



On soft capacities, quasi-stationary distributions and the pathwise approach to metastability

Alexandre Gaudilliere, A. Bianchi, P Milanesi

► To cite this version:

Alexandre Gaudilliere, A. Bianchi, P Milanesi. On soft capacities, quasi-stationary distributions and the pathwise approach to metastability. Journal of Statistical Physics, 2020. hal-01850845v2

HAL Id: hal-01850845

<https://hal.science/hal-01850845v2>

Submitted on 30 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On soft capacities, quasi-stationary distributions and the pathwise approach to metastability

A. Bianchi*

A. Gaudillière†

P. Milanesi‡

Abstract

Motivated by the study of the metastable stochastic Ising model at subcritical temperature and in the limit of a vanishing magnetic field, we extend the notion of (κ, λ) -capacities between sets, as well as the associated notion of soft-measures, to the case of *overlapping* sets. We recover their essential properties, sometimes in a stronger form or in a simpler way, relying on weaker hypotheses. These properties allow to write the main quantities associated with reversible metastable dynamics, e.g. asymptotic transition and relaxation times, in terms of objects that are associated with two-sided variational principles. We also clarify the connection with the classical “pathwise approach” by referring to temporal means on the appropriate time scale.

MSC 2010: primary: 60J27, 60J45, 60J75; secondary: 82C20.

Keywords: Soft capacity, soft measure, quasi-stationary measure, restricted ensemble, metastability, potential theory, relaxation time.

Acknowledgements: A. G. and P. M. thank Maria Eulalia Vares, the Universidade federal do Rio de Janeiro and the Università di Padova for the kind hospitality which gave us the opportunity to lay the foundations of this work.

1 Model and main results

In the present paper we consider a general Markovian model for a metastable dynamics and we show how, under mild hypotheses, “soft measures and capacities” associated with a covering of the configuration space allow for a description of the metastable state provide sharp estimates of the relaxation time, fast relaxation to local equilibria, and enable to establish the asymptotic exponential law of the transition time to equilibrium. The ultimate goal of this paper is to provide a mathematical framework to prove a convergence in law to an exponential distribution of the transition time to equilibrium for the metastable kinetic Ising model at any subcritical temperature and in the limit of a vanishing magnetic field. In the companion paper [GMV20], we establish such a convergence by working out the model-dependent part of the proof. In addition, we present an explicit connection with the classical “pathwise approach to metastability”, and we provide a comparison with the techniques appeared in the recent literature of abstract metastable dynamics.

Before stating the main results, and to better illustrate the ideas leading to the present research, we start by discussing the qualitative behavior of a concrete Markovian model. We consider a continuous time Glauber dynamics in two-dimensional finite boxes of area diverging as $1/h^2$, where $h \ll 1$ is the magnetic field, as studied in [GMV20]. By Glauber dynamics we mean a single spin flip dynamics that is reversible with respect to the Gibbs measure of the Ising model, thus including Metropolis and heat-bath dynamics (see [GMV20] for precise definitions). Figure 1 shows a sample of such a Metropolis dynamics started from the “all minus configuration”, with periodic boundary conditions, and with an Hamiltonian that is twice¹ the Hamiltonian of [GMV20], up to boundary conditions.

*Università degli Studi di Padova, Dipartimento di Matematica, Via Trieste, 63, 35121 Padova, Italy, e-mail: bianchi@math.unipd.it

†Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France, e-mail: alexandre.gaudilliere@math.cnrs.fr

‡Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Marseille, France, e-mail: paolo.milanesi@univ-amu.fr

¹This is because [GMV20] sticks to the conventions of its main reference [SS98], where the authors introduced a relatively unusual factor $1/2$ in the Hamiltonian. Since Figure 1 illustrates a dynamics ran at inverse temperature $\beta = 2/3$ without such a convention, it would correspond to a trajectory sampled at inverse temperature $\beta' = 4/3$, i.e., to a twice smaller temperature, with the convention of [SS98] and [GMV20].

Figure 1: Snapshots of a kinetic Ising model at temperature $T = 1.5$ and under magnetic field $h = 0.14$ in a 256×256 box with periodic boundary conditions. Minus spins are yellow and plus spins are red. We took pictures at times 471, 7482 and 13403 for the first line, at times 14674, 15194, 15432, 15892 and 16558 for the second line, and times 17328, 23645 and 40048 for the last line.

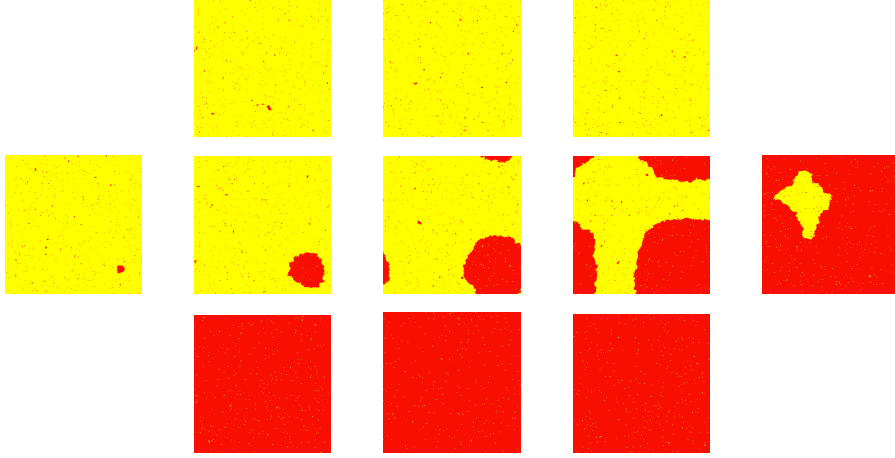


Figure 1 is the kind of picture to bear in mind when reading the present paper. Here are some of its most relevant features.

- The last three pictures can be seen as different samples from the equilibrium Gibbs measure. Since $h > 0$, it is concentrated on configurations where plus spins dominate. One has to wait for some “transition to equilibrium” to start observing non-anomalous configurations with respect to this equilibrium state.
- The first three pictures can be seen as different samples from a single distribution concentrated on configurations where minus spins prevail. Notice that they represent typical configurations of a “local equilibrium”, or metastable state, that is very different from the equilibrium stable state.
- The transition to equilibrium is triggered by the nucleation of a “supercritical droplet” (in the first picture, see the tiny plus-phase droplet that was not large enough to trigger it). Notice that the time needed to invade the whole box is of a smaller order than the time needed to appear —about 2000 and 15000 time units, respectively— as a fluctuation of the metastable state.
- This suggests that the time needed by the system to relax to some local equilibrium (metastable or stable state) is short with respect to the global relaxation time. This is what we call “fast relaxation to local equilibria”.
- This also suggests that the transition to equilibrium can be described as “both late and abrupt”, and it is thus compatible with an asymptotic exponential law of the transition time to equilibrium.

Topping the list of questions posed when trying to build a rigorous mathematical counterpart of such observations:

- i. How can we make sense of such a notion of local or metastable equilibrium? How should it be described?
- ii. How can we define such a random transition time to equilibrium? How should its law be described?

The pathwise approach to metastability introduced in [CGOV84] and fully described in [OV05] suggests using temporal means and stopping times to answer these questions. If

$$X : t \geq 0 \mapsto X(t) \in \mathcal{X}$$

is the Markov process, the metastable character of which we want to investigate, and $f : \mathcal{X} \rightarrow \mathbb{R}$ is an observable, let us denote the associated temporal mean from time $t \geq 0$ and on time scale $\theta > 0$ by

$$A_\theta(t, f) = \frac{1}{\theta} \int_t^{t+\theta} f(X(s)) ds. \quad (1)$$

Let μ be the equilibrium distribution of the process and, for any probability measure ν on \mathcal{X} and any observable f , let us denote by $\nu(f)$ the mean value of f with respect to ν . Within the pathwise approach, answering the previous questions amounts to identifying a probability measure $\tilde{\mu}$, a stopping time T and a deterministic time scale θ such that, for the system started in $\tilde{\mu}$:

- θ is small with respect to $\mathbb{E}_{\tilde{\mu}}[T]$, the mean value of T ;
- $A_\theta(t, f)$ is close to $\tilde{\mu}(f)$ for all $t < T - \theta$ and all observable f with very large probability;
- $A_\theta(t, f)$ is close to $\mu(f)$ for all observables f and all $t \geq T$ of the same order as $\mathbb{E}_{\tilde{\mu}}[T]$, or possibly larger, with very large probability.

The metastable state and the transition time to equilibrium can then be defined by $\tilde{\mu}$ and T , respectively. The pathwise approach also focuses on establishing the convergence in law to an exponential distribution with parameter one of the rescaled transition time $T/\mathbb{E}_{\tilde{\mu}}[T]$. This is to account for a “both late and abrupt” transition to equilibrium, by opposition with the behaviour of a slowly relaxing system for which this relaxation would occur by following a gentle and constant drift². One cannot distinguish between these two behaviours by looking only at the time evolution of the mean value

$$t \geq 0 \longmapsto \mathbb{E}_{\tilde{\mu}}[f(X(t))], \quad f : \mathcal{X} \longrightarrow \mathbb{R},$$

that are associated with the natural Markovian semi-group; hence the terminology “*pathwise* approach”.

As far as the construction of the probability measure $\tilde{\mu}$ and the stopping time T is concerned, the authors of [CGOV84] suggested, by following [LP71], to use a *restricted ensemble*

$$\mu_{\mathcal{R}} = \mu(\cdot \mid \mathcal{R})$$

and the associated exit time $T_{\mathcal{X} \setminus \mathcal{R}}$, for some subset \mathcal{R} of the configuration space \mathcal{X} that can be thought of as a “basin of attraction” of the metastable state. In the case of the kinetic Ising model, Schonmann and Sholsman made a detailed study in the beautiful paper [SS98] of such restricted ensemble and exit time by defining \mathcal{R} as the set of configurations for which the $+$ spins can be enclosed into small enough (subcritical) contours (see [SS98] or the companion paper [GMV20] for more details). Their results allow to prove that the temporal means $A_\theta(t, f)$, for a suitable θ , are close to $\mu_{\mathcal{R}}(f)$ or $\mu(f)$, but only for times t that are very small or very large with respect to $\mathbb{E}_{\mu_{\mathcal{R}}}[T_{\mathcal{X} \setminus \mathcal{R}}]$, which is strongly related with the lack of control of the law of the transition time to equilibrium.

In [BG16] we proposed some modifications of the traditional construction of the metastable state $\tilde{\mu}$ and the transition time to equilibrium T . We worked with quasi-stationary measures to connect the two issues of the description of the metastable state and of the law of the transition time. This led us first to a mathematical formalization of the phenomenon of “local thermalization” or “fast relaxation to local equilibria”, then to conceive of the stable state as a more stable local equilibrium only. As a consequence, we suggested to use a stopping time at which one is ensured³ to be close to this more stable equilibrium, instead of the exit time $T_{\mathcal{X} \setminus \mathcal{R}}$. This eventually led us to the notion of soft measure, which is a quasi-stationary measure (of a trace process) associated with such a stopping time rather than a plain exit time: while quasi-stationary measures are associated with absorption as soon as a Markov process hits some set \mathcal{S} , which acts as an *hard* barrier, we consider absorption, or killing, at some rate λ in \mathcal{S} —our Markov process can then penetrate \mathcal{S} , hence the terminology “*soft* measure”. We showed that, with a suitable set of hypotheses in which \mathcal{S} can be envisaged as a basin of attraction of the stable state, and for a suitable choice of the killing parameter λ , the associated soft measure $\mu_{\mathcal{R}, \lambda_{\mathcal{S}}}^*$ and killing time $T_{\lambda_{\mathcal{S}}}$ (see sections 1.1 and 2.1 for precise definitions) were natural candidates to build the metastable state $\tilde{\mu}$ and the transition time T . This came with an associated notion of “soft capacity” and a solution, as detailed in [BG16] and shortly reviewed in Section 1.2 of the present paper, to difficulties encountered in proving asymptotic exponential laws for transition times and that had also emerged in the “potential-theoretic approach” to metastability.

