, we have to check that $\mu\widehat{L} = 0$. We consider three cases.

- For $x \in V\backslash\{\omega\}$, we have

$$\begin{aligned}\mu\widehat{L}(x) &= \sum_{y\in\widehat{V}} \mu(y)\widehat{L}(y,x)\\ &= \sum_{y\in V} \mu(y)\widetilde{L}(y,x)\\ &= \sum_{y\in V} \widetilde{\pi}(y)\widetilde{L}(y,x)\\ &= 0\end{aligned}$$

- For ω, we have

$$\begin{aligned}\mu\widehat{L}(\omega) &= \mu(\widehat{\omega})\widehat{L}(\bar{\omega},\omega) + \mu(\omega)\widehat{L}(\omega,\omega) + \sum_{y\in V\backslash\{\omega\}} \mu(y)\widetilde{L}(y,\omega)\\ &= a\widetilde{\pi}(\omega) + \widetilde{\pi}(\omega)(\widetilde{L}(\omega,\omega) - \widehat{L}(\omega,\bar{\omega})) + \sum_{y\in V\backslash\{\omega\}} \widetilde{\pi}(y)\widetilde{L}(y,\omega)\\ &= a\widetilde{\pi}(\omega) - \widetilde{\pi}(\omega)a + \sum_{y\in V} \mu(y)\widetilde{L}(y,\omega)\\ &= 0\end{aligned}$$

- For $\bar{\omega}$, we have

$$\begin{aligned} \mu\widehat{L}(\bar{\omega}) &= \mu(\omega)\widehat{L}(\omega,\bar{\omega}) + \mu(\bar{\omega})\widehat{L}(\bar{\omega},\bar{\omega}) \\ &= \widetilde{\pi}(\omega)a - a\widetilde{\pi}(\omega) \\ &= 0 \end{aligned}$$

■

Consider $\widehat{\varphi}$ the unique function solution of the Poisson equation

$$\left\{ \begin{array}{lcl} \widehat{L}[\widehat{\varphi}] &=& \mathbb{1}_{\{\bar{\omega}\}} - \widehat{\pi}(\bar{\omega}) \\ \widehat{\pi}[\widehat{\varphi}] &=& 0 \end{array} \right. \tag{13}$$

Since $\bar{\omega}$ is only in relation with ω, it is possible to make a direct link between φ and $\widehat{\varphi}$.

Lemma 8 *On V, we have*

$$\varphi = \left(1+\frac{1}{a}\right)\widetilde{\pi}(\omega)(\psi - \widetilde{\pi}[\psi])$$

where ψ is the restriction of $\widehat{\varphi}$ to V.

Proof

Let us compute $\widetilde{L}[\psi]$. We consider two cases.

- For $x \in V\backslash\{\omega\}$, we have

$$\begin{aligned} \widetilde{L}[\psi](x) &= \widehat{L}[\widehat{\varphi}](x) \\ &= -\widehat{\pi}(\bar{\omega}) \\ &= -\frac{a}{1+a} \end{aligned}$$

- For ω, we have

$$\begin{aligned} \widetilde{L}[\psi](\omega) &= \widehat{L}[\widehat{\varphi}](\omega) - \widehat{L}(\omega,\bar{\omega})(\widehat{\varphi}(\bar{\omega}) - \widehat{\varphi}(\omega)) \\ &= -\widehat{\pi}(\bar{\omega}) - \widehat{L}(\omega,\bar{\omega})(\widehat{\varphi}(\bar{\omega}) - \widehat{\varphi}(\omega)) \end{aligned}$$

To evaluate the last quantity, note that

$$\begin{aligned} 1 - \widehat{\pi}(\bar{\omega}) &= \widehat{L}[\widehat{\varphi}](\bar{\omega}) \\ &= \widehat{L}(\bar{\omega},\omega)(\widehat{\varphi}(\omega) - \widehat{\varphi}(\bar{\omega})) \end{aligned}$$

so that

$$\begin{aligned} \widetilde{L}[\psi](\omega) &= -\widehat{\pi}(\bar{\omega}) + \frac{\widehat{L}(\omega,\bar{\omega})}{\widehat{L}(\bar{\omega},\omega)}(1 - \widehat{\pi}(\bar{\omega})) \\ &= -\frac{a}{1+a} + \frac{a}{\widetilde{\pi}(\omega)}\frac{1}{1+a} \\ &= \frac{a}{(1+a)\widetilde{\pi}(\omega)} - \frac{a}{1+a} \end{aligned}$$

Thus we have

$$\begin{aligned} \widetilde{L}[\psi] &= \frac{a}{(1+a)\widetilde{\pi}(\omega)}\mathbb{1}_{\{\omega\}} - \frac{a}{1+a} \\ &= \frac{a}{(1+a)\widetilde{\pi}(\omega)}(\mathbb{1}_{\{\omega\}} - \widetilde{\pi}(\omega)) \end{aligned}$$

It means that $\widetilde{\pi}(\omega)\frac{1+a}{a}(\psi - \widetilde{\pi}[\psi])$ is solution to the Poisson equation (7), which amounts to the announced result.

■

Consider $\widehat{\Lambda}$ the multi-set consisting of the spectrum of $-\widehat{L}$ with its algebraic multiplicities. By irreducibility of $\widehat{L}$, $0 \in \widehat{\Lambda}$ with multiplicity 1. Denote $\widehat{\Lambda}_* := \widehat{\Lambda}\backslash\{0\}$ and

$$\widehat{\Sigma}_* \quad := \quad \sum_{\lambda \in \widehat{\Lambda}_*} \frac{1}{\lambda}$$

Applying Proposition 6 to $\widehat{\varphi}$ the solution of the Poisson equation (13), we get that

$$\|\widehat{\varphi}\|_\infty \quad \leqslant \quad \widehat{\Sigma}_* \tag{14}$$

From Lemma 8, we deduce that

$$\|\varphi\|_\infty \quad \leqslant \quad 2\left(1+\frac{1}{a}\right)\widetilde{\pi}(\omega)\widehat{\Sigma}_* \tag{15}$$

Theorem 5 will be a consequence of

Proposition 9 *Assume that all the eigenvalues of $L^\dagger$ are of (algebraic) multiplicity 1. Then we have*

$$\lim_{a\to+\infty} \widehat{\Sigma}_* \quad = \quad \Sigma^\dagger$$

In fact we think this convergence holds without the assumption that the eigenvalues of $L^\dagger$ are of (algebraic) multiplicity 1. The proof of Theorem 5 would then be immediate. Nevertheless the proof of Proposition 9 without its multiplicity assumption requires more care than is really necessary for our purpose. Before proving Proposition 9, let us deduce Theorem 5 in general:

Proof of Theorem 5

Let $\mathcal{I}$ be the (convex) set of all irreducible generators on V and $\mathcal{I}_0$ be the subset of $K \in \mathcal{I}$ such that all the eigenvalues of $K^\dagger := (K(x,y))_{x,y\in V^\dagger}$ are distinct. Let us check that $\mathcal{I}_0$ is dense in $\mathcal{I}$. Fix some $K \in \mathcal{I}$ and $\epsilon > 0$. Consider $\mathcal{B}$ the set of matrices $\widetilde{K} := (\widetilde{K}(x,y))_{x,y\in V}$ such that

$$\begin{aligned}
\forall\ x \neq y \in V, \qquad & K(x,y) \ < \ \widetilde{K}(x,y) \ < \ K(x,y)+\epsilon \\
\forall\ x \in V, \qquad & \widetilde{K}(x,x) \ = \ -\sum_{y\in V\backslash\{x\}} K(x,y)
\end{aligned}$$

Clearly, $\mathcal{B} \subset \mathcal{I}$ and to obtain the desired density, it is sufficient to show that $(\mathcal{B} \cap \mathcal{I}_0)^\dagger \neq \emptyset$, where $(\mathcal{B} \cap \mathcal{I}_0)^\dagger$ is the image of $\mathcal{B} \cap \mathcal{I}_0$ by the mapping $\mathcal{I} \ni K \mapsto K^\dagger$. Note, on one hand, that $(\mathcal{B} \cap \mathcal{I}_0)^\dagger = \mathcal{B}^\dagger \cap \mathcal{J}$, where $\mathcal{J}$ is the set of $V^\dagger \times V^\dagger$-matrices whose eigenvalues are distinct, and on the other hand, that $\mathcal{B}^\dagger$ is an open subset in the set of $V^\dagger \times V^\dagger$-matrices. It is then well-known that $\mathcal{J}$ is dense in the set of all $V^\dagger \times V^\dagger$-matrices, this ends the proof of the density of $\mathcal{I}_0$ in $\mathcal{I}$.

Let $\widetilde{L} \in \mathcal{I}$ be fixed as in Theorem 5 and consider $(\widetilde{L}^{(n)})_{n\in\mathbb{N}}$ be a sequence of elements of $\mathcal{I}_0$ converging toward $\widetilde{L}$. We denote by $(\widetilde{\pi}^{(n)})_{n\in\mathbb{N}}$ and $(\varphi^{(n)})_{n\in\mathbb{N}}$ the corresponding sequences of invariant probability measures and solutions to the Poisson equation (7). Resorting, for all $n \in \mathbb{N}$, to the explicit tree description of $\widetilde{\pi}^{(n)}$ in terms of $\widetilde{L}^{(n)}$ (see e.g. Lemma 3.1 of Chapter 6 of Freidlin and Wentzell [9]), we get

$$\lim_{n\to\infty} \widetilde{\pi}^{(n)} \quad = \quad \widetilde{\pi}$$

Taking into account the uniqueness of the solution of (7), it easily follows that

$$\lim_{n\to\infty} \varphi^{(n)} \quad = \quad \varphi$$

and in particular

$$\lim_{n\to\infty} \left\|\varphi^{(n)}\right\|_\infty = \|\varphi\|_\infty$$

From (15) and Proposition 9, we have for any $n \in \mathbb{N}$,

$$\left\|\varphi^{(n)}\right\|_\infty \leqslant 2\widetilde{\pi}^{(n)}(\omega)\Sigma^{(n)\dagger} \tag{16}$$

where $\Sigma^{(n)\dagger}$ is the trace of the inverse of the matrix $-\widetilde{L}^{(n)\dagger}$. Taking the inverse of a matrix is a continuous operation (among invertible matrices), so we deduce that

$$\lim_{n\to\infty} \Sigma^{(n)\dagger} = \Sigma^\dagger$$

Finally passing to the limit in (16), we get the desired bound.

■

The phenomenon behind the convergence of Proposition 9 is that for $a > 0$ large, $|V|-1$ eigenvalues from $\widehat{\Lambda}_*$ converge toward the eigenvalues of $\Lambda^\dagger$ and the remaining eigenvalue from $\widehat{\Lambda}_*$ diverges toward $+\infty$.

First, let us give a non-linear characterization of the spectrum of $\widehat{\Lambda}_*$.

Lemma 10 *A complex number $z \in \mathbb{C}\backslash\{\widetilde{\pi}(\omega)\}$ is an eigenvalue of $-\widehat{L}$ if and only if there exists a function $f \neq 0$ on V such that*

$$\widetilde{L}[f] = -zf - \frac{az}{\widetilde{\pi}(\omega) - z} f(\omega)\mathbb{1}_{\{\omega\}} \tag{17}$$

The number $\widetilde{\pi}(\omega)$ is an eigenvalue of $-\widehat{L}$ if and only if it is also an eigenvalue of $-L^\dagger$.

Proof

Consider an eigenvalue $\widehat{\lambda} \in \mathbb{C}\backslash\{\widetilde{\pi}(\omega)\}$ of $-\widehat{L}$ and $\widehat{f}$ a corresponding eigenfunction. Denote f the restriction of $\widehat{f}$ to V. Applying the relation $\widehat{L}[\widehat{f}] = -\widehat{\lambda}\widehat{f}$ at $\bar{\omega}$, we get

$$\widetilde{\pi}(\omega)(f(\omega) - \widehat{f}(\bar{\omega})) = -\widehat{\lambda}\widehat{f}(\bar{\omega}) \tag{18}$$

namely

$$f(\omega) = \left(1 - \frac{\widehat{\lambda}}{\widetilde{\pi}(\omega)}\right)\widehat{f}(\bar{\omega}) \tag{19}$$

and since $\widehat{\lambda} \neq \widetilde{\pi}(\omega)$, we deduce that

$$f(\omega) - \widehat{f}(\bar{\omega}) = -\frac{\widehat{\lambda}}{\widetilde{\pi}(\omega) - \widehat{\lambda}} f(\omega) \tag{20}$$

From

$$\widehat{L}\widehat{f}(\omega) = -\widehat{\lambda}\widehat{f}(\omega)$$

we deduce

$$\widetilde{L}[f](\omega) + a(\widehat{f}(\bar{\omega}) - f(\omega)) = -\widehat{\lambda}f(\omega)$$

i.e., taking into account (20),

$$\widetilde{L}[f](\omega) \quad = \quad -\widehat{\lambda} f(\omega)-\frac{a\widehat{\lambda}}{\widetilde{\pi}(\omega)-\widehat{\lambda}}f(\omega)$$

For $x\in V^{\dagger}$, we have $\widetilde{L}[f](x)=\widehat{L}[\widehat{f}](x)=-\widehat{\lambda} f(x)$, so that (17) is satisfied on V. Note that if $f=0$, then from (19) we would get $\widehat{f}(\bar{\omega})=0$ (recall that $\widehat{\lambda}\neq\widetilde{\pi}(\omega)$) and by consequence $\widehat{f}=0$, which is not allowed.

Conversely, consider $z\in\mathbb{C}\backslash\{\widetilde{\pi}(\omega)\}$ and a function $f\neq 0$ on V such that (17) is satisfied. Defining $\widehat{f}$ via

$$\forall\, x\in\widehat{V},\qquad \widehat{f}(x) \;\coloneqq\; \begin{cases} f(x) & \text{, if } x\in V\\ \frac{\widetilde{\pi}(\omega)f(\omega)}{\widetilde{\pi}(\omega)-z} & \text{, if } x=\bar{\omega}\end{cases}$$

and reversing the above computations, we get that $\widehat{f}$ is an eigenvector of $\widehat{L}$ associated to the eigenvalue $-z$.