²In the case of the kinetic Ising model, one can instead think of two opposite strong drifts: one towards the metastable state and one towards the stable state. The nucleation of the supercritical droplet occurs by fluctuation *against* the first drift and the system follows the second drift afterwards. This is coherent with the fast relaxation to local equilibria.

³Note that for the kinetic Ising model, despite the strong drift towards equilibrium at the appearance of the first supercritical droplet, there is still some time to wait before the relaxation to be achieved: the pictures on the second line of Figure 1 show atypical configurations with respect to the equilibrium measure. The exit from \mathcal{R} does not coincide with global thermalization.

Unfortunately, we could not prove the convergence in law of a rescaled transition time to an exponential random variable in the aforementioned case of the kinetic Ising model within the framework of [BG16]. The problem came from the fact that, in [BG16], we assumed that the sets \mathcal{R} and \mathcal{S} form a partition of the configuration space. But as shown in [GMV20], the tools to control local relaxation times, that is the crucial point of the model-dependent part of the study, involve some mesoscopic and macroscopic analysis that is incompatible with such a neat separation at the microscopic level. It turned out that removing the disjointness hypothesis of \mathcal{R} and \mathcal{S} from the framework of [BG16], with the aim of improving the global proposed strategy and eventually prove an asymptotic exponential law in [GMV20], required new ideas. This was the first motivation for writing the present paper. In doing so, we introduced some simplifications and strengthened some estimates with respect to [BG16], and we also provided an explicit connection with the classical pathwise approach.

At last, we have to make a final remark on the asymptotic exponential law for the kinetic Ising model. We prove such an asymptotic result in the companion paper [GMV20], where we consider the kinetic Ising model in a finite box of area diverging as $1/h^2$, in the limit of a vanishing magnetic field $h > 0$. Even though the proof of such a result was missing, this was to be expected and nothing to be surprised at. But in the case of infinite volume dynamics, as the one considered in [SS98], the situation is more delicate, and it is not so clear whether an asymptotic exponential law for the transition time to equilibrium is the correct conjecture for the both late and abrupt transition. Indeed, as shown by Shonmann and Shlosman, the transition mechanism of infinite volume dynamics does not only involve the nucleation of supercritical droplets, which should occur after an asymptotic exponential time, but also the coalescence of such growing droplets on a time scale that is logarithmically equivalent to the nucleation time scale. Thus, to decide whether the asymptotic exponential law still holds when the times involved in both mechanisms are sequentially added, a precise control on the prefactors of their mean values is needed. This underlines the relevance of the sharp estimates of spectral gap and mean transition time (see Theorem 1 in Section 1.1) in terms of soft capacities inherited from the classical ones that were introduced by [BEGK02], [BEGK04] and [BEGK05] in metastability studies.

The paper is organized as follows. In Section 1.1 we present our main result, in Section 1.2 we discuss the relations with the more recent literature on abstract metastable dynamics, in Section 2 we collect all the material that can be directly imported from [BG16] and we simplify some of its proofs, and in sections 3, 4 and 5 we give the proofs of our main theorems.

1.1 Local and global thermalizations

Let X be an irreducible continuous time Markov process on a finite configuration space

$$\mathcal{X} = \mathcal{R} \cup \mathcal{S}$$

with \mathcal{R} and \mathcal{S} possibly overlapping and with generator \mathcal{L} defined by

$$(\mathcal{L}f)(x) = \sum_{y \neq x} w(x, y)[f(y) - f(x)], \quad x \in \mathcal{X}, \quad f : \mathcal{X} \rightarrow \mathbb{R}. \quad (2)$$

We assume X to be reversible with respect to a probability measure μ and we denote by \mathcal{D} the Dirichlet form associating with any observable $f : \mathcal{X} \rightarrow \mathbb{R}$

$$\mathcal{D}(f) = \frac{1}{2} \sum_{x, y \in \mathcal{X}} \mu(x) w(x, y) [f(x) - f(y)]^2.$$

We call $1/\gamma$ the relaxation time of the dynamics, so that γ is the spectral gap

$$\gamma = \min_{\substack{f \in \ell^2(\mu), \\ \text{Var}_\mu(f) \neq 0}} \frac{\mathcal{D}(f)}{\text{Var}_\mu(f)}. \quad (3)$$

With any $\mathcal{Y} \subset \mathcal{X}$ we associate the restricted dynamics $X_{\mathcal{Y}}$ in \mathcal{Y} with generator defined by

$$(\mathcal{L}_{\mathcal{Y}}f)(x) = \sum_{y \in \mathcal{Y} \setminus \{x\}} w(x, y)[f(y) - f(x)], \quad x \in \mathcal{Y}, \quad f : \mathcal{Y} \rightarrow \mathbb{R}, \quad (4)$$

which is reversible with respect to the restricted ensemble

$$\mu_{\mathcal{Y}} = \mu(\cdot \mid \mathcal{Y}).$$

If $X_{\mathcal{Y}}$ is also an irreducible process we denote by $\gamma_{\mathcal{Y}}$ its spectral gap. The associated exit rate (see Section 2 for an explanation of the terminology) is

$$\phi_{\mathcal{Y}}^* = \min_{\substack{f \neq 0, \\ f|_{\mathcal{X} \setminus \mathcal{Y}} = 0}} \frac{\mathcal{D}(f)}{E_{\mu}[f^2]}.$$

Let σ_{λ} be an exponential random variable of parameter λ , independent of X , and $\ell_{\mathcal{S}}(t)$ denote the local time of X in \mathcal{S} , that is

$$\ell_{\mathcal{S}}(t) = \int_0^t \mathbf{1}_{\{X(s) \in \mathcal{S}\}} ds.$$

We then define $T_{\lambda_{\mathcal{S}}}$ as the first time $t \geq 0$ such that $\ell_{\mathcal{S}}(t)$ reaches σ_{λ} :

$$T_{\lambda_{\mathcal{S}}} = \min\{t \geq 0 : \ell_{\mathcal{S}}(t) \geq \sigma_{\lambda}\}.$$

In other words, $T_{\lambda_{\mathcal{S}}}$ is the absorption time of the process X killed at rate $\lambda_{\mathcal{S}}$ with

$$\lambda_{\mathcal{S}}(x) = \lambda \mathbf{1}_{\{x \in \mathcal{S}\}}, \quad x \in \mathcal{X}. \quad (5)$$

We define analogously $T_{\kappa_{\mathcal{R}}}$, which is associated with the killing rate $\kappa_{\mathcal{R}}$, from an exponential random variable σ_{κ} of parameter κ , independent of X and σ_{λ} , and by using $\ell_{\mathcal{R}}$, the local time in \mathcal{R} . With

$$V_{\kappa}^{\lambda}(x) = \mathbb{P}_x(T_{\kappa_{\mathcal{R}}} < T_{\lambda_{\mathcal{S}}}), \quad x \in \mathcal{X},$$

the associated soft capacity, or (κ, λ) -capacity is

$$C_{\kappa}^{\lambda}(\mathcal{R}, \mathcal{S}) = \mathcal{D}(V_{\kappa}^{\lambda}) + \kappa \mu(\mathcal{R}) E_{\mu_{\mathcal{R}}} \left[(V_{\kappa}^{\lambda} - 1)^2 \right] + \lambda \mu(\mathcal{S}) E_{\mu_{\mathcal{S}}} \left[(V_{\kappa}^{\lambda} - 0)^2 \right].$$

We assume to be in some asymptotic regime, meaning that the dynamics depends on some parameters, as temperature, volume or magnetic field going to some finite or infinite limit, and we will write $f \ll g$ and $f \sim g$ whenever $0 \leq f \leq g$ and the ratio f/g goes to zero and one, respectively, in the considered asymptotic regime. To state the first main result of this paper we introduce

Hypotheses (H): We assume \mathcal{R} and \mathcal{S} to be such that

1. the associated restricted processes $X_{\mathcal{R}}$, $X_{\mathcal{R} \setminus \mathcal{S}}$, $X_{\mathcal{S}}$ and $X_{\mathcal{S} \setminus \mathcal{R}}$ are all irreducible;
2. $\phi_{\mathcal{R} \setminus \mathcal{S}}^* \ll \gamma_{\mathcal{R}}$ and $\phi_{\mathcal{S} \setminus \mathcal{R}}^* \ll \gamma_{\mathcal{S}}$;
3. $\mu(\mathcal{S}) \geq \mu(\mathcal{R})$ and $\phi_{\mathcal{R} \setminus \mathcal{S}}^* \ll \gamma_{\mathcal{S}}$.

Since exit rates $\phi_{\mathcal{Y}}^*$ are associated with some mean exit time from \mathcal{Y} (see Section 2.1), such hypotheses amount in particular to say that “local relaxation times” $1/\gamma_{\mathcal{Y}}$ are small with respect to these mean exit times. To check (H) we need upper bounds on exit rates and lower bounds on spectral gaps. While the former are usually quite easy to get, since exit times are given by an infimum in some variational principle, getting the latter is the core of the difficulty. This is the crux of the companion paper [GMV20]. However, if the hypothesis set (H) is in force, even from a rough control of local relaxation times we can get sharp estimates of the global one:

Theorem 1. Assuming hypotheses (H) and choosing κ and λ such that

$$\phi_{\mathcal{R} \setminus \mathcal{S}}^* \ll \kappa \ll \gamma_{\mathcal{R}} \quad \text{and} \quad \phi_{\mathcal{R} \setminus \mathcal{S}}^* \vee \phi_{\mathcal{S} \setminus \mathcal{R}}^* \ll \lambda \ll \gamma_{\mathcal{S}},$$

it holds

$$\frac{1}{\gamma} \sim \frac{\mu(\mathcal{R})\mu(\mathcal{S})}{C_{\kappa}^{\lambda}(\mathcal{R}, \mathcal{S})} \quad \text{and} \quad \mathbb{E}_{\mu_{\mathcal{R}}}[T_{\lambda_{\mathcal{S}}}] \sim \frac{\mu(\mathcal{R})}{C_{\kappa}^{\lambda}(\mathcal{R}, \mathcal{S})}.$$

Comments:

- i. This gives a way to compute the precise asymptotic value of the relaxation time $1/\gamma$. As soon as κ and λ are chosen in the right window, the spectral gap γ scales, up to the multiplicative term $\mu(\mathcal{R})\mu(\mathcal{S})$, like the soft capacity $C_\kappa^\lambda(\mathcal{R}, \mathcal{S})$ that satisfies a two-sided variational principle. As we will see in Section 2, any test function and flow will give upper and lower bound for $C_\kappa^\lambda(\mathcal{R}, \mathcal{S})$. These are then translated in upper and lower bounds for γ .
- ii. We already mentioned that the main difficulty in proving the validity of the hypothesis set (H) usually lies in bounding $\gamma_{\mathcal{R}}$ and $\gamma_{\mathcal{S}}$ from below. It is often the case that, once such lower bounds are established, there is a natural way to build a suitable flow to get useful lower bounds of the soft capacity: see [BG16] and [GMV20].
- iii. We will see with the next theorem that $T_{\lambda_{\mathcal{S}}}$ can be identified, under slightly stronger hypotheses, as a transition time to equilibrium if $\mu(\mathcal{R}) \ll \mu(\mathcal{S})$, or to a more stable local equilibrium in the general case $\mu(\mathcal{R}) \leq \mu(\mathcal{S})$. Theorem 1 says in particular that, when starting from the restricted ensemble $\mu_{\mathcal{R}}$, its expected value and the relaxation time have the same asymptotic behaviour up to the multiplicative factor $1/\mu(\mathcal{S})$, which lies between 1 and 2 in the latter case and goes to 1 in the former one. To prove this result we will focus on the local time in \mathcal{R} up to transition

$$T_{\lambda_{\mathcal{S}}}^{\mathcal{R}} = \ell_{\mathcal{R}}(T_{\lambda_{\mathcal{S}}})$$

and its expected value

$$\frac{1}{\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*} = \mathbb{E}_{\mu_{\mathcal{R}, \lambda_{\mathcal{S}}}^*} [T_{\lambda_{\mathcal{S}}}^{\mathcal{R}}]$$

for the dynamics started from the soft measure $\mu_{\mathcal{R}, \lambda_{\mathcal{S}}}^*$ defined in Section 2.1. For λ ranging from 0 to $+\infty$ the soft measure family provides a continuous interpolation from the restricted ensemble $\mu_{\mathcal{R}}$ to the quasi-stationary distribution $\mu_{\mathcal{R} \setminus \mathcal{S}}^*$ associated with the process immediately killed when hitting \mathcal{S} . We will see in Section 2 that hypotheses (H) implies that our soft measures are all close together, each of them providing a metastable state model. And we will prove that

$$E_{\mu_{\mathcal{R}}} [T_{\lambda_{\mathcal{S}}}] \sim \frac{1}{\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}$$

for any λ in the allowed window of Theorem 1.