Next assume that $\widetilde{\pi}(\omega)$ is an eigenvalue of $-\widehat{L}$, let $\widehat{f}$ be an associated eigenvalue and denote f the restriction of $\widehat{f}$ to V. From (18), we deduce that $f(\omega)=0$. Furthermore, we have for $x\in V^{\dagger}$, $L^{\dagger}[f](x)=\widetilde{L}[f](x)=\widehat{L}[\widehat{f}](x)=-\widetilde{\pi}(\omega)f(x)$. It follows that f is an eigenfunction of $L^{\dagger}$ associated to the eigenvalue $-\widetilde{\pi}(\omega)$. Conversely, if $\widetilde{\pi}(\omega)$ is an eigenvalue of $-L^{\dagger}$ with associated eigenvector f, it is sufficient to consider the function $\widehat{f}$ defined by

$$\forall\, x\in\widehat{V},\qquad \widehat{f}(x) \;\coloneqq\; \begin{cases} f(x) & \text{, if } x\in V^{\dagger}\\ 0 & \text{, if } x\in\{\omega,\bar{\omega}\}\end{cases}$$

to get that $\widehat{L}[\widehat{f}]=-\widetilde{\pi}(\omega)\widehat{f}$.

■

There is probably an extension of Lemma 10 concerning the Jordan blocs of $\widehat{L}$, but such a result will not be useful for us, due to the multiplicity assumption in Proposition 9. Under this hypothesis, we will see below that for $a>0$ large enough, all the eigenvalues of $\widehat{L}$ are distinct. The following result is the crucial step in this direction.

Lemma 11 *Consider $\eta>0$ and $\lambda\neq 0$ an eigenvalue of $-L^{\dagger}$. Under the assumption of Proposition 9, there exists $A>0$ large enough such that for all $a>A$, there exists an eigenvalue of $-\widehat{L}$ in the complex disk of center λ and radius η.*

Proof

If $\lambda=\widetilde{\pi}(\omega)$, according to Lemma 10, λ is also an eigenvalue of $-\widehat{L}$ for all $a>0$. From now on, assume that $\lambda\neq\widetilde{\pi}(\omega)$. There is another situation where the result is obvious. Denote μ the (non-negative) measure on $V^{\dagger}$ given by $(\widetilde{L}(\omega,x))_{x\in V^{\dagger}}$. Let ξ be an eigenvector of $-L^{\dagger}$ associated to λ. If $\mu[\xi]=0$, then λ is also an eigenvalue of $-\widehat{L}$ for all $a>0$. Indeed, note that (17) applied at ω amounts to

$$\mu[f]+\widetilde{L}(\omega,\omega)f(\omega) \quad = \quad -\left(z+\frac{az}{\widetilde{\pi}(\omega)-z}\right)f(\omega) \tag{21}$$

(recall that $\widetilde{L}(\omega,\omega)=-\sum_{x\in V^{\dagger}}\widetilde{L}(\omega,x)$).

Thus considering f defined by

$$\forall\, x\in V,\qquad f(x) \quad = \quad \begin{cases} \xi(x) & \text{, if } x\in V^{\dagger}\\ 0 & \text{, if } x=\omega\end{cases}$$

we get that (21) is satisfied.

Since $f(\omega) = 0$, (17) is just asking for $L^\dagger[f](x) = -zf(x)$ for $x \in V^\dagger$, and this is true with $z = \lambda$.

Let us now consider the situation where $\mu[\xi] \neq 0$. Up to normalizing ξ, we furthermore assume that $\mu[\xi] = 1$. We are looking for a solution (z, f) of (17) equally normalized by $\mu[f] = 1$.

Let us change the notations, defining $\epsilon \coloneqq 1/a$, $r \coloneqq af(\omega)$ and $g \coloneqq (g(x))_{x\in V^\dagger} \coloneqq (f(x))_{x\in V^\dagger}$. The condition $\mu[f] = 1$ translates into $\mu[g] = 1$ and (17) with $\mu[f] = 1$ is equivalent to the system

$$\left\{ \begin{array}{rcll} L^\dagger[g](x) + \epsilon\widetilde{L}(x,\omega)r + zg(x) &=& 0, & \forall\, x \in V^\dagger \\ 1 + \left(\frac{z}{\widetilde{\pi}(\omega)-z} + \epsilon(z + \widetilde{L}(\omega,\omega))\right) r &=& 0 & \\ \mu[g] &=& 1 & \end{array} \right. \tag{22}$$

Consider

$$\mathcal{D} \quad \coloneqq \quad \left\{(\epsilon, z, g) \in [0, +\infty) \times (\mathbb{C}\backslash\{\widetilde{\pi}(\omega)\}) \times \mathbb{R}^{V^\dagger} \; : \; \frac{z}{\widetilde{\pi}(\omega) - z} + \epsilon(z + \widetilde{L}(\omega,\omega)) \neq 0\right\}$$

and define the mapping $F \coloneqq (F_x(\epsilon, z, g))_{x\in V} : \mathcal{D} \to \mathbb{R}^V$ via

$$\forall\, x \in V,\ \forall\, (\epsilon, z, g) \in \mathcal{D},$$
$$F_x(\epsilon, z, g) \quad \coloneqq \quad \left\{ \begin{array}{ll} L^\dagger[g](x) - \epsilon\widetilde{L}(x,\omega)\frac{\widetilde{\pi}(\omega)-z}{z+\epsilon(z+\widetilde{L}(\omega,\omega))(\widetilde{\pi}(\omega)-z)} + zg(x) & \text{, when } x \in V^\dagger \\ \mu[g] & \text{, when } x = \omega \end{array} \right.$$

With this notation, the system (22) can written

$$F(\epsilon, z, g) \quad = \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

where the 1 corresponds to the ω coordinate (and 0 is the null vector in $\mathbb{R}^{V^\dagger}$).

Note that

$$F(0, \lambda, \xi) \quad = \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

thus the implicit function theorem enables us to deduce the desired theorem as soon as we will have shown that the Jacobian matrix $\nabla F \coloneqq (\partial_z F, (\nabla_{g(x)} F)_{x\in V^\dagger})$ is non degenerate at the point $(0, \lambda, \xi)$. We compute that

$$\begin{array}{rrcl} & \partial_z F(0, \lambda, \xi) &=& \begin{pmatrix} \xi \\ 0 \end{pmatrix} \\ \forall\, x \in V^\dagger, & \nabla_{g(x)} F(0, \lambda, \xi) &=& \begin{pmatrix} L^\dagger(\cdot, x) + \lambda\delta_x \\ \mu(x) \end{pmatrix} \end{array}$$

To check that $\nabla F(0, \lambda, \xi)$ is invertible, consider $(s, h) \in \mathbb{R} \times \mathbb{R}^{V^\dagger}$ such that

$$\nabla F(0, \lambda, \xi) \cdot (s, h) \quad = \quad 0$$

According to the above computations, this equation can be written under the following system:

$$\left\{ \begin{array}{rcll} L^\dagger[h](x) + \lambda h(x) + s\xi(x) &=& 0, & \forall\, x \in V^\dagger \\ \mu[h] &=& 0 & \end{array} \right. \tag{23}$$

Under the assumption of Proposition 9, the equation

$$L^\dagger[h] + \lambda h \quad = \quad -s\xi$$

implies that h belongs to the vector space generated by ξ. To see it, just decompose h into a basis of $\mathbb{R}^{V^\dagger}$ consisting of eigenvectors of $L^\dagger$ and take into account that the multiplicity of $-\lambda$ is one. It follows that $L^\dagger[h]+\lambda h=0$ and thus $s=0$. Let $b\in\mathbb{R}$ be such that $h=b\xi$. We deduce that $\mu[h]=b\mu[\xi]=b$, so the second equation of (23) implies that $b=0$ and finally $h=0$. Thus we have $(s,h)=(0,0)$ and $\nabla F(0,\lambda,\xi)$ is non degenerate, as desired.

■

Remark 12 It is the nonlinearity of (17) that leads to the above technical arguments. Had a traditional linear eigenproblem been considered, we could have directly resorted to the results of Kato [12]. Note nevertheless that for $\epsilon=0$, ∇F has the same form as if we had been treating a usual linear eigenproblem.

□

We can now come to the

Proof of Proposition 9

Define

$$\epsilon \quad\coloneqq\quad \min\{|\lambda|\wedge|\lambda-\lambda'|\,:\,\lambda\neq\lambda'\in\Lambda^\dagger\}$$

which is a positive quantity according to the assumption on the multiplicity of the elements of $\Lambda^\dagger$ and to the fact that $0\notin\Lambda^\dagger$. Considering $\eta\coloneqq\epsilon/2$ in Lemma 11, we deduce that there exists $A>0$ such that for any $a>A$ and any $\lambda\in\Lambda^\dagger$, there exists an eigenvalue of $-\widehat{L}$ in the disk centered at λ of radius η. By definition of η, this eigenvalue is not null and all these eigenvalues are distinct for different $\lambda\in\Lambda^\dagger$. This gives us $\mathrm{card}(V^\dagger)=\mathrm{card}(V)-1$ distinct elements from $\widehat{\Lambda}_*$. To see that the missing element is going to infinity as a goes to infinity, it sufficient to consider the trace of $-\widehat{L}$, which is equal to

$$\mathrm{tr}(-\widehat{L}) \quad=\quad a+\widetilde{\pi}(\omega)+\mathrm{tr}(-\widetilde{L}) \quad=\quad a+\widetilde{\pi}(\omega)-\sum_{x\in V}\widetilde{L}(x,x)$$

These observations imply the convergence stated in Proposition 9, as well as the fact that for $a>0$ large enough, all the eigenvalues of $\widehat{L}$ are distinct.

■

3 Exit time

Our main goal here is to prove Theorem 1 via manipulations of Poisson equations and taking into account the estimate of Theorem 5.

Instead of working with the vanishing X_x, for $x\in V$, it is often more convenient to resort to conservative Markov processes, obtained by adding a cemetery point to the state space. So let be given $\infty\notin V$ and associate to it an ergodic Markov generator $\breve{L}\coloneqq(\breve{L}(x,y))_{x,y\in\breve{V}}$ on $\breve{V}\coloneqq V\sqcup\{\infty\}$ via

$$\forall\ x\neq y\in\breve{V},\qquad \breve{L}(x,y)\quad\coloneqq\quad\begin{cases} L(x,y) & \text{, if } x,y\in V\\ -\Big(L(x,x)+\sum_{y\in V\setminus\{x\}}L(x,y)\Big) & \text{, if } x\in V \text{ and } y=\infty\\ a\nu(y) & \text{, if } x=\infty\end{cases}$$

where $a>0$ is fixed for the moment being. This "extension" of the absorbed Markov generator L into an ergodic Markov generator $\breve{L}$ is completely different to the passage from $\widetilde{L}$ to $\widehat{L}$ in the previous section. In some sense, the former is *global* while the latter was *local*.

For $x\in V$, let $\breve{X}_x\coloneqq(\breve{X}_x(t))_{t\geqslant 0}$ be a Markov process starting from x and whose generator is $\breve{L}$, and consider the absorption time

$$\breve{\tau}_x\quad\coloneqq\quad\inf\{t\geqslant 0\,:\,\breve{X}_x(t)=\infty\}$$

Note that the stopped processes $(X_x(t \wedge \tau^x))_{t \geqslant 0}$ and $(\breve{X}_x(t \wedge \breve{\tau}^x))_{t \geqslant 0}$ have the same law and in particular τ_x and $\breve{\tau}_x$ have the same distribution.

The interest of L over $\breve{L}$ is that we can consider $\breve{\psi}$ the function on $\breve{V}$ solution of the Poisson equation

$$\begin{cases} \breve{L}[\breve{\psi}] & = & \mathbb{1}_{\{\infty\}} - \breve{\pi}(\infty) \\ \breve{\pi}[\breve{\psi}] & = & 0 \end{cases}$$

where $\breve{\pi}$ is the invariant probability of $\breve{L}$ and $\mathbb{1}_{\{\infty\}}$ is the indicator function of ∞.

Let us apply to $\breve{\psi}$ the martingale problem associated with $\breve{L}$. We have for any $x \in \breve{V}$ and $t \geqslant 0$,

$$\begin{aligned} \breve{\psi}(\breve{X}_x(t)) &= \breve{\psi}(\breve{X}_x(0)) + \int_0^t \breve{L}[\breve{\psi}](\breve{X}_x(s))\, ds + \breve{M}_t \\ &= \breve{\psi}(x) + \int_0^t \mathbb{1}_{\{\infty\}}(\breve{X}_x(s)) - \breve{\pi}(\infty)\, ds + \breve{M}_t \end{aligned}$$

where $(\breve{M}_t)_{t \geqslant 0}$ is a martingale. Replace t by $t \wedge \breve{\tau}_x$, to get

$$\begin{aligned} \breve{\psi}(\breve{X}_x(t \wedge \breve{\tau}_x)) &= \breve{\psi}(x) + \int_0^{t \wedge \breve{\tau}_x} \mathbb{1}_{\{\infty\}}(\breve{X}_x(s)) - \breve{\pi}(\infty)\, ds + \breve{M}_{t \wedge \breve{\tau}_x} \\ &= \breve{\psi}(x) - \breve{\pi}(\infty)(t \wedge \breve{\tau}_x) + \breve{M}_{t \wedge \breve{\tau}_x} \end{aligned}$$

Taking expectations, we obtain

$$\mathbb{E}[\breve{\psi}(\breve{X}_x(t \wedge \breve{\tau}_x))] = \breve{\psi}(x) - \breve{\pi}(\infty)\mathbb{E}[t \wedge \breve{\tau}_x] \tag{24}$$

Before going further, let us explain heuristically how (24) can be exploited. The underlying principle is that under appropriate conditions, $\breve{\psi}$ is close to $-(1 - \breve{\pi}(\infty))a^{-1}\mathbb{1}_{\{\infty\}}$. So if carelessly we replace $\breve{\psi}$ by $-(1 - \breve{\pi}(\infty))a^{-1}\mathbb{1}_{\{\infty\}}$ in (24), we get for any $x \in V$,

$$\forall\ t \geqslant 0, \qquad \mathbb{E}[\mathbb{1}_{\{\infty\}}(\breve{X}_x(t \wedge \breve{\tau}_x))] \approx \mathbb{1}_{\{\infty\}}(x) + \frac{a\breve{\pi}(\infty)}{1 - \breve{\pi}(\infty)}\mathbb{E}[t \wedge \breve{\tau}_x]$$

namely

$$\forall\ t \geqslant 0, \qquad \mathbb{P}[\breve{\tau}_x \leqslant t] \approx \frac{a\breve{\pi}(\infty)}{1 - \breve{\pi}(\infty)}\mathbb{E}[t \wedge \breve{\tau}_x] \tag{25}$$

A true identity in (25) would imply that $\breve{\tau}_x$ is a exponential random variable of parameter $a\breve{\pi}(\infty)/(1-\breve{\pi}(\infty))$ (see e.g. [15]). These approximative considerations also suggest an identification of $a\breve{\pi}(\infty)/(1-\breve{\pi}(\infty))$ with λ_0. Indeed, consider $\breve{X} \coloneqq (\breve{X}(t))_{t \geqslant 0}$ a Markov process whose initial law is ν and whose generator is $\breve{L}$. One property of the quasi-stationary distribution ν is that the first hitting time $\breve{\tau}$ of ∞ by $\breve{X}$ is distributed as an exponential random variable of parameter λ_0 and thus

$$\forall\ t \geqslant 0, \qquad \mathbb{P}[\breve{\tau} \leqslant t] = \lambda_0 \mathbb{E}[t \wedge \breve{\tau}]$$

Comparing with (25), which is also "valid" when $\breve{X}_x$ is replaced by $\breve{X}$, we get $a\breve{\pi}(\infty)/(1-\breve{\pi}(\infty)) \approx \lambda_0$. It would follow that $\breve{\tau}_x$ is almost an exponential random variable of parameter λ_0 for all $x \in V$.

We now come to more rigorous computations. We begin by computing $\breve{\pi}$ in terms of λ_0 and ν:

Lemma 13 *We have*

$$\begin{aligned} \breve{\pi} &= \frac{\nu + (\lambda_0/a)\delta_\infty}{1 + \lambda_0/a} \\ \sum_{x \in V} \nu(x)\breve{L}(x, \infty) &= \lambda_0 \end{aligned}$$

Proof

It is sufficient to show that $\breve{\pi}\breve{L} = 0$, where $\breve{\pi} \coloneqq \nu + (\lambda_0/a)\delta_\infty$. We begin by showing that

$$\nu\breve{L} \quad = \quad -\lambda_0\nu + \lambda_0\delta_\infty \tag{26}$$

where ν is extended as a probability on $\breve{V}$ giving the mass 0 to ∞. Indeed, note that for $x \in V$, we have

$$\begin{aligned} \nu\breve{L}(x) &= \sum_{y\in V} \nu(y)\breve{L}(y,x) \\ &= \sum_{y\in V} \nu(y)L(y,x) \\ &= \nu L(x) \\ &= -\lambda_0\nu(x) \end{aligned}$$

It follows that there exists a number $\alpha \in \mathbb{R}$ such that $\nu\breve{L} = -\lambda_0\nu + \alpha\delta_\infty$. To compute α, note that $\breve{L}[\mathbb{1}_{\breve{V}}] = 0$, so that $\nu[\breve{L}[\mathbb{1}_{\breve{V}}]] = 0$ and $\alpha = \lambda_0$, proving (26).

As a consequence, we get that

- for $y \neq \infty$,

$$\begin{aligned} \sum_{x\in \breve{V}} \breve{\pi}(x)\breve{L}(x,y) &= (\lambda_0/a)\breve{L}(\infty,y) + \sum_{x\in V} \nu(x)\breve{L}(x,y) \\ &= \lambda_0\nu(y) + (\nu\breve{L})(y) \\ &= \lambda_0\nu(y) - \lambda_0\nu(y) \\ &= 0 \end{aligned}$$

- for ∞,

$$\begin{aligned} \sum_{x\in \breve{V}} \breve{\pi}(x)\breve{L}(x,\infty) &= (\lambda_0/a)\breve{L}(\infty,\infty) + \sum_{x\in V} \nu(x)\breve{L}(x,\infty) \\ &= -\lambda_0 + (\nu\breve{L})(\infty) \\ &= -\lambda_0 + \lambda_0 \\ &= 0 \end{aligned}$$

The previous computation also shows the last equality of the above lemma.

■

As suggested by the heuristic presented before Lemma 13, the function

$$\breve{\phi} \quad \coloneqq \quad \breve{\psi} + \frac{1-\breve{\pi}(\infty)}{a}\mathbb{1}_{\{\infty\}}$$

should play an important role. Note that according to Lemma 13, we have

$$a\frac{\breve{\pi}(\infty)}{1-\breve{\pi}(\infty)} \quad = \quad a\frac{\lambda_0}{a} \quad = \quad \lambda_0$$

in concordance with the "arguments" preceding Lemma 13.

Lemma 14 *We have*

$$\breve{L}[\breve{\phi}] \quad = \quad \frac{(1-\breve{\pi}(\infty))}{a}\sum_{x\in V}\breve{L}(x,\infty)\mathbb{1}_{\{x\}} - \breve{\pi}(\infty)\mathbb{1}_V$$

Proof

Note that

$$\breve{L}[\mathbb{1}_{\{\infty\}}] \;=\; \sum_{x\in V}\breve{L}(x,\infty)\mathbb{1}_{\{x\}} - a\mathbb{1}_{\{\infty\}}$$

Indeed, we compute that

$$\begin{aligned}\breve{L}[\mathbb{1}_{\{\infty\}}](\infty) &= \breve{L}(\infty,\infty)\\ &= -\sum_{y\in V}\breve{L}(\infty,y)\\ &= -a\sum_{y\in V}\nu(y)\\ &= -a\end{aligned}$$

and we clearly have for any $x\in V$, $\breve{L}[\mathbb{1}_{\{\infty\}}](x) = \breve{L}(x,\infty)$.

Taking into account that by definition

$$\breve{L}[\breve{\psi}] \;=\; (1-\breve{\pi}(\infty))\mathbb{1}_{\{\infty\}} - \breve{\pi}(\infty)\mathbb{1}_V$$

we deduce that

$$\begin{aligned}\breve{L}[\breve{\phi}] &= \breve{L}[\breve{\psi}] + \frac{1-\breve{\pi}(\infty)}{a}\breve{L}[\mathbb{1}_{\{\infty\}}]\\ &= \frac{(1-\breve{\pi}(\infty))}{a}\sum_{x\in V}\breve{L}(x,\infty)\mathbb{1}_{\{x\}} - \breve{\pi}(\infty)\mathbb{1}_V\end{aligned}$$

■

It follows that $\breve{L}[\breve{\phi}](\infty)=0$, namely

$$\breve{\phi}(\infty) \;=\; \nu[\breve{\phi}] \tag{27}$$

This observation leads us to introduce a new generator $\widetilde{L}$ on V. Denote $\mathcal{F}(V)$ the set of real functions defined on V. Any $f\in\mathcal{F}(V)$ is extended into a function $\widetilde{f}$ on $\breve{V}$ by imposing

$$\widetilde{f}(\infty) \;:=\; \nu[f]$$

We consider the generator $\widetilde{L}$ given by

$$\forall\ f\in\mathcal{F}(V),\ \forall\ x\in V,\qquad \widetilde{L}[f](x) \;:=\; \breve{L}[\widetilde{f}](x)$$

The generator $\widetilde{L}$ is the Steklov operator associated to $\breve{L}$ and to the "boundary" V of $\breve{V}$, since $\widetilde{f}$ can be seen as the "harmonic extension" of f (for more details about this point of view, see [10]). It follows that the invariant probability measure of $\widetilde{L}$ is the normalization of the restriction of $\breve{\pi}$ to V, namely ν. More explicitly, $\widetilde{L}$ is described by (2).

Denote ϕ the restriction of $\breve{\phi}$ on V. Due to (27), $\widetilde{\phi}$ coincides with $\breve{\phi}$, so by definition of $\widetilde{L}$, we get on V:

$$\widetilde{L}[\phi] \;=\; \frac{(1-\breve{\pi}(\infty))}{a}\sum_{x\in V}\breve{L}(x,\infty)\mathbb{1}_{\{x\}} - \breve{\pi}(\infty)\mathbb{1}_V$$

or equivalently

$$\widetilde{L}[\phi] \;=\; \frac{1}{a+\lambda_0}\sum_{x\in V}\breve{L}(x,\infty)(\mathbb{1}_{\{x\}} - \nu(x)) \tag{28}$$

As in (8), for any $x \in V$, consider the solution φ^x of the Poisson equation

$$\left\{ \begin{array}{rcl} \widetilde{L}[\varphi^x] & = & \mathbb{1}_{\{x\}} - \nu(x) \\ \nu[\varphi^x] & = & 0 \end{array} \right. \tag{29}$$

so that (28) implies that

$$\phi \quad = \quad \nu[\phi] + \frac{1}{a+\lambda_0} \sum_{x \in V} \breve{L}(x,\infty)\varphi^x \tag{30}$$

Indeed, we get

$$\widetilde{L}\left[\phi - \frac{1}{a+\lambda_0} \sum_{x \in V} \breve{L}(x,\infty)\varphi^x\right] \quad = \quad 0$$

thus by irreducibility of $\widetilde{L}$, ϕ and $\frac{1}{a+\lambda_0}\sum_{x\in V} \breve{L}(x,\infty)\varphi^x$ coincide up to an additive constant, which is necessarily $\nu[\phi]$.

The next result shows that ϕ can be completely expressed in terms of $(\varphi^x)_{x\in V}$.

Lemma 15 *We have*

$$\phi \quad = \quad \frac{\lambda_0}{(a+\lambda_0)^2} + \frac{1}{a+\lambda_0} \sum_{x \in V} \breve{L}(x,\infty)\varphi^x$$

Proof

It follows from Lemma 13 and (27) that

$$\begin{array}{rcl} \breve{\pi}[\breve{\phi}] & = & \dfrac{a\nu[\phi] + \lambda_0\breve{\phi}(\infty)}{a+\lambda_0} \\ & = & \nu[\phi] \end{array}$$

By definition of $\breve{\phi}$, we also have

$$\begin{array}{rcl} \breve{\pi}[\breve{\phi}] & = & \breve{\pi}[\breve{\psi}] + \dfrac{1-\breve{\pi}(\infty)}{a}\breve{\pi}(\infty) \\ & = & \dfrac{1-\breve{\pi}(\infty)}{a}\breve{\pi}(\infty) \\ & = & \dfrac{\lambda_0/a}{(1+\lambda_0/a)^2}\dfrac{1}{a} \\ & = & \dfrac{\lambda_0}{(a+\lambda_0)^2} \end{array}$$

so we get

$$\nu[\phi] \quad = \quad \frac{\lambda_0}{(a+\lambda_0)^2}$$

and finally the announced result.

■

Lemma 15 shows that to estimate ϕ (and by consequence the crucial $\breve{\phi}$), we just need to investigate the solutions φ^x of the Poisson equation (29), for $x \in V$ such that $L(x,\infty) > 0$, namely for $x \in \delta V$.

With the notation of the introduction and from Theorem 5, we have

$$\forall\ \omega \in \delta V, \qquad \|\varphi^\omega\|_\infty \quad \leqslant \quad 2\nu(\omega)\Sigma^\dagger_\omega \tag{31}$$

Putting together the above computations, we get:

Corollary 16 *We have*

$$\left\|\breve{\phi}\right\|_\infty \leqslant \frac{\lambda_0}{(a+\lambda_0)^2}+\frac{2\lambda_0}{a+\lambda_0}\Sigma^\dagger$$

Proof

By definition of ϕ, we have

$$\begin{aligned}\left\|\breve{\phi}\right\|_\infty &= \|\phi\|_\infty \vee \left|\breve{\phi}(\infty)\right| \\ &= \|\phi\|_\infty \vee |\nu[\phi]| \\ &= \|\phi\|_\infty\end{aligned}$$

It follows from Lemma 15 and (31) that

$$\begin{aligned}\|\phi\|_\infty &\leqslant \frac{\lambda_0}{(a+\lambda_0)^2}+\frac{1}{a+\lambda_0}\sum_{\omega\in\delta V}\breve{L}(\omega,\infty)\left\|\varphi^\omega\right\|_\infty \\ &= \frac{\lambda_0}{(a+\lambda_0)^2}+\frac{2}{a+\lambda_0}\sum_{\omega\in\delta V}\nu(\omega)\breve{L}(\omega,\infty)\Sigma^\dagger_\omega \\ &\leqslant \frac{\lambda_0}{(a+\lambda_0)^2}+\frac{2\lambda_0}{a+\lambda_0}\sum_{\omega\in\delta V}\Sigma^\dagger_\omega\,\zeta(\omega) \\ &= \frac{\lambda_0}{(a+\lambda_0)^2}+\frac{2\lambda_0}{a+\lambda_0}\Sigma^\dagger\end{aligned}$$

where the last identity of Lemma 13, as well as (4) and (5), were taken into account.

■

Coming back to (24), we deduce that for any $x\in V$ and any $t\geqslant 0$,

$$\left|\mathbb{E}[\mathbb{1}_\infty(\breve{X}_x(t\wedge\breve{\tau}_x))]-\mathbb{1}_\infty(x)-\frac{a\breve{\pi}(\infty)}{1-\breve{\pi}(\infty)}\mathbb{E}[t\wedge\breve{\tau}_x]\right| \leqslant 2\frac{a}{1-\breve{\pi}(\infty)}\left\|\breve{\phi}\right\|_\infty$$

namely

$$\begin{aligned}|\mathbb{P}[\breve{\tau}_x\leqslant t]-\lambda_0\mathbb{E}[t\wedge\breve{\tau}_x]| &\leqslant \frac{2a}{1-\breve{\pi}(\infty)}\left(\frac{\lambda_0}{(a+\lambda_0)^2}+\frac{2\lambda_0}{a+\lambda_0}\Sigma^\dagger\right) \\ &= 2(a+\lambda_0)\left(\frac{\lambda_0}{(a+\lambda_0)^2}+\frac{2\lambda_0}{a+\lambda_0}\Sigma^\dagger\right) \\ &= \frac{2\lambda_0}{(a+\lambda_0)}+4\lambda_0\Sigma^\dagger\end{aligned}$$

Remark that the l.h.s., λ_0 and $\Sigma^\dagger$ do not depend on the choice of a, so we can let a go to infinity to get

$$|\mathbb{P}[\breve{\tau}_x\leqslant t]-\lambda_0\mathbb{E}[t\wedge\breve{\tau}_x]| \leqslant 4\lambda_0\Sigma^\dagger \tag{32}$$

which is the desired bound of Theorem 1.