- iv. The proof of Theorem 1 is deferred to Section 3. In this section we will give explicit upper and lower bounds for $\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*$ and γ that will bound explicitly the multiplicative corrections with respect to the asymptotic values.
- v. About these upper and lower bounds, the latter are the most relevant. Both the exit rate $\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*$ and the spectral gap γ are indeed *infima* given by some variational principles, so that upper bound are simpler to get. In particular, the lower bound given in Section 3 on γ is a Poincaré inequality that can be generalized to the case of more than one metastable state (cf. Proposition 3.1). The upper bounds indicate however that our lower bounds catch the right orders.
- vi. The windows given in Theorem 1 for κ and λ can in principle be enlarged for the conclusion to hold. For example, as far as the lower bound on κ is concerned, the explicit upper and lower bound of Section 3 will scale like the given asymptotic values for $\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*/\kappa \ll 1$. We will see in Section 2 that $\kappa \gg \phi_{\mathcal{R} \setminus \mathcal{S}}^*$ is only a sufficient condition for having $\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*/\kappa \ll 1$. Our hypotheses (H) can be relaxed in the same way for the conclusion to hold.
- vii. When going from the case $\mathcal{R} \cap \mathcal{S} = \emptyset$ of [BG16] to the case $\mathcal{R} \cap \mathcal{S} \neq \emptyset$ of the present paper, some of the arguments of [BG16] are immediately generalized, while others require more work. The new arguments of Section 3 not only allow for weaker hypotheses, they also lead to a better upper bound for $\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*$, already in the special case $\mathcal{S} = \mathcal{X} \setminus \mathcal{R}$.

To connect hypotheses (H) with the pathwise approach framework, we set $T_0 = 0$, $T_1 = T_{\lambda_S}$, we denote by T_2 the absorption time associated with the killing rate $\kappa_{\mathcal{R}}$ after T_1 , that is

$$T_2 = T_1 + T_{\kappa_{\mathcal{R}}} \circ \Theta_{T_1}$$

—where Θ_t is the usual shift operator, such that, for all $s \geq 0$, $(\Theta_t(X))(s) = X(t+s)$ — and we set for all $i \geq 0$,

$$\begin{aligned} T_{2i+1} &= T_{2i} + T_{\lambda_S} \circ \Theta_{T_{2i}}, \\ T_{2i+2} &= T_{2i+1} + T_{\kappa_{\mathcal{R}}} \circ \Theta_{T_{2i+1}}, \end{aligned}$$

using at each step independent copies of the exponential variables σ_{κ} and σ_{λ} . Since we will work with mixing times rather than relaxation times associated with restricted processes, we introduce

$$\chi_{\mathcal{R}} = \max_{x \in \mathcal{R}} \frac{1}{\mu_{\mathcal{R}}(x)}, \quad \chi_{\mathcal{S}} = \max_{x \in \mathcal{S}} \frac{1}{\mu_{\mathcal{S}}(x)}, \quad (6)$$

and a strengthened hypothesis set.

Hypotheses (H'): *Hypotheses (H) hold together with $(\ln \chi_{\mathcal{S}})/\gamma_{\mathcal{S}} \ll 1/\phi_{\mathcal{R} \setminus \mathcal{S}}^*$ and $(\ln \chi_{\mathcal{R}})/\gamma_{\mathcal{R}} \ll 1/\phi_{\mathcal{R} \setminus \mathcal{S}}^*$.*

Recalling Equation (1) it holds:

Theorem 2. *Assuming hypotheses (H') and choosing κ and λ such that*

$$\phi_{\mathcal{R} \setminus \mathcal{S}}^* \ll \kappa \ll \gamma_{\mathcal{R}} / \ln \chi_{\mathcal{R}} \quad \text{and} \quad \phi_{\mathcal{R} \setminus \mathcal{S}}^* \ll \lambda \ll \gamma_{\mathcal{S}} / \ln \chi_{\mathcal{S}},$$

there is a positive $\theta \ll 1/\phi_{\mathcal{R}, \lambda_S}^$ such that, for all $\delta > 0$, for all $n > 0$ and for all observable $f : \mathcal{X} \rightarrow \mathbb{R}$,*

$$\mathbb{P}_{\mu_{\mathcal{R}}} \left(\begin{array}{l} \forall i < n, T_{2i+1} > T_{2i} + \theta, T_{2i+2} > T_{2i+1} + \theta, \\ \forall t \in [T_{2i}, T_{2i+1} - \theta[, \forall t' \in [T_{2i+1}, T_{2i+2} - \theta[, \\ |A_{\theta}(t, f) - \mu_{\mathcal{R}}(f)| \leq \delta, |A_{\theta}(t', f) - \mu_{\mathcal{S}}(f)| \leq \delta \end{array} \right) \sim 1.$$

Comments:

- i. We will give the proof of this theorem in Section 5. We will make explicit a possible choice for θ and we will give explicit estimates of such a probability.
- ii. Let us consider the empirical distribution process defined by

$$A_{\theta}(t) = \frac{1}{\theta} \int_t^{t+\theta} \delta_{X(s)} ds, \quad t \geq 0.$$

If one is interested in convergence of A_{θ} , one can consider (as in [OV05], Chapter 4) a slight modification of the process $A_{\theta}(t)$. With the above notation, let us consider the modification

$$\tilde{A}_{\theta}(t) = \begin{cases} \tilde{A}_{\theta}(T_k - \theta) & \text{if } t \in [T_k - \theta, T_k[\text{ for some } k, \\ A_{\theta}(t) & \text{otherwise.} \end{cases}$$

Then Theorem 2, together with a convergence in law of the rescaled T_k , provides the key tool to prove convergence in Skorokhod topology of \tilde{A}_{θ} .

- iii. In the case $\mu(\mathcal{R}) \ll \mu(\mathcal{S})$, we will have (recall Theorem 1 and its associated third comment)

$$\frac{1}{\gamma} \sim \mathbb{E}_{\mu_{\mathcal{R}}} [T_{\lambda_S}] \sim \frac{1}{\phi_{\mathcal{R}, \lambda_S}^*} \ll \frac{1}{\phi_{\mathcal{S}, \kappa_{\mathcal{R}}}^*} \sim \mathbb{E}_{\mu_{\mathcal{S}}} [T_{\kappa_{\mathcal{R}}}]$$

with $\phi_{\mathcal{S}, \kappa_{\mathcal{R}}}^*$ symmetrically defined with respect to $\phi_{\mathcal{R}, \lambda_S}^*$. The rescaled process on time unit $1/\phi_{\mathcal{R}, \lambda_S}^*$ reaches equilibrium after time $\phi_{\mathcal{R}, \lambda_S}^* T_1$ of order 1, this equilibrium can be described by $\mu_{\mathcal{S}}$, and there is no oscillation on this time scale between the two states $\mu_{\mathcal{R}}$ and $\mu_{\mathcal{S}}$. Our system will go back to $\mu_{\mathcal{R}}$ only after a much longer time of order $1/\phi_{\mathcal{S}, \kappa_{\mathcal{R}}}^*$.

As far as we are concerned with the asymptotic exponential law of T_{λ_S} and $T_{\kappa_{\mathcal{R}}}$, when starting from $\mu_{\mathcal{R}}$ and $\mu_{\mathcal{S}}$, respectively, we can adapt and even simplify the arguments of [BG16]. This is a simple consequence of the following facts, that are the contents of Lemma 2.1, Lemma 2.3 and Proposition 2.6. (See also Proposition 2.7 for a quantitative statement.)

- i. $T_{\lambda_S}^{\mathcal{R}} = \ell_{\mathcal{R}}(T_{\lambda_S})$ is an exact exponential random variable, of rate $\phi_{\mathcal{R},\lambda_S}^*$, when starting from $\mu_{\mathcal{R},\lambda_S}^*$. This is the reason to work with such quasi-stationary distributions.
- ii. Since $\ell_{\mathcal{S}\setminus\mathcal{R}}(T_{\lambda_S})$ is dominated by the exponential random variable σ_{λ} , taking λ in the right window and passing to the expected value, it holds that

$$\frac{1}{\lambda} \ll \frac{1}{\phi_{\mathcal{R}\setminus\mathcal{S}}^*} \leq \frac{1}{\phi_{\mathcal{R},\lambda_S}^*}.$$

Hence, since $T_{\lambda_S} = \ell_{\mathcal{R}}(T_{\lambda_S}) + \ell_{\mathcal{S}\setminus\mathcal{R}}(T_{\lambda_S})$, and together with point i., we deduce the asymptotic law of T_{λ_S} when starting from $\mu_{\mathcal{R},\lambda_S}^*$.

- iii. The restricted ensemble $\mu_{\mathcal{R}}$ and the soft measure $\mu_{\mathcal{R},\lambda_S}^*$ are close in total variation distance.

Our last main result is a “local thermalization property” which implies the existence of an asymptotic exponential law for stopping times T_{λ_S} or $T_{\kappa_{\mathcal{R}}}$, independently of the choice of the initial distribution. As in [BG16] we can use T_{λ_S} and $T_{\kappa_{\mathcal{R}}}$ to build a stopping time T^* of order $1/(\kappa \wedge \lambda)$ at most and such that the law of $X(T^*)$ is close to $\mu_{\mathcal{R},\lambda_S}^*$ or $\mu_{\mathcal{S},\kappa_{\mathcal{R}}}^*$ by conditioning on two complementary events. Since we are here mainly interested in convergence in law of the transitions times, closeness in total variation distance will be enough. We can then build T^* in a simpler way, and with weaker hypotheses, than in [BG16].

We recall that the previously mentioned and symmetric exit rates $\phi_{\mathcal{R},\lambda_S}^*$ and $\phi_{\mathcal{S},\kappa_{\mathcal{R}}}^*$ are precisely defined in Section 2.1. Their asymptotic behaviour is precisely controlled in terms of soft capacities by Theorem 5 in Section 3.2.

Theorem 3. *Assuming hypotheses (H') and choosing κ and λ such that*

$$\phi_{\mathcal{R}\setminus\mathcal{S}}^* \ll \kappa \ll \gamma_{\mathcal{R}}/\ln \chi_{\mathcal{R}} \quad \text{and} \quad \phi_{\mathcal{R}\setminus\mathcal{S}}^* \ll \lambda \ll \gamma_{\mathcal{S}}/\ln \chi_{\mathcal{S}},$$

there are three stopping times $T_{\mathcal{R}}^$, $T_{\mathcal{S}}^*$ and $T^* = T_{\mathcal{R}}^* \wedge T_{\mathcal{S}}^*$ such that, whatever the starting distribution of X :*

- i. *the expected value of T^* is smaller than or equal to $2/(\kappa \wedge \lambda)$;*
- ii. *with $\pi_{\mathcal{R}}^*$ and $\pi_{\mathcal{S}}^*$ the laws of $X(T^*)$ conditioned to $\{T^* = T_{\mathcal{R}}^*\}$ and $\{T^* = T_{\mathcal{S}}^*\}$, respectively, it holds*

$$d_{\text{TV}}(\mu_{\mathcal{R}}, \pi_{\mathcal{R}}^*) \leq \frac{3\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}}\sqrt{\chi_{\mathcal{R}}}}{2\kappa} \right]_+ \right) \ll 1,$$

and

$$d_{\text{TV}}(\mu_{\mathcal{S}}, \pi_{\mathcal{S}}^*) \leq \frac{3\lambda}{\gamma_{\mathcal{S}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{S}}\sqrt{\chi_{\mathcal{S}}}}{2\lambda} \right]_+ \right) \ll 1;$$

- iii. *conditionally to $\{T^* = T_{\mathcal{R}}^*\}$ or $\{T^* = T_{\mathcal{S}}^*\}$, the random variable $\phi_{\mathcal{R},\lambda_S}^* T_{\lambda_S} \circ \Theta_{T^*}$ or $\phi_{\mathcal{S},\kappa_{\mathcal{R}}}^* T_{\kappa_{\mathcal{R}}} \circ \Theta_{T^*}$, respectively, converge in law to an exponential variable or parameter 1;*

- iv. *it holds*

$$\mathbb{P}_{\nu}(T^* = T_{\mathcal{R}}^* = T_{\kappa_{\mathcal{R}}} < T_{\lambda_S}) \geq 1 - 3\mathbb{P}_{\nu}(T_{\lambda_S} < T_{\kappa_{\mathcal{R}}})$$

and

$$\mathbb{P}_{\nu}(T^* = T_{\mathcal{S}}^* = T_{\kappa_{\mathcal{S}}} < T_{\lambda_{\mathcal{R}}}) \geq 1 - 3\mathbb{P}_{\nu}(T_{\lambda_{\mathcal{R}}} < T_{\kappa_{\mathcal{S}}}),$$

so that these probabilities converge to 1 when $\mathbb{P}_{\nu}(T_{\lambda_S} < T_{\kappa_{\mathcal{R}}}) \ll 1$ and $\mathbb{P}_{\nu}(T_{\kappa_{\mathcal{R}}} < T_{\lambda_S}) \ll 1$, respectively.