Instead of Theorem 5, we could have used Proposition 6 in the proof of Corollary 16. Then we end up with

$$\left\|\breve{\phi}\right\|_\infty \leqslant \frac{\lambda_0}{(a+\lambda_0)^2}+\frac{\widetilde{\Sigma}_*}{a+\lambda_0}\sum_{\omega\in\delta V}\breve{L}(\omega,\infty) \tag{33}$$

where, as in Subsection 2.1,

$$\begin{aligned}\widetilde{\Sigma}_* &\coloneqq \sum_{\lambda\in\widetilde{\Lambda}_*}\frac{1}{\lambda}\\ \widetilde{\Lambda}_* &\coloneqq \widetilde{\Lambda}\backslash\{0\}\end{aligned}$$

and $\widetilde{\Lambda}$ is the multiset consisting of the spectrum of $-\widetilde{L}$ with its algebraic multiplicities (which contains 0 with multiplicity 1, by Markovianity and irreducibility).

From (33), we deduce as above, that for any $x\in V$,

$$|\mathbb{P}[\breve{\tau}_x \leqslant t]-\lambda_0\mathbb{E}[t\wedge\breve{\tau}_x]| \leqslant \frac{2\lambda_0}{(a+\lambda_0)}+2\widetilde{\Sigma}_*\sum_{\omega\in V}\breve{L}(\omega,\infty)$$

Letting a go to infinity, we get an alternative bound to Theorem 1:

$$|\mathbb{P}[\tau_x \leqslant t]-\lambda_0\mathbb{E}[t\wedge\tau_x]| \leqslant 2\widetilde{\Sigma}_*\sum_{\omega\in\delta V}\left(|L(\omega,\omega)|-\sum_{y\in V}L(\omega,y)\right) \tag{34}$$

Let us give an alternative description of $\widetilde{\Sigma}_*$. Consider Λ the multiset consisting of the spectrum of $-L$ with its algebraic multiplicities. By irreducibility of L, $\lambda_0\in\Lambda$ with multiplicity 1, but 0 does not belong to Λ, because L is a strictly sub-Markovian generator. Denote $\Lambda_*\coloneqq\Lambda\backslash\{\lambda_0\}$ and

$$\Sigma_* \coloneqq \sum_{\lambda\in\Lambda_*}\frac{1}{\lambda} \tag{35}$$

We have

$$\Sigma_* = \widetilde{\Sigma}_*$$

This result is an immediate consequence the following result, which is interesting in itself.

Proposition 17 *We have* $\widetilde{\Lambda}_*=\Lambda_*$.

Proof

Consider $\lambda\in\Lambda_*$ and let f be an eigenvector associated to λ for $-L$: we have $L[f]=-\lambda f$. Extend f into $\breve{f}$, the function on $\breve{V}$ coinciding with f on V and such that $\breve{f}(\infty)=0$. Then on V, we have $L[f]=\breve{L}[\breve{f}]$. It follows from (1) that

$$\begin{aligned}\nu[f] &= -\frac{1}{\lambda}\nu[L[f]] \\ &= \frac{\lambda_0}{\lambda}\nu[f]\end{aligned} \tag{36}$$

Since $\lambda\neq\lambda_0$, we deduce that $\nu[f]=0$, namely $\widetilde{f}=\breve{f}$ and $L[f]=\widetilde{L}[f]$. Thus $\lambda\in\widetilde{\Lambda}$ and since $\lambda\neq 0$, we get $\lambda\in\widetilde{\Lambda}_*$.

A similar reasoning is also valid if we consider a multiplicity of λ coming from a Jordan block of $-L$. Indeed, it is sufficient to see that if $L[f]=-\lambda f+g$, with $\nu[g]=0$, then $\nu[f]=0$. This is true, since (36) still holds.

It follows that apart from their respective eigenvalues 0 and λ_0, $\widetilde{L}$ and L have the same spectral structure, namely $\widetilde{\Lambda}_*=\Lambda_*$.

■

Thus (34) can be rewritten under the form

$$|\mathbb{P}[\breve{\tau}_x \leqslant t] - \lambda_0\mathbb{E}[t \wedge \breve{\tau}_x]| \quad \leqslant \quad 2\Sigma_* \sum_{\omega\in\delta V} \left(|L(\omega,\omega)| - \sum_{y\in V} L(\omega,y) \right) \tag{37}$$

This bound is more explicit in terms of L, since it only uses its spectrum (and not the spectra of the $L_\omega^\dagger$ for $\omega \in \delta V$) and is generically as good as Theorem 1 on two-point state spaces at small temperature, see Remark 25 of Section 5. But in Remark 29, we will check on an example that this is no longer true for larger state spaces.

Remark 18 The partial equality of spectra presented in Proposition 17 suggests that there could exist an intertwining between L and $\widetilde{L}$, namely we could find a Markov kernel K from V to V such that either

$$\widetilde{L}K \quad = \quad KL \tag{38}$$

or

$$LK \quad = \quad K\widetilde{L} \tag{39}$$

Nevertheless this is wrong: for (38), multiply on the left by ν, the invariant probability of $\widetilde{L}$, to get $\nu KL = 0$, meaning that the probability νK is invariant for L. But there is no such invariant probability, since L is strictly sub-Markovian. Concerning (39), multiply on the left by the quasi-stationary measure ν to obtain $-\lambda_0\nu K = \nu K\widetilde{L}$. Since νK is a probability distribution, it is not 0, so that it is an eigenvector of $\widetilde{L}$ associated to the eigenvalue $-\lambda_0$. It follows that $\lambda_0 \in \widetilde{\Sigma}_*$, namely $\lambda_0 \in \Sigma_*$, a contradiction.

Yet there exists an intertwining relation from $\breve{L}$ to $\widetilde{L}$, i.e. a Markov kernel K (also called a **link**) from $\breve{V}$ to V such that $\breve{L}K = K\widetilde{L}$. Furthermore there is such a relation with K of rank $|V| - 1$. Indeed, note that the spectrum $\breve{\Lambda}$ of $-\breve{L}$ is equal to $\Lambda \sqcup \{0\}$ as multisets: $0 \in \breve{\Lambda}$ by Markovianity and the eigenvectors of L are extended into eigenvectors of $\breve{L}$ by imposing they vanish at ∞ (the same is true for the vectors associated to Jordan blocks). Following the arguments of [18], a link K of rank $|V|-1$ can be constructed by perturbing the Markov kernel from $\breve{V}$ to V whose lines are all equal to ν. As shown in general by Diaconis and Fill [6], an intertwining relation from an absorbed process to an ergodic process can be used to construct strong stationary times from absorption times. Here this is quite simple: from a Markov process $\breve{X}$ associated to $\breve{L}$, construct a Markov process $\widetilde{X}$ associated to $\widetilde{L}$ by redistributing the position according to ν instead of hitting ∞. It appears then that the absorption time for $\breve{X}$ (i.e. the hitting time of ∞) is a strong stationary time for $\widetilde{X}$.

□

4 Exit position

Here we prove Theorem 3. The arguments follow those of the previous section, with similar notations, that coincide should we have $\partial V = \{\infty\}$. We preferred to separate the treatment of the exit time and of the exit position for the sack of clarity for the former.

As in Section 3, we begin by transforming $\bar{L}$ into an ergodic Markov generator $\breve{L} \coloneqq (\breve{L}(x,y))_{x,y\in\bar{V}}$. Let be given a positive number $a > 0$. We define $\breve{L}$ by only modifying the rows indexed by ∂V:

$$\forall\ x \neq y \in \bar{V}, \qquad \breve{L}(x,y) \quad \coloneqq \quad \begin{cases} L(x,y) & \text{, if } x \in V \text{ and } y \in \bar{V} \\ a\nu(y) & \text{, if } x \in \partial V \text{ and } y \in \bar{V} \end{cases}$$

where we recall that ν is the quasi-stationary measure of the sub-Markovian generator L.

By irreducibility, $\breve{L}$ admits a unique invariant probability $\breve{\pi}$. Let us compute it in terms of ν and μ:

Lemma 19 *We have*

$$\breve{\pi} \quad = \quad \frac{a\nu + \lambda_0 \sum_{\omega\in\partial V} \mu(\omega)\delta_\omega}{a+\lambda_0}$$

Proof

Define

$$\forall\ \omega \in \partial V, \qquad \alpha_\omega \quad \coloneqq \quad \frac{1}{a}\sum_{x\in V} \nu(x)\bar{L}(x,\omega)$$

We begin by showing that $\breve{\pi}\breve{L} = 0$, where $\breve{\pi} \coloneqq \nu + \sum_{\omega\in\partial V} \alpha_\omega \delta_\omega$.

- For any $\omega_0 \in \partial V$, we have

$$\begin{aligned}
\breve{\pi}\breve{L}(\omega_0) &= \sum_{x\in V} \nu(x)\breve{L}(x,\omega_0) + \sum_{\omega\in\partial V} \alpha_\omega L(\omega,\omega_0) \\
&= \sum_{x\in V} \nu(x)\bar{L}(x,\omega_0) + \alpha_{\omega_0} L(\omega_0,\omega_0) \\
&= a\alpha_{\omega_0} - \alpha_{\omega_0} a \\
&= 0
\end{aligned}$$

- For any $x_0 \in V$, we have

$$\begin{aligned}
\breve{\pi}\breve{L}(x_0) &= \sum_{x\in V} \nu(x)\breve{L}(x,x_0) + \sum_{\omega\in\partial V} \alpha_\omega L(\omega,x_0) \\
&= \sum_{x\in V} \nu(x)L(x,x_0) + \sum_{\omega\in\partial V} \alpha_\omega a\nu(x_0) \\
&= -\lambda_0\nu(x_0) + \nu(x_0)\sum_{\omega\in\partial V} \alpha_\omega a \\
&= \left(\sum_{\omega\in\partial V} \alpha_\omega a - \lambda_0\right)\nu(x_0)
\end{aligned}$$

  To conclude that $\breve{\pi}\breve{L}(x_0) = 0$, it remains to see that

$$\sum_{\omega\in\partial V} a\alpha_\omega \quad = \quad \lambda_0 \tag{40}$$

  By definition, we have

$$\begin{aligned}
\sum_{\omega\in\partial V} a\alpha_\omega &= \sum_{\omega\in\partial V}\sum_{x\in V} \nu(x)\bar{L}(x,\omega) \\
&= \sum_{x\in V} \nu(x) \sum_{\omega\in\partial V} \bar{L}(x,\omega) \\
&= -\sum_{x\in V} \nu(x) \sum_{y\in V} \bar{L}(x,y) \\
&= -\sum_{x\in V} \nu(x) \sum_{y\in V} L(x,y) \\
&= -\sum_{x,y\in V} \nu(x)L(x,y) \\
&= \lambda_0 \sum_{y\in V} \nu(y) \\
&= \lambda_0
\end{aligned}$$

Recalling the definition of μ in (6), note that

$$\forall\ \omega \in \partial V, \qquad \alpha_\omega \;=\; \frac{Z\mu(\omega)}{a}$$

From (40), we deduce that $Z = \lambda_0$, so that $\breve{\pi} = \nu + \frac{\lambda_0}{a}\sum_{\omega\in\partial V}\mu(\omega)\delta_\omega$ and the desired result follows by normalization.

■

For any $\omega \in \partial V$, consider the solution $\breve{\psi}_\omega$ of the Poisson equation:

$$\begin{cases} \breve{L}[\breve{\psi}_\omega] &= \mathbb{1}_{\{\omega\}} - \breve{\pi}(\omega) \\ \breve{\pi}[\breve{\psi}_\omega] &= 0 \end{cases}$$

As in the previous section, the idea is that ψ_ω is close to $-\frac{1}{a}(\mathbb{1}_\omega - \breve{\pi}(\omega)\mathbb{1}_{\partial V})$, when $\lambda_0\Sigma^\dagger$ is small. Let us first heuristically deduce Theorem 3 from this belief.

For $x \in V$, let $\breve{X}_x := (\breve{X}_x(t))_{t\geqslant 0}$ be a Markov process associated to the generator $\breve{L}$ and starting from x. Let $\breve{\tau}_x$ be its hitting time of ∂V. From the martingale problem satisfied by $\breve{X}_x$, there exists a martingale $(\breve{M}_t)_{t\geqslant 0}$ such that for any $t \geqslant 0$,

$$\begin{aligned} \breve{\psi}_\omega(\breve{X}_x(t\wedge\breve{\tau}_x)) &= \breve{\psi}_\omega(\breve{X}_x(0)) + \int_0^{t\wedge\breve{\tau}_x} \breve{L}[\breve{\psi}_\omega](\breve{X}_x(s))\,ds + \breve{M}_{t\wedge\breve{\tau}_x} \\ &= \breve{\psi}_\omega(x) + \int_0^{t\wedge\breve{\tau}_x} \mathbb{1}_{\{\omega\}}(\breve{X}_x(s)) - \breve{\pi}(\omega)\,ds + \breve{M}_{t\wedge\breve{\tau}_x} \\ &= \breve{\psi}_\omega(x) - \breve{\pi}(\omega)(t\wedge\breve{\tau}_x) + \breve{M}_{t\wedge\breve{\tau}_x} \end{aligned}$$

Taking expectation, we get

$$\mathbb{E}[\breve{\psi}_\omega(\breve{X}_x(t\wedge\breve{\tau}_x))] \;=\; \breve{\psi}_\omega(x) - \breve{\pi}(\omega)\mathbb{E}[t\wedge\breve{\tau}_x] \tag{41}$$

According to the expected behavior of ψ_ω, we should have

$$\forall\ t\geqslant 0, \qquad \mathbb{E}[\mathbb{1}_{\{\omega\}}(\breve{X}_x(t\wedge\breve{\tau}_x)) - \breve{\pi}(\omega)\mathbb{1}_{\partial V}(\breve{X}_x(t\wedge\breve{\tau}_x))] \;\approx\; \mathbb{1}_{\{\omega\}}(x) - \breve{\pi}(\omega)\mathbb{1}_{\partial V}(x) + a\breve{\pi}(\omega)\mathbb{E}[t\wedge\breve{\tau}_x]$$

namely

$$\forall\ t\geqslant 0, \qquad \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\ \breve{\tau}_x \leqslant t] - \breve{\pi}(\omega)\mathbb{P}[\breve{\tau}_x\leqslant t] \;\approx\; a\breve{\pi}(\omega)\mathbb{E}[t\wedge\breve{\tau}_x]$$

According to Section 3, we also have

$$\mathbb{E}[t\wedge\breve{\tau}_x] \;\approx\; \frac{\mathbb{P}[\breve{\tau}_x\leqslant t]}{\lambda_0}$$

so that

$$\forall\ \omega\in\partial V,\ \forall\ t\geqslant 0, \qquad \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\ \breve{\tau}_x \leqslant t] - \breve{\pi}(\omega)\mathbb{P}[\breve{\tau}_x\leqslant t] \;\approx\; \frac{a\breve{\pi}(\omega)}{\lambda_0}\mathbb{P}[\breve{\tau}_x\leqslant t]$$

i.e.

$$\forall\ \omega\in\partial V,\ \forall\ t\geqslant 0, \qquad \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\ \breve{\tau}_x \leqslant t] \;\approx\; \frac{\lambda_0 + a}{a}\breve{\pi}(\omega)\mathbb{P}[\breve{\tau}_x\leqslant t]$$

It would mean that $\breve{X}_x(\breve{\tau}_x)$ and $\breve{\tau}_x$ are almost independent and the distribution of the former is given by

$$\begin{aligned} \forall\ \omega\in\partial V, \qquad \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega] &\approx \frac{\lambda_0 + a}{a}\breve{\pi}(\omega) \\ &= \mu(\omega) \end{aligned}$$

where Lemma 19 was taken into account.

Let us now come to more rigorous computations. As suggested by the above heuristic, for any fixed $\omega \in \partial V$, we should investigate the function

$$\breve{\phi}_\omega \quad := \quad \breve{\psi}_\omega + \frac{1}{a}(\mathbb{1}_{\{\omega\}} - \breve{\pi}(\omega)\mathbb{1}_{\partial V})$$

Lemma 20 *We have*

$$\breve{L}[\breve{\phi}_\omega] \quad = \quad \frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))\mathbb{1}_{\{x\}} - \breve{\pi}(\omega)\mathbb{1}_V$$

where $\bar{L}(x,\partial V) = \sum_{w\in\partial V}\bar{L}(x,w)$.

Proof
Note that

$$\breve{L}[\mathbb{1}_{\{\omega\}} - \breve{\pi}(\omega)\mathbb{1}_{\partial V}] \quad = \quad \sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))\mathbb{1}_{\{x\}} - a(\mathbb{1}_{\{\omega\}} - \breve{\pi}(\omega)\mathbb{1}_{\partial V}) \tag{42}$$

Indeed, we compute that for any $w \in \partial V$,

$$\begin{aligned}
\breve{L}[\mathbb{1}_{\{\omega\}}](w) &= \breve{L}(w,\omega)\\
&= \mathbb{1}_{\{\omega\}}(w)\breve{L}(\omega,\omega)\\
&= -\mathbb{1}_{\{\omega\}}(w)\sum_{y\in V}\breve{L}(\omega,y)\\
&= -a\mathbb{1}_{\{\omega\}}(w)\sum_{y\in V}\nu(y)\\
&= -a\mathbb{1}_{\{\omega\}}(w)
\end{aligned}$$

and similarly, for any $w \in \partial V$,

$$\begin{aligned}
\breve{L}[\mathbb{1}_{\partial V}](w) &= \sum_{w'\in\partial V}\breve{L}(w,w')\\
&= \breve{L}(w,w)\\
&= -a
\end{aligned}$$

Furthermore, we clearly have for any $x \in V$, $\breve{L}[\mathbb{1}_{\{\omega\}} - \breve{\pi}(\omega)\mathbb{1}_{\partial V}](x) = \bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V)$ and (42) follows.

Taking into account that by definition

$$\breve{L}[\breve{\psi}_\omega] \quad = \quad \mathbb{1}_{\{\omega\}} - \breve{\pi}(\omega)\mathbb{1}_{\partial V} - \breve{\pi}(\omega)\mathbb{1}_V$$

we deduce that

$$\begin{aligned}
\breve{L}[\breve{\phi}_\omega] &= \breve{L}[\breve{\psi}_\omega] + \frac{1}{a}\breve{L}[\mathbb{1}_{\{\omega\}} - \breve{\pi}(\omega)\mathbb{1}_{\partial V}]\\
&= \frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \bar{L}(x,\partial V))\mathbb{1}_{\{x\}} - \breve{\pi}(\omega)\mathbb{1}_V
\end{aligned}$$

■

It follows that $\breve{L}[\breve{\phi}_\omega](w) = 0$ for any $w \in \partial V$, namely

$$\forall\ w \in \partial V, \qquad \breve{\phi}_\omega(w) \quad = \quad \nu[\breve{\phi}_\omega] \tag{43}$$

This observation leads us to introduce a new generator $\widetilde{L}$ on V. Any $f \in \mathcal{F}(V)$ is extended into a function $\widetilde{f}$ on $\bar{V}$ by imposing

$$\forall\ w \in \partial V, \qquad \widetilde{f}(w) \ \coloneqq \ \nu[f]$$

We consider the generator $\widetilde{L}$ given by

$$\forall\ f \in \mathcal{F}(V), \forall\ x \in V, \qquad \widetilde{L}[f](x) \ \coloneqq \ \breve{L}[\widetilde{f}](x)$$

Again, the generator $\widetilde{L}$ is the Steklov operator associated to $\breve{L}$ and to the "boundary" V of $\bar{V}$. It follows that the invariant probability measure of $\widetilde{L}$ is the normalization of the restriction of $\breve{\pi}$ to V, namely ν. As in Section 3, $\widetilde{L}$ is described by (2).

Denote ϕ_ω the restriction of $\breve{\phi}_\omega$ on V. Due to (43), $\widetilde{\phi}_\omega$ coincides with $\breve{\phi}_\omega$, so by definition of $\widetilde{L}$, we get on V:

$$\widetilde{L}[\phi_\omega] \ = \ \frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))\mathbb{1}_{\{x\}} - \breve{\pi}(\omega)\mathbb{1}_V$$

or equivalently, using that $\nu[\widetilde{L}[\phi_\omega]] = 0$,

$$\widetilde{L}[\phi_\omega] \ = \ \frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))(\mathbb{1}_{\{x\}} - \nu(x)) \tag{44}$$

Recalling that for any $x \in V$, φ^x is the solution of the Poisson equation (29), (44) implies that

$$\phi_\omega \ = \ \nu[\phi_\omega] + \frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))\varphi^x$$

Indeed, we get

$$\widetilde{L}\left[\phi_\omega - \frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))\varphi^x\right] \ = \ 0$$

thus by irreducibility of $\widetilde{L}$, ϕ_ω and $\frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))\varphi^x$ coincide up to an additive constant, which is necessarily $\nu[\phi_\omega]$.

The next result shows that ϕ_ω can be completely expressed in terms of $(\varphi^x)_{x\in V}$.

Lemma 21 *We have*

$$\phi_\omega \ = \ \frac{1}{a+\lambda_0}\breve{\pi}(\omega) + \frac{1}{a}\sum_{x\in V}(\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V))\varphi^x$$

Proof

It follows from Lemma 19 and (43) that

$$\begin{aligned}\breve{\pi}[\breve{\phi}_\omega] \ &= \ \frac{a\nu[\phi_\omega] + \lambda_0\sum_{w\in\partial V}\mu(w)\phi_\omega(w)}{a+\lambda_0}\\ &= \ \nu[\phi_\omega]\end{aligned}$$

By definition of $\breve{\phi}_\omega$, we also have

$$\begin{aligned}\breve{\pi}[\breve{\phi}_\omega] \ &= \ \breve{\pi}[\breve{\psi}_\omega] + \frac{1}{a}(\breve{\pi}(\omega) - \breve{\pi}(\omega)\breve{\pi}(\partial V))\\ &= \ \frac{\breve{\pi}(V)}{a}\breve{\pi}(\omega)\\ &= \ \frac{\breve{\pi}(\omega)}{a+\lambda_0}\end{aligned}$$

so we get

$$\nu[\phi_\omega] \;=\; \frac{1}{a+\lambda_0}\breve{\pi}(\omega)$$

and finally the announced result.

■

With the notation of the introduction and from Theorem 5, we have

$$\left\{\begin{array}{lrcl} \forall\ x \notin \delta V, & \bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V) & = & 0 \\ \forall\ x \in \delta V, & \|\varphi^x\|_\infty & \leqslant & 2\nu(x)\Sigma_x^\dagger \end{array}\right. \tag{45}$$

so putting together the above computations, we get:

Corollary 22 *We have*

$$\left\|\breve{\phi}_\omega\right\|_\infty \;\leqslant\; \left(\frac{\lambda_0}{(a+\lambda_0)^2} + \frac{2a+4\lambda_0}{a(a+\lambda_0)}\Sigma^\dagger\lambda_0\right)\mu(\omega)$$

Proof
By definition of ϕ_ω, we have

$$\begin{aligned} \left\|\breve{\phi}_\omega\right\|_\infty &= \|\phi_\omega\|_\infty \vee \max\left\{\left|\breve{\phi}_\omega(w)\right| \,:\, w \in \partial V\right\} \\ &= \|\phi_\omega\|_\infty \vee |\nu[\phi_\omega]| \\ &= \|\phi_\omega\|_\infty \end{aligned}$$

It follows from Lemma 21 and (45) that

$$\begin{aligned} \|\phi_\omega\|_\infty &\leqslant \frac{1}{a+\lambda_0}\breve{\pi}(\omega) + \frac{1}{a}\sum_{x\in V}\left|\bar{L}(x,\omega) - \breve{\pi}(\omega)\bar{L}(x,\partial V)\right| \|\varphi^x\|_\infty \\ &\leqslant \frac{1}{a+\lambda_0}\breve{\pi}(\omega) + \frac{2}{a}\sum_{x\in\delta V}\nu(x)(\bar{L}(x,\omega) + \breve{\pi}(\omega)\bar{L}(x,\partial V))\Sigma_x^\dagger \\ &= \frac{1}{a+\lambda_0}\breve{\pi}(\omega) + \frac{2}{a}Z\Sigma^\dagger(\chi(\omega) + \breve{\pi}(\omega)) \\ &= \frac{\lambda_0}{(a+\lambda_0)^2}\mu(\omega) + \frac{2}{a}\lambda_0\Sigma^\dagger\left(\chi(\omega) + \frac{\lambda_0}{a+\lambda_0}\mu(\omega)\right) \end{aligned}$$

where in the fourth line, we used that $\lambda_0 = Z$, as seen at the end of the proof of Lemma 19.

■

Coming back to (41), we deduce that for any $\omega \in \partial V$, $x \in V$ and $t \geqslant 0$,

$$\left|\mathbb{E}[\mathbb{1}_\omega(\breve{X}_x(t\wedge\breve{\tau}_x)) - \breve{\pi}(\omega)\mathbb{1}_{\partial V}(\breve{X}_x(t\wedge\breve{\tau}_x))] - a\breve{\pi}(\omega)\mathbb{E}[t\wedge\breve{\tau}_x]\right| \;\leqslant\; 2a\left\|\breve{\phi}_\omega\right\|_\infty$$

namely

$$\begin{aligned} &\left|\mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\, \breve{\tau}_x \leqslant t] - \breve{\pi}(\omega)\mathbb{P}[\breve{\tau}_x \leqslant t] - a\breve{\pi}(\omega)\mathbb{E}[t\wedge\breve{\tau}_x]\right| \\ &\qquad\leqslant\; 2\left(\frac{a\lambda_0}{(a+\lambda_0)^2}\mu(\omega) + 2\lambda_0\Sigma^\dagger\left(\chi(\omega) + \frac{\lambda_0}{a+\lambda_0}\mu(\omega)\right)\right) \end{aligned} \tag{46}$$

In this bound, the term $\breve{\pi}(\omega)$ depends on $a > 0$, as we have

$$\breve{\pi}(\omega) \quad = \quad \frac{\lambda_0 \mu(\omega)}{a + \lambda_0}$$

Thus, letting a go to infinity in (46), we get

$$\left| \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\, \breve{\tau}_x \leqslant t] - \lambda_0 \mu(\omega) \mathbb{E}[t \wedge \breve{\tau}_x] \right| \quad \leqslant \quad 4\Sigma^\dagger \lambda_0 \chi(\omega) \tag{47}$$

Taking (32) into account, we obtain

$$\left| \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\, \breve{\tau}_x \leqslant t] - \mu(\omega) \mathbb{P}[\breve{\tau}_x \leqslant t] \right| \quad \leqslant \quad 8\Sigma^\dagger \lambda_0 \chi(\omega) \tag{48}$$

which shows that the exit time and position are almost independent when $\Sigma^\dagger \lambda_0$ is small. To end up with the desired bound of Theorem 3, it remains to use Theorem 1.