Comments:

- i. We will give the proof of this theorem in Section 4, together with the explicit construction of T^* and explicit estimates on all error terms.
- ii. This theorem implies an asymptotic exponential law for the rescaled time $\phi_{\mathcal{R},\lambda_S}^* T_{\lambda_S}$ when X starts from any measure ν such that $\mathbb{P}_\nu(T_{\lambda_S} < T_{\kappa\mathcal{R}}) \ll 1$. We will give a direct proof of this fact in Section 2 (see Proposition 2.8). One of the reasons why we needed stronger hypotheses in [BG16] was that we were also interested there in events like $\mathbb{P}_\nu(T_{\lambda_S} > t)$ for times t that were large with respect to $1/\gamma_{\mathcal{R}}$ but of a smaller order than $1/\phi_{\mathcal{R},\lambda_S}^*$.
- iii. Theorem 3 implies an analogue of Theorem 2 where T_0 is replaced by T^* , T_1 by $T^* + T_{\lambda_S} \circ \Theta_{T^*}$ and $\mathbb{P}_{\mu_{\mathcal{R}}}$ by $\mathbb{P}_x(\cdot \mid T^* = T_{\mathcal{R}}^*)$, for any x in \mathcal{X} .

1.2 The entropy issue

The problem of describing metastable states goes back to [Max75]. But we will limit our discussion of the related literature to the last five decades by starting from [LP71]. Since such a review, albeit already biased, was developed and detailed in [BG16], we tried to make it short in the present paper.

After the introduction in [LP71] of restricted ensembles for describing metastable states, the pathwise approach introduced in [CGOV84] was mainly built on [FW84] and its large deviations techniques. It entangled the first two main questions—that of describing the metastable state and that of characterizing the transition time to equilibrium—with a third one, that of giving a detailed picture of the typical paths followed along the transition⁴. One of the main difficulties faced when dealing with each particular model was to give *upper* bounds for the expected value of the transition time to equilibrium. In the case of “very low temperature” dynamics or “rare transition limits”, this question is related with that of giving a detailed description of the “(virtual) energy landscape”. Both reversible and non-reversible dynamics were considered within the pathwise approach, we refer to [OV05] for a full picture and to [CNS15] for what we believe to be the most detailed and general account, along Freidlin and Wentzell’s line, on this main difficulty (disentangled from the third question) in the case of very low temperature systems.

Potential-theoretic tools introduced by [BEGK02], [BEGK04] and [BEGK05] in metastability studies brought in particular two major improvements:

- they related the estimation of mean transition times with capacity estimates for which variational principles were available, without reference to any energy landscape;
- the upper and lower bounds obtained from these variational principles, led in many cases to sharp estimates of mean transition times.

Basic hypotheses expressed in terms of mean hitting times and capacities allowed for asymptotic exponential laws for transition times between local equilibria. We refer to [BdH15] for a full account on these techniques, initially restricted to the reversible case. These potential-theoretic tools were incorporated in [BL10] and [BL11], where the authors proposed to deal with metastable dynamics by proving, using martingale equations, a convergence in law to a Markovian dynamics in Skorohod topology of the label evolution associated with a trace process. From [GL11] and [Lan14] to [LMS19], this was extended to non-reversible dynamics.

When we wrote the first version of [BG16], the vast majority of metastability studies were restricted either to very low temperature dynamics, or to Markovian models that could be naturally mapped to dynamics with very low entropy metastable states, such as distributions that are concentrated around a single point of the configuration space. There were two kinds of difficulties when dealing with metastable states for which entropy really matters: local relaxation was difficult to prove and there was no clear road map to follow. This was the case for the kinetic Ising model at *fixed* subcritical temperature, that is the entropy issue. As far as the general theory was concerned, when looking for example at potential-theoretic tools that linked the estimation of mean transition times to capacities (for which variational formulas were available) one came across two difficulties among others: the formula used harmonic measures that were not naturally linked with metastable states, and it required estimates for a mean electrostatic potential that did not satisfy any variational formula. It seemed also that one

⁴In the case of our kinetic Ising model this amounts to describing the shape of subcritical, critical and subcritical droplets. The techniques presented in [GMV20] allow for proving that the critical droplet is Wulff-shaped and this question is still largely open for subcritical and supercritical droplets.

could get asymptotic exponential laws only if one was able to solve these difficulties. By introducing soft measures and capacities, we reduced in [BG16] these difficulties of the general theory to a single model-dependant issue: getting (rough) upper bounds for the relaxation times of restricted dynamics. Theorem 1 of the present paper indeed provides an asymptotic formula for the mean transition time started in the metastable state, and not from a harmonic measure⁵, that does not require any estimate on the mean potential $\mu(V_\kappa^\lambda)$. We transposed the difficulties of the general theory to that of checking in each particular model the hypotheses (H) or (H') , namely to getting upper bounds for relaxation times of restricted dynamics. Such solutions were immediately adapted to the framework of the “martingale approach to metastability” [BL15]. Armendáriz, Grosskinsky and Loulakis could then study the metastable condensing zero-range process in the thermodynamic limit [AGL17].

We already mentioned that we could not directly apply the results of [BG16] to deal with the kinetic Ising model: [BG16] requires \mathcal{R} and \mathcal{S} to form a partition of the configuration space, but the mesoscopic and macroscopic analysis we make in [GMV20] to control the relaxation times of the restricted dynamics does not allow for such a neat separation at the microscopic level. In the martingale approach framework, local equilibria are however associated with usually far apart subsets \mathcal{R}_i of the configuration space. This space is written as a disjoint union

$$\mathcal{X} = \mathcal{N} \cup \bigcup_{i \in I} \mathcal{R}_i,$$

where \mathcal{N} is a subset of \mathcal{X} in which the system is expected to spend a negligible time on the time scale that allows for the transition to equilibrium. With such a framework, we might still be able to bound the relaxation times on the restricted dynamics in \mathcal{R}_i (say taking $\mathcal{R}_0 \subset \mathcal{R}$ and $\mathcal{R}_1 \subset \mathcal{S}$, with \mathcal{R} and \mathcal{S} defined in [GMV20]) and we could avoid the issue of configurations which in [GMV20] belong to both \mathcal{R} and \mathcal{S} : such critical configurations would have been in \mathcal{N} . But this would have only displaced some difficulties. As previously mentioned, one of the main issues is that of getting upper bounds for the mean transition time. This is turned by Theorem 1, and in the martingale approach, into a problem of lower bounds for a (soft) capacity. Even though associated with a variational principle, this is often one of the most delicate point. As mentioned in the second comment of Theorem 1, such lower bounds are however obtained simply, in [GMV20], from the control of the local relaxation times. In getting such lower bounds for the soft capacities, we do use the fact that \mathcal{R} and \mathcal{S} are “neighbouring” sets (even overlapping in this case). Using overlapping \mathcal{R} and \mathcal{S} , instead of far apart \mathcal{R}_i together with a negligible set \mathcal{N} can be a matter of taste. But there is an issue when estimating the relevant capacities.

It is worth noting at this point that in going from restricted ensemble to soft measures, we start changing our representation for the metastable set, by considering overlapping \mathcal{R} and \mathcal{S} we stop considering only local equilibria with disjoint support, and our soft capacities, that could be considered as our main objects, now only depend on some killing rates like λ_S and not λ and S separately. We could actually consider capacities associated with full support killing rates, and much more general local equilibria. Similar ideas have been considered in [ACGM] with concrete applications [ACGM20] in a quite different field.

As far as the exponential law in itself is concerned [FMNS15], [BG16] and [FMN⁺16] contain similar detailed and quantitative results, obtained with rather similar techniques. The first paper is limited to low entropy metastable states and uses recurrence to an associated reference configuration, the second paper explains how this recurrence property can be replaced by a fast relaxation hypothesis to a quasi-stationary measure and such is the approach of the third paper. By contrast we only focus in the present paper on convergence in law results, which are still quantitative but less precise and associated with simpler constructions. But as in [BG16] we also study, on the one hand, the convergence to a stable equilibrium and not only the exit time of a domain, which is the main object of interest in [FMNS15] and [FMN⁺16], and we are interested, on the other hand, in getting sharp bounds for mean transition times by relying on capacity estimates: this is where our reversibility hypothesis matters, while [FMNS15] and [FMN⁺16] do not assume reversibility.

We close this discussion on the recent literature of abstract metastable dynamics by listing, in order of importance, the improvement of the present paper with respect to [BG16]:

- By relaxing the previous hypothesis $\mathcal{R} \cap \mathcal{S} = \emptyset$, we provided a framework to prove, in the companion paper [GMV20], the asymptotic exponential law of the two-dimensional kinetic Ising model, at any subcritical temperature and in the limit of vanishing magnetic field;

⁵This is however an asymptotic formula and there is no equality.

- we made explicit the connection with the classical pathwise approach by referring to temporal means and identifying explicitly an associated time scale θ ;
- we simplified the construction of the short time scale stopping time T^* at which the system is distributed “as $\mu_{\mathcal{R}}$ or $\mu_{\mathcal{S}}$ ” independently of the starting measure;
- we simplified the proof of the asymptotic exponential law;
- we obtained a sharper quantitative control of the mean transition time to equilibrium.

2 Basic definitions and properties

2.1 Soft measures and killed process

Let X be an irreducible Markov process, with generator \mathcal{L} as in Equation (2). We assume X to be reversible with respect to a probability measure μ . For \mathcal{R} and \mathcal{S} such that $\mathcal{X} = \mathcal{R} \cup \mathcal{S}$, and for any given finite $\lambda \geq 0$ we build in this section the soft measure $\mu_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ from the trace $X_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ on \mathcal{R} of the process X killed at rate $\lambda_{\mathcal{S}}(x)$ in each x in \mathcal{X} . To facilitate the reading, we use the upper index $*$ to denote objects associated with killed processes.