As in Corollary 16, instead of Theorem 5, we could have used Proposition 6 in the proof of Corollary 22. It leads to

$$\left| \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\, \breve{\tau}_x \leqslant t] - \lambda_0 \mu(\omega) \mathbb{E}[t \wedge \breve{\tau}_x] \right| \quad \leqslant \quad 2\Sigma_* \sum_{y \in \delta V} \bar{L}(y, \omega)$$

and taking (34) into account,

$$\left| \mathbb{P}[\breve{X}_x(\breve{\tau}_x) = \omega,\, \breve{\tau}_x \leqslant t] - \mu(\omega) \mathbb{P}[\breve{\tau}_x \leqslant t] \right| \quad \leqslant \quad 2\Sigma_* (\bar{L}(\delta V, \omega) + \mu(\omega) \bar{L}(\delta V, \partial V))$$

where for any disjoint $A, B \subset \bar{V}$, $\bar{L}(A, B) \coloneqq \sum_{y \in A, z \in B} \bar{L}(y, z)$.

For the reason presented in Section 3, these bounds seem less interesting than (47) and (48) respectively.

5 Simple examples at small temperature

Here we illustrate the metastability phenomenon in the simplest situation, namely a two-point state space at small temperature. This benchmark will enable us to see that Corollary 4 is quite sharp, contrary to Theorem 1. Resorting to a 3-point state space, we also underline the difference between the estimates of Theorem 5 and Proposition 6.

On the state space $V \coloneqq \{0, 1\}$, let be given a family $(L_\beta)_{\beta \geqslant 0} \coloneqq ((L_\beta(x, y))_{x, y \in V})_{\beta \geqslant 0}$ of irreducible strictly sub-Markovian generators. The parameter $\beta \geqslant 0$ is to be seen as an inverse temperature and we are interested in the asymptotic regime when β goes to infinity (namely the temperature goes to 0_+). As in the introduction, denote $\breve{V} \coloneqq \{0, 1, \infty\}$ and to avoid reference to $\breve{L}_\beta$, for $\beta \geqslant 0$, we adopt the simplified notation $L_\beta(x, \infty) \coloneqq -L_\beta(x, x) - L_\beta(x, 1 - x)$, for any $x \in \{0, 1\}$. Thus we have

$$\forall\ \beta \geqslant 0, \qquad L_\beta \quad \coloneqq \quad \begin{pmatrix} -L_\beta(0,1) - L_\beta(0,\infty) & L_\beta(0,1) \\ L_\beta(1,0) & -L_\beta(1,0) - L_\beta(1,\infty) \end{pmatrix}$$

Furthermore, to get a more convenient landscape in the setting of Theorem 3 and Corollary 4, let us split ∞ into the two points -1 and 2, and consider on the state space $\bar{V} \coloneqq \{-1, 0, 1, 2\}$, the following absorbing extension $\bar{L}_\beta$ of L_β, for any given $\beta \geqslant 0$:

$$\bar{L}_\beta \quad \coloneqq \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ L_\beta(0,\infty) & -L_\beta(0,1) - L_\beta(0,\infty) & L_\beta(0,1) & 0 \\ 0 & L_\beta(1,0) & -L_\beta(1,0) - L_\beta(1,\infty) & L_\beta(1,\infty) \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We assume the existence of the following limits

$$\forall\ (x,y) \in V \times \check{V},\ x \neq y, \qquad W(x,y) \ \coloneqq\ -\lim_{\beta\to+\infty} \frac{1}{\beta} \ln(L_\beta(x,y)) \ \in\ [0,+\infty) \tag{49}$$

Let us simplify the notations and define

$$\begin{array}{llllllllllll} a & \coloneqq & L_\beta(0,\infty), & b & \coloneqq & L_\beta(0,1), & c & \coloneqq & L_\beta(1,0), & d & \coloneqq & L_\beta(1,\infty) \\ \widetilde{a} & \coloneqq & W(0,\infty), & \widetilde{b} & \coloneqq & W(0,1), & \widetilde{c} & \coloneqq & W(1,0), & \widetilde{d} & \coloneqq & W(1,\infty) \end{array}$$

Except when explicitly said otherwise, in this section we suppose:

$$\text{The numbers } \widetilde{a}, \widetilde{b}, \widetilde{c} \text{ and } \widetilde{d} \text{ are all distinct and } \widetilde{a} + \widetilde{c} \neq \widetilde{b} + \widetilde{d} \tag{50}$$

Furthermore, up to exchanging 0 and 1, we assume that $\widetilde{b} > \widetilde{c}$.

Before applying Theorem 1 and Corollary 4, let us check directly if metastability holds or not, by considering the different possible situations.

- Case (1) where $\widetilde{c} < \widetilde{a} < \widetilde{b}$ and $\widetilde{c} < \widetilde{d}$.

  Taking into account the probabilistic description of a Markov process $X_0 \coloneqq (X_0(t))_{t\geqslant 0}$ associated to the generator $\bar{L}_\beta$ and starting from 0, X_0 stays in $X_0(0) = 0$ for a exponential time τ_1 of parameter $a + b \sim a$, because $\widetilde{b} > \widetilde{a}$. The position $X_0(\tau_1)$ is equal to -1 with probability $a/(a+b)$ and to 1 with probability $b/(a+b)$. Thus, up to an exponentially small error (in β), starting from 0, the exit time is an exponential variable of parameter a and the exit position is -1.

  Similarly since $\widetilde{c} < \widetilde{d}$, starting from 1 and up to an exponentially small error, the process X_1 waits an exponential time of parameter c before jumping in 0. From 0, the process behaves like X_0. Since $\widetilde{a} > \widetilde{c}$, the time to jump from 1 to 0 is negligible with respect to the time to jump from 0 to -1. It follows that up to an exponentially small error, again the exit time is an exponential variable of parameter a and the exit position is -1. Thus the exit behavior is independent of the initial state: the metastability phenomenon occurs.

- Case (2) where $\widetilde{a} < \widetilde{b}$ and $\widetilde{c} > \widetilde{d}$.

  Starting from 0 the situation is similar to Case (1). Starting from 1, the process X_1 waits an exponential variable of parameter d before jumping to 2, up to an exponentially small error. The metastability phenomenon does not occur, since the distribution of the exit position strongly depends on the initial point.

- Case (3) where $\widetilde{a} > \widetilde{b}$ and $\widetilde{c} > \widetilde{d}$.

  As in case (2), starting from 1, the process X_1 waits an exponential variable of parameter d before jumping to 2, up to an exponentially small error. Starting from 0, the process X_0 waits an exponential variable of parameter b before jumping to 1, before jumping to 2 after an exponential variable of parameter d, all that up to an exponentially small error. The metastability phenomenon does not occur, because the exit time from 0 is much longer than the exit time from 1.

- Case (4) where $\widetilde{a} > \widetilde{b}$, $\widetilde{c} < \widetilde{d}$.

  As above, all the following statements are up to an exponentially small error. Starting from 1, the process X_1 first reaches 0. Thus the exit position distribution will not depend on the initial state. Furthermore the time to reach 0 from 1 is much smaller than the time to get out of 0 (and first to reach 1). This is sufficient to insure metastability (consider the quasi-stationary distribution as initial distribution, the exit time will be the same as the exit time starting from 0, according to the above arguments). The exit position distribution will be concentrated on 2 (respectively -1), if $\widetilde{a} > \widetilde{b} - \widetilde{c} + \widetilde{d}$ (resp. $\widetilde{a} < \widetilde{b} - \widetilde{c} + \widetilde{d}$).

Let us denote by $\mathcal{M}$ the set of $(\tilde{a},\tilde{b},\tilde{c},\tilde{d}) \in \mathbb{R}_+^4$ satisfying (50) and for which metastability holds, namely corresponding to Cases (1) and (4) above. It is not difficult to check that $\mathcal{M}$ is the set of $(\tilde{a},\tilde{b},\tilde{c},\tilde{d}) \in \mathbb{R}_+^4$ satisfying (50) and $\tilde{c} < \tilde{a} \wedge \tilde{b}$.

For any $\beta \geqslant 0$, consider $\lambda_0(\beta)$ and $\Sigma^\dagger(\beta)$ the quantities associated to L_β as in the introduction. The following result shows that the metastability of Cases (1) and (4) is recovered from Theorem 3 and Corollary 4.

Proposition 23 *When $(\tilde{a},\tilde{b},\tilde{c},\tilde{d}) \in \mathcal{M}$, we have*

$$\lim_{\beta\to+\infty} \beta^{-1} \ln\left(\lambda_0(\beta)\Sigma^\dagger(\beta)\right) \quad < \quad 0 \tag{51}$$

(in particular the l.h.s. limit exists).

Note that for $(\tilde{a},\tilde{b},\tilde{c},\tilde{d}) \in \mathbb{R}_+^4$ satisfying (50) and corresponding to Cases (2) and (3), (51) cannot hold, otherwise we could conclude to metastability. Thus Theorem 3 and Corollary 4 are quite sharp, since they enable to recover the domain of coefficients $(\tilde{a},\tilde{b},\tilde{c},\tilde{d}) \in \mathbb{R}_+^4$ leading to metastability, at least under (50).

Let us start the proof of Proposition 23 by obtaining the behavior of $\lambda_0(\beta)$ at small temperature:

Lemma 24 *As soon as the limits in (49) exist, we have*

$$-\lim_{\beta\to+\infty} \beta^{-1} \ln(\lambda_0(\beta)) \quad = \quad \min(\tilde{a}+\tilde{c},\tilde{a}+\tilde{d},\tilde{b}+\tilde{d}) - \min(\tilde{a},\tilde{c},\tilde{d})$$

In particular, when $(\tilde{a},\tilde{b},\tilde{c},\tilde{d}) \in \mathcal{M}$, we deduce that

$$-\lim_{\beta\to+\infty} \beta^{-1} \ln(\lambda_0(\beta)) \quad = \quad \tilde{a} \wedge (\tilde{b}-\tilde{c}+\tilde{d})$$

and in Case (1), the l.h.s. is $\tilde{a}$.

Proof

For fixed $\beta \geqslant 0$, the two eigenvalues of L_β are the roots of the characteristic polynomial

$$(X-a-b)(X-c-d)-bc \quad = \quad X^2-(a+b+c+d)X+ac+ad+bd$$

Since $\lambda_0(\beta)$ is the smallest of them, we get

$$\begin{aligned}
\lambda_0(\beta) &= \frac{1}{2}(a+b+c+d-\sqrt{\triangle}) \\
&= \frac{1}{2}\frac{(a+b+c+d)^2-\triangle}{a+b+c+d+\sqrt{\triangle}} \\
&= 2\frac{ac+ad+bd}{a+b+c+d+\sqrt{\triangle}}
\end{aligned} \tag{52}$$

where the discriminant is given by

$$\begin{aligned}
\triangle &\coloneqq (a+b+c+d)^2-4(ac+ad+bd) \\
&= (a+b-c-d)^2+4bc
\end{aligned}$$

It is clear that

$$-\lim_{\beta\to+\infty} \beta^{-1} \ln(ac+ad+bd) \quad = \quad \min(\tilde{a}+\tilde{c},\tilde{a}+\tilde{d},\tilde{b}+\tilde{d})$$

Note that

$$\begin{aligned} -\lim_{\beta\to+\infty}\beta^{-1}\ln(a+b+c+d) &= \min(\widetilde{a},\widetilde{b},\widetilde{c},\widetilde{d}) \\ &= \min(\widetilde{a},\widetilde{c},\widetilde{d}) \end{aligned}$$

and, since $(a+b-c-d)^2+4bc \leqslant 4a^2+6b^2+6c^2+4d^2$, that

$$-\lim_{\beta\to+\infty}\beta^{-1}\ln(\sqrt{(a+b-c-d)^2+4bc}) \;\geqslant\; \min(\widetilde{a},\widetilde{b},\widetilde{c},\widetilde{d})$$

so we deduce that

$$-\lim_{\beta\to+\infty}\beta^{-1}\ln(a+b+c+d+\sqrt{\triangle}) \;=\; \min(\widetilde{a},\widetilde{c},\widetilde{d})$$

The announced results are an immediate consequence of (52).

When $(\widetilde{a},\widetilde{b},\widetilde{c},\widetilde{d})\in\mathcal{M}$, we have $\widetilde{c}=\min(\widetilde{a},\widetilde{c},\widetilde{d})$ and $\widetilde{d}>\widetilde{c}$, so that

$$\begin{aligned} -\lim_{\beta\to+\infty}\beta^{-1}\ln(\lambda_0(\beta)) &= \min(\widetilde{a},\widetilde{a}-\widetilde{c}+\widetilde{d},\widetilde{b}-\widetilde{c}+\widetilde{d}) & (53)\\ &= \min(\widetilde{a},\widetilde{b}-\widetilde{c}+\widetilde{d}) & (54) \end{aligned}$$

In Case (1), we have $\widetilde{b}\geqslant\widetilde{a}$ and $\widetilde{d}\geqslant\widetilde{c}$, so that $\min(\widetilde{a},\widetilde{b}-\widetilde{c}+\widetilde{d})=\widetilde{a}$.

In Case (4), both alternatives $\widetilde{a}>\widetilde{b}-\widetilde{c}+\widetilde{d}$ and $\widetilde{a}<\widetilde{b}-\widetilde{c}+\widetilde{d}$ are possible.

■

Remark 25 For $\beta\geqslant 0$, let $\Sigma_*(\beta)$ be defined as in (35): we have $\Sigma_*(\beta)=1/\lambda_1(\beta)$, where $\lambda_1(\beta)=\frac{1}{2}(a+b+c+d+\sqrt{\triangle})$ is the other eigenvalue of L_β. It is clear that

$$\lim_{\beta\to+\infty}\beta^{-1}\ln(\lambda_1(\beta)) \;=\; -(\widetilde{a}\wedge\widetilde{b}\wedge\widetilde{c}\wedge\widetilde{d})$$

Note that in the present context, the r.h.s. of (37) is just

$$2\Sigma_*(\beta)(L_\beta(0,\infty)+L_\beta(1,\infty)) \;=\; \frac{2}{\lambda_1(\beta)}(\widetilde{a}+\widetilde{d})$$

so that

$$\lim_{\beta\to+\infty}\beta^{-1}\ln\Big(2\Sigma_*(\beta)(L_\beta(0,\infty)+L_\beta(1,\infty))\Big) \;=\; \widetilde{a}\wedge\widetilde{b}\wedge\widetilde{c}\wedge\widetilde{d}-\widetilde{a}\wedge\widetilde{d}$$

and the r.h.s. is negative if and only $\widetilde{b}\wedge\widetilde{c}<\widetilde{a}\wedge\widetilde{d}$, i.e. $\widetilde{c}<\widetilde{a}\wedge\widetilde{d}$.