Formally, let $T_{\mathcal{R}}$ be the hitting time of \mathcal{R} , and $T_{\lambda_{\mathcal{S}}}$ be the absorption time of X killed at rate $\lambda_{\mathcal{S}}$ defined in Equation (5). Equivalently⁶, we can associate any $x \in \mathcal{X}$ with an independent Poisson process of rate $\lambda_{\mathcal{S}}(x)$, and define $T_{\lambda_{\mathcal{S}}}$ as the first arrival time t of the Poisson process associated with $x = X(t)$. The trace process $X_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ is then the Markov process on \mathcal{R} that, on one hand, is killed in x in \mathcal{R} at rate

$$e_{\mathcal{R},\lambda_{\mathcal{S}}}^*(x) = \lambda_{\mathcal{S}}(x) + \sum_{z \neq x} w(x,z) \mathbb{P}_z(T_{\mathcal{R}} > T_{\lambda_{\mathcal{S}}}) = \lambda_{\mathcal{S}}(x) + \sum_{z \notin \mathcal{R}} w(x,z) \mathbb{P}_z(T_{\mathcal{R}} > T_{\lambda_{\mathcal{S}}}), \quad (7)$$

where the second term accounts for the rate with which X escapes from \mathcal{R} and is killed before returning in it, and that, on the other hand, jumps from x to a distinct $y \in \mathcal{R}$ at rate

$$w_{\mathcal{R},\lambda_{\mathcal{S}}}(x,y) = \sum_{z \neq x} w(x,z) \mathbb{P}_z(X(T_{\mathcal{R}}) = y, T_{\mathcal{R}} < T_{\lambda_{\mathcal{S}}}) = w(x,y) + \sum_{z \notin \mathcal{R}} w(x,z) \mathbb{P}_z(X(T_{\mathcal{R}}) = y, T_{\mathcal{R}} < T_{\lambda_{\mathcal{S}}}). \quad (8)$$

It is associated with the sub-Markovian generator $\mathcal{L}_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ that acts on functions $f : \mathcal{R} \rightarrow \mathbb{R}$ according to

$$(\mathcal{L}_{\mathcal{R},\lambda_{\mathcal{S}}}^* f)(x) = -e_{\mathcal{R},\lambda_{\mathcal{S}}}^*(x) f(x) + \sum_{y \in \mathcal{R} \setminus \{x\}} w_{\mathcal{R},\lambda_{\mathcal{S}}}(x,y) (f(y) - f(x)), \quad x \in \mathcal{R}. \quad (9)$$

Notice that the irreducibility hypothesis (cf. (H)) on X_R , having generator given by Equation (4), implies the irreducibility of $X_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ associated with Equation (9). Then, by Perron-Frobenius Theorem, $\mathcal{L}_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ has only negative eigenvalues and the smallest eigenvalue $\phi_{\mathcal{R},\lambda_{\mathcal{S}}}^* > 0$ of $-\mathcal{L}_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ is non-degenerate and associated with a left eigenvector $\mu_{\mathcal{R},\lambda_{\mathcal{S}}}^*$ which is a probability on \mathcal{R} . We call it the *soft measure* associated with the killing rate $\lambda_{\mathcal{S}}$ and it is a special kind of quasi-stationary measure. We denote by $\ell_{\mathcal{R}}$ the local time in \mathcal{R} :

$$\ell_{\mathcal{R}}(t) = \int_0^t \mathbf{1}_{\{X(u) \in \mathcal{R}\}} du, \quad t \geq 0.$$

It is a non-decreasing continuous function of the time t and we call $\ell_{\mathcal{R}}^{-1}$ its right-continuous inverse:

$$\ell_{\mathcal{R}}^{-1}(s) = \inf \{t \geq 0 : \ell_{\mathcal{R}}(t) > s\}, \quad s \geq 0.$$

We then have $X_{\mathcal{R},\lambda_{\mathcal{S}}}^*(s) = X \circ \ell_{\mathcal{R}}^{-1}(s)$ for all $s < T_{\lambda_{\mathcal{S}}}^{\mathcal{R}} = \ell_{\mathcal{R}}(T_{\lambda_{\mathcal{S}}})$ and the following lemma.

Lemma 2.1. *For all $s \geq 0$ and all x and y in \mathcal{R} it holds*

$$\begin{aligned} \mathbb{P}_{\mu_{\mathcal{R},\lambda_{\mathcal{S}}}^*} (X \circ \ell_{\mathcal{R}}^{-1}(s) = x \mid T_{\lambda_{\mathcal{S}}}^{\mathcal{R}} > s) &= \mu_{\mathcal{R},\lambda_{\mathcal{S}}}^*(x), \\ \lim_{s \rightarrow \infty} \mathbb{P}_x (X \circ \ell_{\mathcal{R}}^{-1}(s) = y \mid T_{\lambda_{\mathcal{S}}}^{\mathcal{R}} > s) &= \mu_{\mathcal{R},\lambda_{\mathcal{S}}}^*(y), \end{aligned}$$

⁶This is an alternative construction that does not use the exponential variable σ_{λ} of Section 1.1 and allows for less homogeneous killing rates.

and

$$\mathbb{P}_{\mu_{\mathcal{R},\lambda\mathcal{S}}^*} (T_{\lambda\mathcal{S}}^{\mathcal{R}} > s) = e^{-s\phi_{\mathcal{R},\lambda\mathcal{S}}^*}.$$

The exit rate is also given by

$$\phi_{\mathcal{R},\lambda\mathcal{S}}^* = \mu_{\mathcal{R},\lambda\mathcal{S}}^* (e_{\mathcal{R},\lambda\mathcal{S}}^*).$$

These formulas are consequences of the fact that, for all $s \geq 0$, every function f and every probability ν on \mathcal{R} ,

$$\mathbb{E}_{\nu} \left[f(X_{\mathcal{R},\lambda\mathcal{S}}^*(s)) \mathbf{1}_{\{T_{\lambda\mathcal{S}}^{\mathcal{R}} > s\}} \right] = \nu \left(e^{s\mathcal{L}_{\mathcal{R},\lambda\mathcal{S}}^*} f \right).$$

We refer to Lemma 2.12 in [BG16] for more details.

As in [BG16] (Lemma 2.13) and assuming hypotheses (H) , when λ goes from 0 to $+\infty$ we get a continuous interpolation between the restricted ensemble $\mu_{\mathcal{R}}$ and the quasi-stationary measure $\mu_{\mathcal{R}\setminus\mathcal{S}}^*$. Indeed, when $\lambda = 0$ our soft measure $\mu_{\mathcal{R},\lambda\mathcal{S}}^*$ is nothing but the invariant measure of the trace process on \mathcal{R} , which coincides with $\mu(\cdot|\mathcal{R})$ as a consequence of our reversibility hypothesis. And, if $\lambda = +\infty$ we are led to define $X_{\mathcal{R},\lambda\mathcal{S}}^*$ as $X_{\mathcal{R}\setminus\mathcal{S}}^*$, obtained from X by killing it instantaneously in \mathcal{S} . This process is associated with the sub-Markovian generator $\mathcal{L}_{\mathcal{R}\setminus\mathcal{S}}^*$ defined by

$$\left(\mathcal{L}_{\mathcal{R}\setminus\mathcal{S}}^* f \right) (x) = -e_{\mathcal{R}\setminus\mathcal{S}}^*(x) f(x) + \sum_{\substack{y \in \mathcal{R}\setminus\mathcal{S}, \\ y \neq x}} w(x,y) (f(y) - f(x)), \quad f : \mathcal{R}\setminus\mathcal{S} \rightarrow \mathbb{R}, \quad x \in \mathcal{R},$$

with, for all x in $\mathcal{R}\setminus\mathcal{S}$,

$$e_{\mathcal{R}\setminus\mathcal{S}}^*(x) = \sum_{z \in \mathcal{S}} w(x,z). \quad (10)$$

Since $X_{\mathcal{R}\setminus\mathcal{S}}^*$ is assumed to be irreducible, the quasi-stationary measure $\mu_{\mathcal{R}\setminus\mathcal{S}}^*$ is unambiguously defined. Contrarily to $X_{\mathcal{R}\setminus\mathcal{S}}^*$, which cannot penetrate \mathcal{S} , $\mu_{\mathcal{R},\lambda\mathcal{S}}^*$ and $X_{\mathcal{R},\lambda\mathcal{S}}^*$ are associated with a process which can visit the killing zone \mathcal{S} . As mentioned earlier, this is the explanation for the terminology *soft* measure.

We denote by $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ the standard norm and scalar product in $\ell^2(\mu)$. For any $\mathcal{A} \subset \mathcal{X}$ we denote by $\mu_{\mathcal{A}}$ the conditional probability measure $\mu_{\mathcal{A}} = \mu(\cdot|\mathcal{A})$, and by $\|\cdot\|_{\mathcal{A}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{A}}$ the standard norm and scalar product in $\ell^2(\mu_{\mathcal{A}})$. We also denote by $e_{\mathcal{A}}^*$, as in Equation (10), the escape rate

$$e_{\mathcal{A}}^*(x) = \sum_{z \notin \mathcal{A}} w(x,z)$$

for x in \mathcal{A} .

Proposition 2.2. *When λ grows to $+\infty$,*

$$\phi_{\mathcal{R},\lambda\mathcal{S}}^* = \min_{\substack{f: \mathcal{R} \rightarrow \mathbb{R} \\ f \neq 0}} \frac{\langle f, -\mathcal{L}_{\mathcal{R},\lambda\mathcal{S}}^* f \rangle_{\mathcal{R}}}{\|f\|_{\mathcal{R}}^2} = \min_{f > 0} \frac{\langle f, -\mathcal{L}_{\mathcal{R},\lambda\mathcal{S}}^* f \rangle_{\mathcal{R}}}{\|f\|_{\mathcal{R}}^2} \quad (11)$$

continuously grows to $\phi_{\mathcal{R}\setminus\mathcal{S}}^$. In particular, for all $\lambda \geq 0$, it holds*

$$\phi_{\mathcal{R},\lambda\mathcal{S}}^* \leq \phi_{\mathcal{R}\setminus\mathcal{S}}^* = \min_{\substack{f: \mathcal{R}\setminus\mathcal{S} \rightarrow \mathbb{R} \\ f \neq 0}} \frac{\langle f, -\mathcal{L}_{\mathcal{R}\setminus\mathcal{S}}^* f \rangle_{\mathcal{R}\setminus\mathcal{S}}}{\|f\|_{\mathcal{R}\setminus\mathcal{S}}^2}. \quad (12)$$

Proof: The first equality in Equation (11) comes from the fact that $\mathcal{L}_{\mathcal{R},\lambda\mathcal{S}}^*$ is a self-adjoint operator. The last equality in Equation (11) is a consequence of Perron-Frobenius Theorem. Continuity follows from Lemma 2.1 and continuity properties of $\mu_{\mathcal{R},\lambda\mathcal{S}}^*$ and $e_{\mathcal{R},\lambda\mathcal{S}}^*$. Finally, for any positive function f on \mathcal{R} it holds

$$\begin{aligned} \langle f, -\mathcal{L}_{\mathcal{R},\lambda\mathcal{S}}^* f \rangle_{\mathcal{R}} &= - \sum_{x \in \mathcal{R}} \sum_{y \in \mathcal{R} \setminus \{x\}} \mu_{\mathcal{R}}(x) f(x) w_{\mathcal{R},\lambda\mathcal{S}}(x,y) f(y) \\ &\quad + \sum_{x \in \mathcal{R}} \mu_{\mathcal{R}}(x) \left(e_{\mathcal{R},\lambda\mathcal{S}}^*(x) + \sum_{y \in \mathcal{R} \setminus \{x\}} w_{\mathcal{R},\lambda\mathcal{S}}(x,y) \right) f^2(x). \end{aligned}$$

The first term, minus sign included, which accounts for the cross product, is clearly increasing in λ , since $w_{\mathcal{R},\lambda_S}(x,y)$ decreases with λ and f is positive. As far as the second term is concerned, using the definitions (7) and 8, we compute

$$\begin{aligned}
e_{\mathcal{R},\lambda_S}^*(x) + \sum_{y \in \mathcal{R} \setminus \{x\}} w_{\mathcal{R},\lambda_S}(x,y) &= \lambda_S(x) + \sum_{z \notin \mathcal{R}} w(x,z) \mathbb{P}_z(T_{\mathcal{R}} > T_{\lambda_S}) \\
&+ \sum_{y \in \mathcal{R} \setminus \{x\}} w(x,y) + \sum_{y \in \mathcal{R} \setminus \{x\}} \sum_{z \notin \mathcal{R}} w(x,z) \mathbb{P}_z(X(T_{\mathcal{R}}) = y, T_{\mathcal{R}} < T_{\lambda_S}) \\
&= \lambda_S(x) + \sum_{z \notin \mathcal{R}} w(x,z) \mathbb{P}_z(T_{\mathcal{R}} > T_{\lambda_S}) + \sum_{y \in \mathcal{R} \setminus \{x\}} w(x,y) \\
&+ \sum_{z \notin \mathcal{R}} w(x,z) \left(1 - \mathbb{P}_z(T_{\mathcal{R}} > T_{\lambda_S}) - \mathbb{P}_z(X(T_{\mathcal{R}}) = x, T_{\mathcal{R}} < T_{\lambda_S})\right) \\
&= \lambda_S(x) + w(x) - \sum_{z \notin \mathcal{R}} w(x,z) \mathbb{P}_z(X(T_{\mathcal{R}}) = x, T_{\mathcal{R}} < T_{\lambda_S}),
\end{aligned}$$

where in the last line we set $w(x) = \sum_{z \neq x} w(x,z)$. The whole coefficient is then increasing in λ and the monotonicity of $\phi_{\mathcal{R},\lambda_S}^*$ follows. \square

Using simple test functions in such variational principles we get the following upper bounds on $\phi_{\mathcal{R},\lambda_S}^*$.

Lemma 2.3. *It holds, for all $\lambda \geq 0$,*

$$\phi_{\mathcal{R},\lambda_S}^* \leq \phi_{\mathcal{R} \setminus \mathcal{S}}^* \leq \mu_{\mathcal{R} \setminus \mathcal{S}}(e_{\mathcal{R} \setminus \mathcal{S}}^*), \quad (13)$$

$$\phi_{\mathcal{R},\lambda_S}^* \leq \lambda \mu_{\mathcal{R}}(\mathcal{R} \cap \mathcal{S}) + \mu_{\mathcal{R}}(e_{\mathcal{R}}^*), \quad (14)$$

$$\mathbb{E}_{\mu_{\mathcal{R}}}[T_{\lambda_S}^{\mathcal{R}}] \leq \frac{1}{\phi_{\mathcal{R},\lambda_S}^*}. \quad (15)$$

Proof: Formula (13) is obtained from Formula (12) with the constant function equal to 1 on $\mathcal{R} \setminus \mathcal{S}$ used as a test function. We use the constant function $\mathbf{1}_{\mathcal{R}}$ equal to 1 on \mathcal{R} in Equation (11) and we bound by 1 the last probability appearing in Equation (7) to get Inequality (14). Finally, to get Inequality (15), we observe that $1/\phi_{\mathcal{R},\lambda_S}^*$ is the largest eigenvalue of the Green kernel $\mathcal{G}_{\lambda_S}^{\mathcal{R}} = (-\mathcal{L}_{\mathcal{R},\lambda_S})^{-1}$, so that, writing ℓ_y for the local time in y ,

$$\frac{1}{\phi_{\mathcal{R},\lambda_S}^*} = \max_{\|f\|_{\mathcal{R}}=1} \langle f, \mathcal{G}_{\lambda_S}^{\mathcal{R}} f \rangle_{\mathcal{R}} \geq \langle \mathbf{1}_{\mathcal{R}}, \mathcal{G}_{\lambda_S}^{\mathcal{R}} \mathbf{1}_{\mathcal{R}} \rangle_{\mathcal{R}} = \sum_{x,y \in \mathcal{R}} \mu_{\mathcal{R}}(x) \mathbb{E}_x[\ell_y(T_{\lambda_S})] = \mathbb{E}_{\mu_{\mathcal{R}}}[\ell_{\mathcal{R}}(T_{\lambda_S})].$$

\square

Remark: While $\phi_{\mathcal{R},\lambda_S}^*$ and $\phi_{\mathcal{R} \setminus \mathcal{S}}^*$ are the means of escape rates according to quasi-stationary measures that can be difficult to compute, the upper bounds given in (13) and (14) refer only to the reversible measure μ .

2.2 Restricted dynamics

For any finite $\lambda \geq 0$, the *restricted dynamics* $X_{\mathcal{R},\lambda_S}$ is obtained from $X_{\mathcal{R},\lambda_S}^*$ or X by removing all killing moves or killing excursions outside \mathcal{R} . It is associated with the generator $\mathcal{L}_{\mathcal{R},\lambda_S}$ that acts on functions $f : \mathcal{R} \rightarrow \mathbb{R}$ according to

$$(\mathcal{L}_{\mathcal{R},\lambda_S} f)(x) = \sum_{y \in \mathcal{R} \setminus \{x\}} w_{\mathcal{R},\lambda_S}(x,y)(f(y) - f(x)), \quad x \in \mathcal{R},$$

with $w_{\mathcal{R},\lambda_S}(x,y)$ defined by Equation (8) for distinct x and y in \mathcal{R} . In particular this process is reversible with respect to the restricted measure $\mu_{\mathcal{R}}$ and it inherits irreducibility from the restricted process $X_{\mathcal{R}}$. We call $\gamma_{\mathcal{R},\lambda_S}$ its spectral gap

$$\gamma_{\mathcal{R},\lambda_S} = \min_{\substack{f: \mathcal{R} \rightarrow \mathbb{R}, \\ \text{Var}_{\mu_{\mathcal{R}}}(f) \neq 0}} \frac{\mathcal{D}_{\mathcal{R},\lambda_S}(f)}{\text{Var}_{\mu_{\mathcal{R}}}(f)}, \quad (16)$$

with $\mathcal{D}_{\mathcal{R},\lambda_S}$ the *Dirichlet form* defined by

$$\mathcal{D}_{\mathcal{R},\lambda_S}(f) = \langle f, -\mathcal{L}_{\mathcal{R},\lambda_S} f \rangle_{\mathcal{R}} = \frac{1}{2} \sum_{\substack{x,y \in \mathcal{R} \\ x \neq y}} c_{\mathcal{R},\lambda_S}(x,y) [f(y) - f(x)]^2, \quad f : \mathcal{R} \rightarrow \mathbb{R},$$

where the *conductances* $c_{\mathcal{R},\lambda_S}(x,y)$ are given by

$$c_{\mathcal{R},\lambda_S}(x,y) = \mu_{\mathcal{R}}(x)w_{\mathcal{R},\lambda_S}(x,y) = \mu_{\mathcal{R}}(y)w_{\mathcal{R},\lambda_S}(y,x).$$

For $\lambda = +\infty$ we define $X_{\mathcal{R},\lambda_S}$ as the limiting process $X_{\mathcal{R}}$, rather than by removing all killing moves from the limiting killed process $X_{\mathcal{R} \setminus S}^*$. In this case we will also write $\mathcal{D}_{\mathcal{R}}$ and $c_{\mathcal{R}}$ in place of $\mathcal{D}_{\mathcal{R},\lambda_S}$ and $c_{\mathcal{R},\lambda_S}$. From Equation (16) as well as the continuity and monotonicity in λ of the conductances $c_{\mathcal{R},\lambda_S}$, which is inherited from that of the rates $w_{\mathcal{R},\lambda_S}(x,y)$, we get

Proposition 2.4. *When λ goes to $+\infty$, $\gamma_{\mathcal{R},\lambda_S}$ continuously decreases to $\gamma_{\mathcal{R}}$. In particular, it holds*

$$\gamma_{\mathcal{R},\lambda_S} \geq \gamma_{\mathcal{R}}$$

for all $\lambda \geq 0$.

We have from propositions 2.2 and 2.4 that the ratio

$$\epsilon_{\mathcal{R},\lambda_S}^* = \frac{1/\gamma_{\mathcal{R},\lambda_S}}{1/\phi_{\mathcal{R},\lambda_S}^*} = \frac{\phi_{\mathcal{R},\lambda_S}^*}{\gamma_{\mathcal{R},\lambda_S}}$$

is decreasing in λ . When $\lambda = +\infty$ we will also write

$$\epsilon_{\mathcal{R},S}^* = \frac{\phi_{\mathcal{R} \setminus S}^*}{\gamma_{\mathcal{R}}} \tag{17}$$

in place of $\epsilon_{\mathcal{R},\lambda_S}^*$ and we get, for all $\lambda \geq 0$,

$$\epsilon_{\mathcal{R},\lambda_S}^* \leq \epsilon_{\mathcal{R},S}^*. \tag{18}$$

Hence $\epsilon_{\mathcal{R},\lambda_S}^*$ tends to be negligible under hypotheses (H).

In this case all our soft measures are close together. Let us introduce some more notation to make this precise. We denote by $h_{\mathcal{R},\lambda_S}^*$ the density of $\mu_{\mathcal{R},\lambda_S}^*$ with respect to $\mu_{\mathcal{R}}$, i.e.,

$$h_{\mathcal{R},\lambda_S}^*(x) = \frac{\mu_{\mathcal{R},\lambda_S}^*(x)}{\mu_{\mathcal{R}}(x)}, \quad x \in \mathcal{R}.$$

In the special case $\lambda = +\infty$ we write $h_{\mathcal{R} \setminus S}^*$ for $h_{\mathcal{R},\lambda_S}^*$:

$$h_{\mathcal{R} \setminus S}^* : x \in \mathcal{R} \mapsto \frac{\mu_{\mathcal{R} \setminus S}^*(x)}{\mu_{\mathcal{R}}(x)},$$

which is supported on $\mathcal{R} \setminus S$. In particular, for any finite λ , we get

$$\phi_{\mathcal{R},\lambda_S}^* = \frac{\langle h_{\mathcal{R},\lambda_S}^*, -\mathcal{L}_{\mathcal{R},\lambda_S}^* h_{\mathcal{R},\lambda_S}^* \rangle_{\mathcal{R}}}{\|h_{\mathcal{R},\lambda_S}^*\|_{\mathcal{R}}^2},$$

and

$$\phi_{\mathcal{R} \setminus S}^* = \frac{\langle h_{\mathcal{R} \setminus S}^*, -\mathcal{L}_{\mathcal{R} \setminus S}^* h_{\mathcal{R} \setminus S}^* \rangle_{\mathcal{R}}}{\|h_{\mathcal{R} \setminus S}^*\|_{\mathcal{R}}^2}.$$

Since $\mathcal{L}_{\mathcal{R} \setminus S}^*$ acts on functions of $\mathcal{R} \setminus S$, to make sense of $\mathcal{L}_{\mathcal{R} \setminus S}^* h_{\mathcal{R} \setminus S}^*$ we used in the previous equation

Convention 2.5. *Any function f on $\mathcal{A} \subset \mathcal{X}$ is identified with the function on its support $\mathcal{B} \subset \mathcal{A}$. In particular it also identified, for any $\mathcal{C} \supset \mathcal{B}$, with the function g on \mathcal{C} that coincides with f on \mathcal{B} and has zero value in $\mathcal{C} \setminus \mathcal{B}$.*

We can now compare our soft measures in $\ell^2(\mu_{\mathcal{R}})$ and $\ell^1(\mu_{\mathcal{R}})$ norms.