Thus it appears that on the two-point state space, (37) is as good as Theorem 1 (as long as the exponential rate is concerned at small temperature). But this is no longer true on state spaces containing at least three points, see Remark 29 below.

□

For $\beta\geqslant 0$, let ν_β be the quasi-stationary distribution associated to L_β.

Lemma 26 *When $(\widetilde{a},\widetilde{b},\widetilde{c},\widetilde{d})\in\mathcal{M}$, we have for large $\beta\geqslant 0$,*

$$\forall\, x\in\{0,1\},\qquad \nu_\beta(x) \;\sim\; \begin{cases} 1 & ,\ \textit{if } x=0\\ b/c & ,\ \textit{if } x=1\end{cases} \tag{55}$$

and in particular,

$$\forall\, x\in\{0,1\},\qquad \lim_{\beta\to+\infty}\beta^{-1}\ln(\nu_\beta(x)) \;=\; \begin{cases} 0 & ,\ \textit{if } x=0\\ \widetilde{c}-\widetilde{b} & ,\ \textit{if } x=1\end{cases}$$

Proof

For $\beta \geqslant 0$, let $\widetilde{L}_\beta$ be associated to L_β and ν_β as in the introduction. More precisely, we have

$$\widetilde{L}_\beta = \begin{pmatrix} -(b + a\nu_\beta(1)) & b + a\nu_\beta(1) \\ c + d\nu_\beta(0) & -(c + d\nu_\beta(0)) \end{pmatrix}$$

Since ν_β is invariant for $\widetilde{L}_\beta$, we deduce that

$$\nu_\beta(0) = \frac{c + d\nu_\beta(0)}{b + a\nu_\beta(1) + c + d\nu_\beta(0)} \tag{56}$$

$$\nu_\beta(1) = \frac{b + a\nu_\beta(1)}{b + a\nu_\beta(1) + c + d\nu_\beta(0)} \tag{57}$$

In both Cases (1) and (4), we have for large $\beta \geqslant 0$,

$$c \gg a \vee b \vee d$$

and we deduce from (56) that

$$\nu_\beta(1) \sim 1$$

From (57), we get

$$\nu_\beta(1) \sim \frac{b + a\nu_\beta(1)}{c}$$

so from $a \ll c$, we deduce

$$\nu_\beta(1) \sim \frac{b}{c}$$

■

We can now come to the

Proof of Proposition 23

For $\beta \geqslant 0$, consider the probability ζ_β defined as in (4). We have

$$\forall\, x \in \{0,1\}, \qquad \zeta_\beta(x) = \begin{cases} \dfrac{a\nu_\beta(0)}{a\nu_\beta(0) + b\nu_\beta(1)} & \text{, if } x = 0 \\ \dfrac{b\nu_\beta(1)}{a\nu_\beta(0) + b\nu_\beta(1)} & \text{, if } x = 1 \end{cases}$$

and we deduce from (55)

$$\forall\, x \in \{0,1\}, \qquad \zeta_\beta(x) \sim \begin{cases} \dfrac{ac}{ac + b^2} & \text{, if } x = 0 \\ \dfrac{b^2}{ac + b^2} & \text{, if } x = 1 \end{cases}$$

For any $\beta \geqslant 0$, we also have, with the notation of the introduction,

$$L^\dagger_{\beta,0} = (-c - d\nu_\beta(0))$$

$$L^\dagger_{\beta,1} = (-b - a\nu_\beta(1))$$

(where the r.h.s. are seen as 1×1-matrices), so that, for large $\beta \geqslant 0$,

$$\begin{array}{rcl} \Sigma_0^\dagger(\beta) & = & \dfrac{1}{c + d\nu_\beta(0)} \;\sim\; \dfrac{1}{c} \\ \Sigma_1^\dagger(\beta) & = & \dfrac{1}{b + a\nu_\beta(1)} \;\sim\; \dfrac{c}{b(c+a)} \;\sim\; \dfrac{1}{b} \end{array}$$

and thus

$$\begin{array}{rcl} \Sigma^\dagger(\beta) & = & \Sigma_0^\dagger(\beta)\zeta_\beta(0) + \Sigma_1^\dagger(\beta)\zeta_\beta(1) \\ & \sim & \dfrac{1}{c}\dfrac{ac}{ac+b^2} + \dfrac{1}{b}\dfrac{b^2}{ac+b^2} \\ & = & \dfrac{a+b}{ac+b^2} \end{array}$$

leading to

$$\lim_{\beta\to+\infty} \beta^{-1} \ln\left(\Sigma^\dagger(\beta)\right) \;=\; (\widetilde{a}+\widetilde{c}) \wedge (2\widetilde{b}) - \widetilde{a} \wedge \widetilde{b}$$

Taking into account (53), we get

$$\lim_{\beta\to+\infty} \beta^{-1} \ln\left(\lambda_0(\beta)\Sigma^\dagger(\beta)\right) \;=\; \delta$$

with

$$\delta \;\coloneqq\; (\widetilde{a}+\widetilde{c}) \wedge (2\widetilde{b}) - \widetilde{a} \wedge \widetilde{b} - \widetilde{a} \wedge (\widetilde{b} - \widetilde{c} + \widetilde{d})$$

In Case (1), we have already seen that the last term in the r.h.s. is $\widetilde{a}$, furthermore we have $\widetilde{a}+\widetilde{c} \leqslant 2\widetilde{b}$ and $\widetilde{a} \leqslant \widetilde{b}$ so that

$$\begin{array}{rcl} \delta & = & \widetilde{a}+\widetilde{c}-\widetilde{a}-\widetilde{a} \\ & = & \widetilde{c}-\widetilde{a} \\ & < & 0 \end{array}$$

In Case (4), the only clear inequality is $\widetilde{b} \leqslant \widetilde{a}$, so that

$$\begin{array}{rcl} \delta & \coloneqq & (\widetilde{a}+\widetilde{c}) \wedge (2\widetilde{b}) - \widetilde{b} - \widetilde{a} \wedge (\widetilde{b} - \widetilde{c} + \widetilde{d}) \\ & = & (\widetilde{a}+\widetilde{c}-\widetilde{b}) \wedge \widetilde{b} - \widetilde{a} \wedge (\widetilde{b} - \widetilde{c} + \widetilde{d}) \end{array}$$

Let us consider two subcases:

- When $\widetilde{a} < \widetilde{b} - \widetilde{c} + \widetilde{d}$, then $\widetilde{a}+\widetilde{c}-\widetilde{b} < \widetilde{b}$, so

$$\begin{array}{rcl} \delta & = & \widetilde{a}+\widetilde{c}-\widetilde{b}-\widetilde{a} \\ & = & \widetilde{c}-\widetilde{b} \;<\; 0 \end{array}$$

- When $\widetilde{a} > \widetilde{b} - \widetilde{c} + \widetilde{d}$, then $\widetilde{a}+\widetilde{c}-\widetilde{b} > \widetilde{b}$, so

$$\begin{array}{rcl} \delta & = & \widetilde{b} - (\widetilde{b} - \widetilde{c} + \widetilde{d}) \\ & = & \widetilde{c} - \widetilde{d} \;<\; 0 \end{array}$$

It both subcases, we get $\delta < 0$, as desired.

■

It is time to discuss about Assumption (50). Consider for example the case where $\widetilde{a} = \widetilde{d}$, and more demandingly, let us assume that $a = d$. Then whatever the initial distribution on $\{0, 1\}$, the exit time is an exponential distribution of parameter a, in particular $\lambda_0(\beta) = a$. It follows that the l.h.s. in the bound of Theorem 1 is zero, while the r.h.s. is positive. This r.h.s. may even be non-vanishing for large $\beta \geqslant 0$. Indeed, note that as soon as $\widetilde{b} \wedge \widetilde{c} > \widetilde{a} = \widetilde{c}$, then the exit position will strongly depend on the initial state: up to an exponential small error, starting from 0 (respectively 1), the process will exit by -1 (resp. 2). Thus from Corollary 4, we have

$$\liminf_{\beta\to\infty} \lambda_0(\beta)\Sigma^{\dagger}(\beta) \quad > \quad 0$$

These observations show that Theorem 1 is not optimal, in the logarithmic scale at small temperature, while we believe that Theorem 3 and Corollary 4 are. As it was mentioned in the introduction, the latter two results do stand for metastability, but not Theorem 1, which is only concerned with the exit time.

Let us now illustrate the difference between the estimates of Theorem 5 and Proposition 6 in the 3-point state space $V \coloneqq \{0, 1, 2\}$. Assume that for all $\beta \geqslant 0$, we are given a birth-and-death Markovian generator

$$\widetilde{L}_\beta \quad \coloneqq \quad \begin{pmatrix} -\widetilde{L}_\beta(0,1) & \widetilde{L}_\beta(0,1) & 0 \\ \widetilde{L}_\beta(1,0) & -\widetilde{L}_\beta(1,0) - \widetilde{L}_\beta(1,2) & \widetilde{L}_\beta(1,2) \\ 0 & \widetilde{L}_\beta(2,1) & -\widetilde{L}_\beta(2,1) \end{pmatrix} \tag{58}$$

As in the above subMarkovian situation, we assume the existence of the following limits

$$\forall\ (x,y) \in V \times V,\ x \neq y, \qquad \widetilde{W}(x,y) \quad \coloneqq \quad -\lim_{\beta\to+\infty} \frac{1}{\beta} \ln(\widetilde{L}_\beta(x,y)) \ \in \ [0, +\infty] \tag{59}$$

and simplify the notations by defining

$$\begin{array}{llllllllllll} a & \coloneqq & \widetilde{L}_\beta(0,1), & b & \coloneqq & \widetilde{L}_\beta(1,0), & c & \coloneqq & \widetilde{L}_\beta(1,2), & d & \coloneqq & \widetilde{L}_\beta(2,1) \\ \widetilde{a} & \coloneqq & \widetilde{W}(0,1), & \widetilde{b} & \coloneqq & \widetilde{W}(1,0), & \widetilde{c} & \coloneqq & \widetilde{W}(1,2), & \widetilde{d} & \coloneqq & \widetilde{W}(2,1) \end{array}$$

For $\beta \geqslant 0$, denote $\widetilde{\pi}_\beta$ the associated reversible probability measure. It is well-known (see for instance Chapter 6 of Freidlin and Wentzell [9]) that the following limits exist:

$$\forall\ x \in V, \qquad U(x) \quad \coloneqq \quad -\lim_{\beta\to+\infty} \beta^{-1} \ln(\widetilde{\pi}_\beta(x)) \tag{60}$$

and the function $U : V \to \mathbb{R}_+$ is called the quasi-potential, it only depends on $\widetilde{W}$ (through finite minimization problems over covering trees).

Let us assume the following inequalities:

$$\widetilde{a} \ > \ \widetilde{b}, \quad \widetilde{c} \ > \ \widetilde{b}, \quad \widetilde{b} \ > \ \widetilde{d} \tag{61}$$

It follows that

$$U(0) \ = \ 0, \quad U(1) \ = \ \widetilde{a} - \widetilde{b}, \quad U(2) \ = \ \widetilde{a} + \widetilde{c} - (\widetilde{b} + \widetilde{d})$$

Let us come back to the setting of Section 7, where all notions now depend on $\beta \geqslant 0$. We consider the case where $\omega = 2$, namely we are interested in φ_β, the solution to the Poisson equation:

$$\begin{cases} \widetilde{L}_\beta[\varphi_\beta] & = & \mathbb{1}_{\{2\}} - \widetilde{\pi}_\beta(2) \\ \widetilde{\pi}_\beta[\varphi_\beta] & = & 0 \end{cases} \tag{62}$$

Let $\Sigma^\dagger(\beta)$ and $\widetilde{\Sigma}^\dagger_*(\beta)$ be the quantities appearing respectively in Theorem 5 and Proposition 6. Since these quantities come from 2×2 and 3×3 matrices, it is clear that for large $\beta \geqslant 0$, $\Sigma^\dagger(\beta)$ has the same logarithmic behavior as the inverse of the first Dirichlet eigenvalue of $(\widetilde{L}_\beta(x,y))_{x,y\in\{0,1\}}$ and $\widetilde{\Sigma}^\dagger_*(\beta)$ has the same logarithmic behavior as the inverse of the spectral gap of $\widetilde{L}_\beta$. Direct computations then lead to

$$\begin{aligned} \lim_{\beta\to+\infty} \beta^{-1}\ln\left(\Sigma^\dagger(\beta)\right) &= \widetilde{a}+\widetilde{c}-\widetilde{b} \\ \lim_{\beta\to+\infty} \beta^{-1}\ln\left(\Sigma^\dagger_*(\beta)\right) &= \widetilde{b} \end{aligned}$$

Remark 27 Such results can also be obtained without computations by extending the path method of Holley and Stroock [11] to the situation "without potential", as in [14], and to the absorbing situation, as in [2].

□

Taking into account that $U(2)=\widetilde{a}+\widetilde{c}-(\widetilde{b}+\widetilde{d})$, we deduce the following behaviors for the bounds of Theorem 5 and Proposition 6:

Proposition 28 *Under the above assumptions, in particular* (61)*, we have*

$$\begin{aligned} \lim_{\beta\to+\infty} \beta^{-1}\ln\left(\widetilde{\pi}_\beta(2)\Sigma^\dagger(\beta)\right) &= \widetilde{d} \\ \lim_{\beta\to+\infty} \beta^{-1}\ln\left(\Sigma^\dagger_*(\beta)\right) &= \widetilde{b} \end{aligned}$$

Since $\widetilde{d}<\widetilde{b}$, for large $\beta\geqslant 0$, the bound of Theorem 5 is much better than that of Proposition 6.

Remark 29 Despite Remark 25, the fact that (37) is based on Proposition 6 rather than on Theorem 5 is a first suggestion that the bound of Theorem 1 should be better than (37). Let us give here an instance at small temperature, by modifying the above three-point example.

For $\beta\geqslant 0$, consider the subMarkov generator L_β defined on $V\coloneqq\{0,1,2\}$ as in (58), except that the underlying process is killed at 2 with rate $L_\beta(2,\infty)$:

$$L_\beta \;\coloneqq\; \begin{pmatrix} -L_\beta(0,1) & L_\beta(0,1) & 0 \\ L_\beta(1,0) & -L_\beta(1,0)-L_\beta(1,2) & L_\beta(1,2) \\ 0 & L_\beta(2,1) & -L_\beta(2,1)-L_\beta(2,\infty) \end{pmatrix} \tag{63}$$

We assume the existence of the following limits (recall that $\breve{V}=\{0,1,2,\infty\}$),

$$\forall\ (x,y)\in V\times\breve{V},\ x\neq y, \qquad W(x,y) \;\coloneqq\; -\lim_{\beta\to+\infty}\frac{1}{\beta}\ln(L_\beta(x,y)) \;\in\; [0,+\infty]$$

simplify the notations by defining

$$\widetilde{a} \;\coloneqq\; \widetilde{W}(0,1),\quad \widetilde{b} \;\coloneqq\; \widetilde{W}(1,0),\quad \widetilde{c} \;\coloneqq\; \widetilde{W}(1,2),\quad \widetilde{d} \;\coloneqq\; \widetilde{W}(2,1),\quad \widetilde{e} \;\coloneqq\; \widetilde{W}(2,\infty)$$

and assume

$$\widetilde{a} \;>\; \widetilde{b} \;>\; 0,\quad \widetilde{c}\;>\;\widetilde{b},\quad \widetilde{d}\;=\;0,\quad \widetilde{e}\;>0 \tag{64}$$

Taking into account the results of [15], metastability at small temperature holds and we have

$$\lim_{\beta\to+\infty}\frac{1}{\beta}\ln(\lambda_0(\beta)) \;=\; -(\widetilde{a}-\widetilde{b}+\widetilde{c}+\widetilde{e}) \tag{65}$$

(it is due to the fact that $\{0,1,2\}$ can be seen as a well of height $\widetilde{a}-\widetilde{b}+\widetilde{c}+\widetilde{e}>0$ in a larger state space (for instance $\breve{V}$ by adding an exponential transition from ∞ to 2) for a reversible Markov generator at small temperature). Furthermore, L_β admits another exponentially small eigenvalue at small temperature, say $\lambda_1(\beta)$, which satisfies

$$\lim_{\beta\to+\infty}\frac{1}{\beta}\ln(\lambda_1(\beta)) \quad = \quad -\widetilde{b}$$

An easy way to get the upper bound, which is the only thing needed in the following arguments, is to apply the variational principle to the vector space generated by $\mathbb{1}_{\{0\}}$ and $\mathbb{1}_{\{0\}}$ in $\mathbb{L}^2(\pi_\beta)$, where π_β is the reversible probability measure associated to the Markovian generator obtained by removing $L_\beta(2,\infty)$ from (63).

Now let us come back to (37). With the corresponding notations, we have

$$\begin{aligned}\lim_{\beta\to+\infty}\frac{1}{\beta}\ln(\Sigma_*(\beta)) \quad &= \quad -\lim_{\beta\to+\infty}\frac{1}{\beta}\ln(\lambda_1(\beta))\\ &= \quad \widetilde{b}\end{aligned}$$

and

$$\begin{aligned}\lim_{\beta\to+\infty}\frac{1}{\beta}\ln\left(\sum_{\omega\in\delta V}\left(|L_\beta(\omega,\omega)|-\sum_{y\in V}L_\beta(\omega,y)\right)\right) \quad &= \quad \lim_{\beta\to+\infty}\frac{1}{\beta}\ln(L_\beta(2,\infty))\\ &= \quad -\widetilde{e}\end{aligned}$$

Thus metastability is not recovered as soon as $\widetilde{e}<\widetilde{b}$, at least under (64).

To be able to apply Theorem 1, we must first understand the behavior at small temperature of the quasi-stationary measure ν_β. In the present particular example (be careful, this is not always true, see the counter-example closing this section), it can be checked the logarithmic behavior of ν_β is the same as for the invariant measure π_β, and thus we get

$$\lim_{\beta\to+\infty}\beta^{-1}\ln(\nu_\beta(0)) \;=\; 0,\quad \lim_{\beta\to+\infty}\beta^{-1}\ln(\nu_\beta(1)) \;=\; -(\widetilde{a}-\widetilde{b}),\quad \lim_{\beta\to+\infty}\beta^{-1}\ln(\nu_\beta(2)) \;=\; -(\widetilde{a}+\widetilde{c}-\widetilde{b})$$

We deduce, with the notations of the introduction,

$$\begin{aligned}\lim_{\beta\to+\infty}\beta^{-1}\ln(\Sigma^\dagger(\beta)) \quad &= \quad \lim_{\beta\to+\infty}\beta^{-1}\ln(\Sigma_2^\dagger(\beta))\\ &= \quad \widetilde{a}+\widetilde{c}-\widetilde{b}\end{aligned}$$

Comparing this convergence with (65), we get that Theorem 1 enables to recover the metastability phenomenon under (64), without the restriction $\widetilde{e}<\widetilde{b}$.

□

To finish, let us briefly consider the extension of the above small temperature considerations to arbitrary finite state space V. We assume that we are given a family $(L_\beta)_{\beta\geqslant 0}$ of irreducible strictly subMarkovian generators on V and that (49) holds in $[0,+\infty]$ (with $\breve{V}\coloneqq V\sqcup\{\infty\}$ and $L_\beta(x,\infty)\coloneqq -\sum_{y\in V}L_\beta(x,y)$, for all $x\in V$).

There is no difficulty with the behavior of $\lambda_0(\beta)$, as we know the validity of

$$\lim_{\beta\to+\infty}\beta^{-1}\ln(\lambda_0(\beta)) \quad = \quad -l$$

where l is the highest depth of a well included in V, for the appropriate definitions of the energy landscape in this context.

More problematic and surprising at first view, is the behavior of the quasi-stationary measure ν_β for large $\beta \geqslant 0$, since the existence of the limits

$$\forall\, x \in V, \qquad W(x) \;:=\; \lim_{\beta\to+\infty} \beta^{-1}\ln(\nu_\beta(x)) \tag{66}$$

is not always true, and when they exist, they may not depend only on $(W(x,y))_{(x,y)\in V\times\check{V}}$. Thus the situation is quite different from the existence of the quasi-potential as in (60), which always exists for irreducible Markovian generators at small temperature (and only depend on the exponential rates of the transitions). Nevertheless, we think that the limits in (66), that could be called "quasi-quasi-potentials", generically exist and only depend on the rates $(W(x,y))_{(x,y)\in V\times\check{V}}$, i.e. up to removing exceptional identities between these rates, as in (50). Maybe the non-validity of (66) is a watered-down instance in the finite setting of the non-uniqueness of quasi-stationary measures in general (see e.g. Example 6.3.1 from Collet, Martínez and San Martín [3]), due to the non-linearity of the equation they solve. We hope to be able to investigate more thoroughly this situation in a future work.

For the moment being, let us conclude by giving a counter-example to (66), in the two-point state space $\{0,1\}$. We begin by a simple computation:

Lemma 30 *Assume that* (49) *holds, with $d > a$ for all $\beta \geqslant 0$, and*

$$\lim_{\beta\to+\infty} \beta^{-1}\ln(d-a) \;=\; -\widetilde{r}$$

with $\widetilde{r} < \widetilde{b} \wedge \widetilde{c}$. Then we get

$$\lim_{\beta\to+\infty} \beta^{-1}\ln(\nu_\beta(1)) \;=\; \widetilde{r} - \widetilde{b}$$

Proof

Let us come back to (56), taking into account that $\nu_\beta(1) = 1 - \nu_\beta(0)$, we get that $x := \nu_\beta(0)$ is solution of the second order equation

$$(d-a)x^2 + (a+b+c-d)x - c \;=\; 0$$

Its solutions are

$$x_\pm \;=\; \frac{1}{2(d-a)}(d-(a+b+c) \pm \sqrt{\triangle})$$

with

$$\triangle \;:=\; (a+b+c-d)^2 + 4c(d-a)$$

Note that the product of these solutions is $-c/(d-a) < 0$, so $\nu_\beta(0)$ is the positive solution: $\nu_\beta(0) = x_+$. We deduce

$$\nu_\beta(0) \;=\; \frac{1}{2(d-a)}(d-(a+b+c) + \sqrt{\triangle})$$

and

$$\begin{aligned}
\nu_\beta(1) &= 1 - \nu_\beta(0)\\
&= \frac{1}{2(d-a)}(d-a+b+c-\sqrt{\triangle})\\
&= \frac{1}{2(d-a)}\frac{(d-a+b+c)^2 - (a+b+c-d)^2 - 4c(d-a)}{d-a+b+c+\sqrt{\triangle}}\\
&= \frac{1}{2(d-a)}\frac{4(d-a)(b+c) - 4c(d-a)}{d-a+b+c+\sqrt{\triangle}}\\
&= \frac{2b}{d-a+b+c+\sqrt{\triangle}}
\end{aligned}$$

Due to $\widetilde{r} < \widetilde{b} \wedge \widetilde{c}$, we have that for large $\beta \geqslant 0$, $b + c \ll d - a$, so that $\nu_\beta(1) \sim b/(d-a)$ and the announced result follows at once.

■

Choose $0 \leqslant \widetilde{a} < \widetilde{r} < \widetilde{b}$ and define for all $\beta \geqslant 0$,

$$a \coloneqq \exp(-\widetilde{a}\beta), \quad b \coloneqq \exp(-\widetilde{b}\beta), \quad c \coloneqq b, \quad d \coloneqq a + \exp(-\widetilde{r}\beta)$$

The conditions of Lemma 30 are satisfied and we get

$$\lim_{\beta\to+\infty} \beta^{-1}\ln(\nu_\beta(1)) = \widetilde{r} - \widetilde{b} \tag{67}$$

where the r.h.s. is not a function of the exponential rates $\widetilde{a} = \widetilde{d}$ and $\widetilde{b} = \widetilde{c}$.

To get the desired counter-example, choose $\widetilde{a} < \widehat{a} < \widehat{b} < \widetilde{b}$ and consider a function $\widetilde{r} : \mathbb{R}_+ \ni \beta \mapsto \widetilde{r}(\beta) \in [\widehat{a}, \widehat{b}]$ with

$$\liminf_{\beta\to+\infty} \widetilde{r}(\beta) = \widehat{a}, \qquad \limsup_{\beta\to+\infty]} \widetilde{r}(\beta) = \widehat{b}$$

From the proof of Lemma 30, we deduce

$$\liminf_{\beta\to+\infty} \nu_\beta(1) = \widehat{a} - \widetilde{b}, \qquad \limsup_{\beta\to+\infty} \nu_\beta(1) = \widehat{b} - \widetilde{b}$$

Acknowledgments: I'm thankful to the IFCAM for the support to the project *Metastability in complex systems* in which took place this research.

References

[1] Anton Bovier and Frank den Hollander. *Metastability*, volume 351 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2015. A potential-theoretic approach.

[2] Daniel Bump, Persi Diaconis, Angela Hicks, Laurent Miclo, and Harold Widom. An exercise(?) in Fourier analysis on the Heisenberg group. *Ann. Fac. Sci. Toulouse Math. (6)*, 26(2):263–288, 2017.

[3] Pierre Collet, Servet Martínez, and Jaime San Martín. *Quasi-stationary distributions*. Probability and its Applications (New York). Springer, Heidelberg, 2013. Markov chains, diffusions and dynamical systems.

[4] Giacomo Di Gesù, Tony Lelièvre, Dorian Le Peutrec, and Boris Nectoux. Sharp asymptotics of the first exit point density. *Ann. PDE*, 5(1):Art. 5, 174, 2019.

[5] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.*, 93(4):1659–1664, 1996.

[6] Persi Diaconis and James Allen Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.

[7] Persi Diaconis and Laurent Miclo. On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.*, 22(3):558–586, 2009.

[8] James Allen Fill. The passage time distribution for a birth-and-death chain: strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.*, 22(3):543–557, 2009.

[9] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1984. Translated from the Russian by Joseph Szücs.

[10] Asma Hassannezhad and Laurent Miclo. Higher order Cheeger inequalities for Steklov eigenvalues. arXiv e-print, May 2017. To appear in *Annales Scientifiques de l'École Normale Supérieure.*

[11] Richard Holley and Daniel Stroock. Simulated annealing via Sobolev inequalities. *Comm. Math. Phys.*, 115(4):553–569, 1988.

[12] Tosio Kato. *Perturbation theory for linear operators.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[13] Francesco Manzo and Elisabetta. Scoppola. Exact results on the first hitting via conditional strong quasi-stationary times and applications to metastability. *Journal of Statistical Physics*, Jan 2019.

[14] Laurent Miclo. Recuit simulé sans potentiel sur un ensemble fini. In *Séminaire de Probabilités, XXVI*, volume 1526 of *Lecture Notes in Math.*, pages 47–60. Springer, Berlin, 1992.

[15] Laurent Miclo. Sur les problèmes de sortie discrets inhomogènes. *Ann. Appl. Probab.*, 6(4):1112–1156, 1996.

[16] Laurent Miclo. On absorption times and Dirichlet eigenvalues. *ESAIM Probab. Stat.*, 14:117–150, 2010.

[17] Laurent Miclo. An absorbing eigentime identity. *Markov Process. Related Fields*, 21(2):249–262, 2015.

[18] Laurent Miclo. On the Markovian similarity. In *Séminaire de Probabilités XLIX*, volume 2215 of *Lecture Notes in Math.*, pages 375–403. Springer, Cham, 2018.

† miclo@math.cnrs.fr

Institut de Mathématiques de Toulouse
Université Paul Sabatier, 118, route de Narbonne
31062 Toulouse cedex 9, France

Toulouse School of Economics,
1, Esplanade de l'université
31080 Toulouse cedex 06, France