Proposition 2.6. *For all $\lambda \in [0, +\infty]$, if $\epsilon_{\mathcal{R}, \lambda_S}^* < 1$, then*

$$\text{Var}_{\mu_{\mathcal{R}}} (h_{\mathcal{R}, \lambda_S}^*) = \|h_{\mathcal{R}, \lambda_S}^* - \mathbb{1}_{\mathcal{R}}\|_{\mathcal{R}}^2 \leq \frac{\epsilon_{\mathcal{R}, \lambda_S}^*}{1 - \epsilon_{\mathcal{R}, \lambda_S}^*}.$$

In particular, with d_{TV} the total variation distance,

$$d_{\text{TV}} (\mu_{\mathcal{R}, \lambda_S}^*, \mu_{\mathcal{R}}) \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda_S}^*}{1 - \epsilon_{\mathcal{R}, \lambda_S}^*}}. \quad (19)$$

Proof: Using $h_{\mathcal{R}, \lambda_S}^*$ as a test function in the variational principle given by Equation (16), we get, for any finite λ ,

$$\begin{aligned} \text{Var}_{\mu_{\mathcal{R}}} (h_{\mathcal{R}, \lambda_S}^*) &\leq \frac{1}{\gamma_{\mathcal{R}, \lambda_S}} \mathcal{D}_{\mathcal{R}, \lambda_S} (h_{\mathcal{R}, \lambda_S}^*) = \frac{1}{\gamma_{\mathcal{R}, \lambda_S}} \langle h_{\mathcal{R}, \lambda_S}^*, -\mathcal{L}_{\mathcal{R}, \lambda_S} h_{\mathcal{R}, \lambda_S}^* \rangle_{\mathcal{R}} \\ &= \frac{1}{\gamma_{\mathcal{R}, \lambda_S}} \left(\langle h_{\mathcal{R}, \lambda_S}^*, -\mathcal{L}_{\mathcal{R}, \lambda_S}^* h_{\mathcal{R}, \lambda_S}^* \rangle_{\mathcal{R}} - \sum_{x \in \mathcal{R}} \mu_{\mathcal{R}}(x) e_{\mathcal{R}}^*(x) h_{\mathcal{R}, \lambda_S}^{*2}(x) \right) \\ &\leq \frac{1}{\gamma_{\mathcal{R}, \lambda_S}} \phi_{\mathcal{R}, \lambda_S}^* \mu_{\mathcal{R}} (h_{\mathcal{R}, \lambda_S}^{*2}) = \frac{\phi_{\mathcal{R}, \lambda_S}^*}{\gamma_{\mathcal{R}, \lambda_S}} (1 + \text{Var}_{\mu_{\mathcal{R}}} (h_{\mathcal{R}, \lambda_S}^*)). \end{aligned}$$

In the limiting case $\lambda = +\infty$ we obtain a similar estimate:

$$\begin{aligned} \text{Var}_{\mu_{\mathcal{R}}} (h_{\mathcal{R} \setminus S}^*) &\leq \frac{1}{\gamma_{\mathcal{R}}} \mathcal{D}_{\mathcal{R}} (h_{\mathcal{R} \setminus S}^*) \\ &= \frac{1}{\gamma_{\mathcal{R}}} \left(\frac{1}{2} \sum_{x, y \in \mathcal{R} \setminus S} c_{\mathcal{R}}(x, y) [h_{\mathcal{R} \setminus S}^*(x) - h_{\mathcal{R} \setminus S}^*(y)]^2 + \sum_{\substack{x \in \mathcal{R} \setminus S \\ y \in S \cap \mathcal{R}}} c_{\mathcal{R}}(x, y) [h_{\mathcal{R} \setminus S}^*(x)]^2 \right) \\ &\leq \frac{1}{\gamma_{\mathcal{R}}} \left(\sum_{x \in \mathcal{R} \setminus S} \mu_{\mathcal{R}}(x) h_{\mathcal{R} \setminus S}^*(x) \sum_{y \in \mathcal{R} \setminus S} w(x, y) [h_{\mathcal{R} \setminus S}^*(x) - h_{\mathcal{R} \setminus S}^*(y)] \right. \\ &\quad \left. + \sum_{x \in \mathcal{R} \setminus S} \mu_{\mathcal{R}}(x) \sum_{y \in S} w(x, y) [h_{\mathcal{R} \setminus S}^*(x)]^2 \right) \\ &= \frac{1}{\gamma_{\mathcal{R}}} \langle h_{\mathcal{R} \setminus S}^*, -\mathcal{L}_{\mathcal{R} \setminus S}^* h_{\mathcal{R} \setminus S}^* \rangle_{\mathcal{R}} \\ &= \frac{1}{\gamma_{\mathcal{R}}} \phi_{\mathcal{R} \setminus S}^* \mu_{\mathcal{R}} (h_{\mathcal{R} \setminus S}^{*2}) = \frac{\phi_{\mathcal{R} \setminus S}^*}{\gamma_{\mathcal{R}}} (1 + \text{Var}_{\mu_{\mathcal{R}}} (h_{\mathcal{R} \setminus S}^*)). \end{aligned}$$

Rearranging the terms, we get the desired inequality. Equation (19) follows then by Jensen inequality. \square

As a consequence, as soon as $\lambda \gg \phi_{\mathcal{R}, \lambda_S}^*$, for example by setting $\lambda \gg \phi_{\mathcal{R} \setminus S}^*$ (in virtue of (13)), we obtain an asymptotic exponential law for T_{λ_S} when starting from $\mu_{\mathcal{R}, \lambda_S}^*$ or $\mu_{\mathcal{R}}$. If $\epsilon_{\mathcal{R}, \lambda_S}^* \ll 1$, then the random time $\phi_{\mathcal{R}, \lambda_S}^* T_{\lambda_S}$ converges in law to an exponential variable of parameter 1.

Proposition 2.7. *For all $\lambda \geq \phi_{\mathcal{R} \setminus S}^*$, or any λ such that $\lambda \geq \phi_{\mathcal{R}, \lambda_S}^*$, it holds*

$$e^{-t} \leq \mathbb{P}_{\mu_{\mathcal{R}, \lambda_S}^*} \left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R}, \lambda_S}^*} \right) \leq e^{-t} \left\{ \exp \left(\sqrt{\phi_{\mathcal{R}, \lambda_S}^* / \lambda} \right) + \exp \left(t - \sqrt{\lambda / \phi_{\mathcal{R}, \lambda_S}^*} \right) \right\}$$

as soon as $\sqrt{\phi_{\mathcal{R}, \lambda_S}^* / \lambda} \leq t$, and, if $\epsilon_{\mathcal{R}, \lambda_S}^* < 1$,

$$\left| \mathbb{P}_{\mu_{\mathcal{R}}} \left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R}, \lambda_S}^*} \right) - \mathbb{P}_{\mu_{\mathcal{R}, \lambda_S}^*} \left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R}, \lambda_S}^*} \right) \right| \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda_S}^*}{1 - \epsilon_{\mathcal{R}, \lambda_S}^*}}.$$

Remark: As in [BG16] we write into brace parenthesis quantities that go to one under suitable hypotheses.

Proof of the proposition: The first inequality is a consequence of the fact that $T_{\lambda_S} \geq T_{\lambda_S}^{\mathcal{R}}$ and that, starting from $\mu_{\mathcal{R},\lambda_S}^*$, the latter is an exponential variable of parameter $\phi_{\mathcal{R},\lambda_S}^*$. As far as the second one is concerned, it holds, with ℓ_S the local time in \mathcal{S} ,

$$T_{\lambda_S} \leq T_{\lambda_S}^{\mathcal{R}} + \ell_S(T_{\lambda_S})$$

so that, for any $0 \leq \beta \leq 1$,

$$\begin{aligned} \mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*} \left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*} \right) &\leq \mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*} \left(T_{\lambda_S}^{\mathcal{R}} > \frac{(1-\beta)t}{\phi_{\mathcal{R},\lambda_S}^*} \right) + \mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*} \left(\ell_S(T_{\lambda_S}) > \frac{\beta t}{\phi_{\mathcal{R},\lambda_S}^*} \right) \\ &= e^{-(1-\beta)t} + e^{-\lambda \beta t / \phi_{\mathcal{R},\lambda_S}^*}. \end{aligned}$$

Choosing

$$\beta = \frac{1}{t} \sqrt{\frac{\phi_{\mathcal{R},\lambda_S}^*}{\lambda}} \leq 1,$$

this gives the desired result. Finally, by considering an optimal coupling between two random variables $X(0)$ and $X'(0)$ of law $\mu_{\mathcal{R}}$ and $\mu_{\mathcal{R},\lambda_S}^*$, such that $X(0) \neq X'(0)$ with probability (recall Proposition 2.6)

$$d_{\text{TV}}(\mu_{\mathcal{R}}, \mu_{\mathcal{R},\lambda_S}^*) \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}}$$

it holds

$$\mathbb{P}_{\mu_{\mathcal{R}}}(\phi_{\mathcal{R},\lambda_S}^* T_{\lambda_S} > t) \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + \mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*}(\phi_{\mathcal{R},\lambda_S}^* T_{\lambda_S} > t)$$

and

$$\mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*}(\phi_{\mathcal{R},\lambda_S}^* T_{\lambda_S} > t) \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + \mathbb{P}_{\mu_{\mathcal{R}}}(\phi_{\mathcal{R},\lambda_S}^* T_{\lambda_S} > t).$$

□

It follows that we also have an asymptotic exponential law for T_{λ_S} when starting from any probability measure ν such that the total variation distance between $\mu_{\mathcal{R}}$ and ν goes to zero. An example of such a distribution ν is given by the law of $X(T_{\kappa_{\mathcal{R}}})$ if $\kappa \ll \gamma_{\mathcal{R}} / \ln \chi_{\mathcal{R}}$ (recall Equation (6)). Hence, if a probability measure π is such that $\mathbb{P}_{\pi}(T_{\kappa_{\mathcal{R}}} > T_{\lambda_S}) \ll 1$, then T_{λ_S} has an asymptotic exponential law of parameter $\phi_{\mathcal{R},\lambda_S}^*$.

Proposition 2.8. *For all $x \in \mathcal{X}$ and all $\kappa > 0$ it holds*

$$d_{\text{TV}}(\mu_{\mathcal{R}}, \mathbb{P}_x(X(T_{\kappa_{\mathcal{R}}}) = \cdot)) \leq \frac{\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\chi_{\mathcal{R}}}}{2\kappa} \right]_+ \right). \quad (20)$$

Also, for any probability π on \mathcal{X} , by setting

$$\delta = \mathbb{P}_{\pi}(T_{\lambda_S} < T_{\kappa_{\mathcal{R}}})$$

and

$$\eta = \delta + \frac{\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\chi_{\mathcal{R}}}}{2\kappa} \right]_+ \right) + \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}},$$

it holds, for all $\lambda \geq \phi_{\mathcal{R} \setminus \mathcal{S}}^*$, or any $\lambda \geq \phi_{\mathcal{R},\lambda_S}^*$,

$$e^{-t} \{1 - \eta e^t\} \leq \mathbb{P}_{\pi} \left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*} \right) \leq e^{-t} \left\{ \exp \left(2 \sqrt{\frac{\phi_{\mathcal{R},\lambda_S}^*}{\kappa \wedge \lambda}} \right) + \left[\eta + 2 \exp \left(- \sqrt{\frac{\kappa \wedge \lambda}{\phi_{\mathcal{R},\lambda_S}^*}} \right) \right] e^t \right\}.$$

Proof: Let us denote by $X^{\mathcal{R}}$ the trace of X on \mathcal{R} . Note that we already gave to this process an heavier notation:

$$X^{\mathcal{R}} = X_{\mathcal{R},0_S}.$$

Since $X(T_{\kappa_{\mathcal{R}}}) \in \mathcal{R}$, the law of $X(T_{\kappa_{\mathcal{R}}})$ is that of $X^{\mathcal{R}}(T_{\kappa_{\mathcal{R}}}^{\mathcal{R}})$ with

$$T_{\kappa_{\mathcal{R}}}^{\mathcal{R}} = \ell_{\mathcal{R}}(T_{\kappa_{\mathcal{R}}}).$$

Since the relaxation time of $X^{\mathcal{R}}$ is smaller than or equal to that of $X_{\mathcal{R}}$ (recall Proposition 2.4), by using reversibility to estimate the $\ell^2(\mu_{\mathcal{R}})$ distance between (the densities with respect to $\mu_{\mathcal{R}}$ of) the law of $X^{\mathcal{R}}(s)$ and $\mu_{\mathcal{R}}$, then Jensen inequality to compare the ℓ^1 and ℓ^2 distances, we have

$$d_{\text{TV}}\left(\mathbb{P}_x\left(X^{\mathcal{R}}(s) = \cdot\right), \mu_{\mathcal{R}}\right) \leq \frac{\sqrt{\chi_{\mathcal{R}}}}{2} e^{-\gamma_{\mathcal{R}} s}.$$

Since $T_{\kappa_{\mathcal{R}}}^{\mathcal{R}}$ is an exponential variable of parameter κ , it then holds, for all non-negative s ,

$$d_{\text{TV}}\left(\mathbb{P}_x\left(X(T_{\kappa_{\mathcal{R}}}) = \cdot\right), \mu_{\mathcal{R}}\right) \leq \kappa s + \frac{\sqrt{\chi_{\mathcal{R}}}}{2} e^{-\gamma_{\mathcal{R}} s}.$$

Optimizing in $s \geq 0$, we get (20).

Next, denoting by ν the law of $X(T_{\kappa_{\mathcal{R}}})$ with starting distribution π , it holds, for all $t \geq 0$,

$$\mathbb{P}_{\nu}\left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*}\right) \leq \mathbb{P}_{\pi}(T_{\lambda_S} < T_{\kappa_{\mathcal{R}}}) + \mathbb{P}_{\pi}\left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*}\right).$$

By using Proposition 2.7 it follows that

$$\mathbb{P}_{\pi}\left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*}\right) \geq e^{-t} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} - \frac{\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\chi_{\mathcal{R}}}}{2\kappa}\right]_+\right) - \delta.$$

Finally, for all $\theta > 0$, we get

$$\mathbb{P}_{\pi}\left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*}\right) \leq \mathbb{P}_{\pi}(T_{\lambda_S} < T_{\kappa_{\mathcal{R}}}) + \mathbb{P}_{\pi}(T_{\lambda_S} \wedge T_{\kappa_{\mathcal{R}}} > \theta) + \mathbb{P}_{\nu}\left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*} - \theta\right).$$

By using Proposition 2.7 and the fact that $T_{\kappa_{\mathcal{R}}} \wedge T_{\lambda_S}$ is dominated by an exponential variable of parameter $\kappa \wedge \lambda$ we get

$$\begin{aligned} \mathbb{P}_{\pi}\left(T_{\lambda_S} > \frac{t}{\phi_{\mathcal{R},\lambda_S}^*}\right) &\leq \delta + e^{-(\kappa \wedge \lambda)\theta} + \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + \frac{\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\chi_{\mathcal{R}}}}{2\kappa}\right]_+\right) + \exp\left[-\sqrt{\lambda/\phi_{\mathcal{R},\lambda_S}^*}\right] \\ &\quad + e^{-t} \exp\left[\phi_{\mathcal{R},\lambda_S}^* \theta + \sqrt{\phi_{\mathcal{R},\lambda_S}^* / \lambda}\right], \end{aligned}$$

which gives our last desired inequality by choosing $\theta = (\phi_{\mathcal{R},\lambda_S}^* (\kappa \wedge \lambda))^{-1/2}$. \square

2.3 Soft capacities

To define our soft capacities we extend the conductance network (\mathcal{X}, c) , with conductances

$$c(x, y) = c(y, x) = \mu(x)w(x, y), \quad x, y \in \mathcal{X}, \quad x \neq y,$$

which is associated with the process X , into a larger network $(\tilde{\mathcal{X}}, \tilde{c})$ by attaching dangling edges (r, \bar{r}) and (s, \check{s}) to each $r \in \mathcal{R}$ and $s \in \mathcal{S}$. Given κ and λ we set to

$$\tilde{c}(r, \bar{r}) = \kappa \mu(r) \quad \text{and} \quad \tilde{c}(s, \check{s}) = \lambda \mu(s)$$

the conductances of these new edges. We call $\bar{\mathcal{R}}$ and $\check{\mathcal{S}}$ the collections of these extra nodes:

$$\bar{\mathcal{R}} = \{\bar{r} : r \in \mathcal{R}\}, \quad \check{\mathcal{S}} = \{\check{s} : s \in \mathcal{S}\}.$$

For any positive $\tilde{\kappa}$ and $\tilde{\lambda}$, by extending the probability measure μ defined on \mathcal{X} into a measure $\tilde{\mu}$ defined on

$$\tilde{\mathcal{X}} = \mathcal{X} \cup \bar{\mathcal{R}} \cup \check{\mathcal{S}}$$

by

$$\tilde{\mu}(\tilde{x}) = \begin{cases} \mu(x) & \text{if } \tilde{x} = x \text{ for some } x \in \mathcal{X}, \\ \kappa\mu(r)/\tilde{\kappa} & \text{if } \tilde{x} = \bar{r} \text{ for some } r \in \bar{\mathcal{R}}, \\ \lambda\mu(s)/\tilde{\lambda} & \text{if } \tilde{x} = \check{s} \text{ for some } s \in \check{\mathcal{S}}, \end{cases}$$

the extended network $(\tilde{\mathcal{X}}, \tilde{c})$ is then associated with the process \tilde{X} with rates $\tilde{w}(\tilde{x}, \tilde{y})$ given, for distinct \tilde{x} and \tilde{y} in $\tilde{\mathcal{X}}$, by

$$\tilde{w}(\tilde{x}, \tilde{y}) = \begin{cases} w(x, y) & \text{if } \tilde{x} = x \text{ and } \tilde{y} = y \text{ for some } x \text{ and } y \text{ in } \mathcal{X}, \\ \kappa & \text{if } \tilde{x} = r \text{ and } \tilde{y} = \bar{r} \text{ for some } r \text{ in } \bar{\mathcal{R}}, \\ \tilde{\kappa} & \text{if } \tilde{x} = \bar{r} \text{ and } \tilde{y} = r \text{ for some } r \text{ in } \bar{\mathcal{R}}, \\ \lambda & \text{if } \tilde{x} = s \text{ and } \tilde{y} = \check{s} \text{ for some } s \text{ in } \check{\mathcal{S}}, \\ \tilde{\lambda} & \text{if } \tilde{x} = \check{s} \text{ and } \tilde{y} = s \text{ for some } s \text{ in } \check{\mathcal{S}}, \\ 0 & \text{otherwise,} \end{cases}$$

and which is reversible with respect to $\tilde{\mu}$. (Note that $\tilde{\mu}$ is not a probability measure.)

Our soft capacities are then the (κ, λ) -capacities of [BG16].

Definition 2.9. The (κ, λ) -capacity $C_\kappa^\lambda(\mathcal{R}, \mathcal{S})$ is the capacity between the sets $\bar{\mathcal{R}}$ and $\check{\mathcal{S}}$ in the network $(\tilde{\mathcal{X}}, \tilde{c})$. According to Dirichlet principle, it is given by

$$\begin{aligned} C_\kappa^\lambda(\mathcal{R}, \mathcal{S}) &= \min \left\{ \frac{1}{2} \sum_{\tilde{x}, \tilde{y} \in \tilde{\mathcal{X}}} \tilde{c}(\tilde{x}, \tilde{y}) \left(\tilde{f}(\tilde{x}) - \tilde{f}(\tilde{y}) \right)^2 : \tilde{f}|_{\bar{\mathcal{R}}} = 1, \tilde{f}|_{\check{\mathcal{S}}} = 0 \right\} \\ &= \min_{f: \mathcal{X} \rightarrow \mathbb{R}} \left\{ \mathcal{D}(f) + \kappa \sum_{r \in \mathcal{R}} \mu(r) (f(r) - 1)^2 + \lambda \sum_{s \in \mathcal{S}} \mu(s) (f(s) - 0)^2 \right\} \\ &= \min_{f: \mathcal{X} \rightarrow \mathbb{R}} \left\{ \mathcal{D}(f) + \kappa \mu(\mathcal{R}) \mathbb{E}_{\mu_{\mathcal{R}}} \left[(f|_{\mathcal{R}} - 1)^2 \right] + \lambda \mu(\mathcal{S}) \mathbb{E}_{\mu_{\mathcal{S}}} \left[(f|_{\mathcal{S}} - 0)^2 \right] \right\}. \end{aligned} \quad (21)$$

Remarks:

- i. Since this definition depends on the conductances only, while it is naturally associated with the Markov process \tilde{X} , it does not depend on the choice of $\tilde{\kappa}$ and $\tilde{\lambda}$. $C_\kappa^\lambda(\mathcal{R}, \mathcal{S})$ depends on two parameters only, κ and λ , in which it is increasing.
- ii. Our soft capacities satisfy a two-sided variational principle. On one side they are given by Definition 2.9 as the infimum of some functional, and any test function will provide an upper bound. On the other side they are, by Thomson principle, the supremum of another functional on unitary flows from $\bar{\mathcal{R}}$ to $\check{\mathcal{S}}$, which are antisymmetric functions of oriented edges with null divergence in \mathcal{X} and total divergence in $\bar{\mathcal{R}}$ and $\check{\mathcal{S}}$ equal to 1 and -1 respectively, i.e., on antisymmetric functions $\tilde{\psi}: \tilde{\mathcal{X}} \times \tilde{\mathcal{X}} \rightarrow \mathbb{R}$ such that $\tilde{\psi}(\tilde{x}, \tilde{y}) = 0$ as soon as $\tilde{c}(\tilde{x}, \tilde{y}) = 0$ and

$$\operatorname{div}_x \tilde{\psi} = \sum_{\tilde{x} \in \tilde{\mathcal{X}}} \tilde{\psi}(x, \tilde{x}) = 0, \quad x \in \mathcal{X},$$

while

$$\sum_{r \in \mathcal{R}} \operatorname{div}_{\bar{r}} \tilde{\psi} = \sum_{r \in \mathcal{R}} \tilde{\psi}(\bar{r}, r) = 1 = - \sum_{s \in \mathcal{S}} \operatorname{div}_{\check{s}} \tilde{\psi} = \sum_{s \in \mathcal{S}} \tilde{\psi}(s, \check{s}).$$

With

$$\tilde{\mathcal{D}}(\tilde{\psi}) = \frac{1}{2} \sum_{\tilde{x}, \tilde{y} \in \tilde{\mathcal{X}}} \frac{\tilde{\psi}(\tilde{x}, \tilde{y})^2}{\tilde{c}(\tilde{x}, \tilde{y})}$$

the energy dissipated in $(\tilde{\mathcal{X}}, \tilde{c})$ by the flow $\tilde{\psi}$ and with $\tilde{\Psi}_1(\bar{\mathcal{R}}, \check{\mathcal{S}})$ the set of unitary flows from $\bar{\mathcal{R}}$ to $\check{\mathcal{S}}$, we get

$$C_\kappa^\lambda(\mathcal{R}, \mathcal{S}) = \max_{\tilde{\psi} \in \tilde{\Psi}_1(\bar{\mathcal{R}}, \check{\mathcal{S}})} \tilde{\mathcal{D}}(\tilde{\psi})^{-1}. \quad (22)$$

Any test flow provides then a lower bound on $C_\kappa^\lambda(\mathcal{R}, \mathcal{S})$.

- iii. By referring to \tilde{X} , the stopping time T_{λ_S} can also be defined as the hitting time of \tilde{S} . By defining $T_{\kappa_{\mathcal{R}}}$ in a symmetric way, the minimum in (21) is realized by the equilibrium potential V_{κ}^{λ} given by

$$V_{\kappa}^{\lambda}(x) = \mathbb{P}_x(T_{\kappa_{\mathcal{R}}} < T_{\lambda_S}), \quad x \in \mathcal{X},$$

whereas the maximum in (22) is realized by the associated normalized current.

As an application of Dirichlet principle we obtain an upper bound on our soft capacities, or of the asymptotic given for the spectral γ in Theorem 1. Let us define

$$\phi_{\kappa}^{\lambda} = \frac{C_{\kappa}^{\lambda}(\mathcal{R}, \mathcal{S})}{\mu(\mathcal{R})\mu(\mathcal{S})}$$

and recall Equation (17).

Lemma 2.10. *If $\mu(\mathcal{S}) \geq \mu(\mathcal{R})$ and $\epsilon_{\mathcal{R}, \mathcal{S}}^* \leq 1$, then, for all $\kappa \geq 0$, it holds*

$$\phi_{\kappa}^{\lambda} \leq \mu(\mathcal{S})^{-1} \phi_{\mathcal{R}, \mathcal{S}}^* \left\{ \frac{1 + \kappa/\gamma_{\mathcal{R}}}{1 - \epsilon_{\mathcal{R}, \mathcal{S}}^*} \right\} \quad \text{and} \quad \mu(\mathcal{S})^{-1} \leq 2. \quad (23)$$

In particular, under the hypotheses of Theorem 1 it holds $\phi_{\kappa}^{\lambda} = O(\phi_{\mathcal{R}, \mathcal{S}}^*)$.

Proof: By taking $h_{\mathcal{R}, \mathcal{S}}^*$ as test function in Equation (21), and following Convention 2.5, we get

$$C_{\kappa}^{\lambda}(\mathcal{R}, \mathcal{S}) \leq \mathcal{D}(h_{\mathcal{R}, \mathcal{S}}^*) + \kappa \mu(\mathcal{R}) \text{Var}_{\mu_{\mathcal{R}}}(h_{\mathcal{R}, \mathcal{S}}^*).$$

The second term is straightforwardly bounded by using Proposition 2.6 which reads

$$\kappa \mu(\mathcal{R}) \text{Var}_{\mu_{\mathcal{R}}}(h_{\mathcal{R}, \mathcal{S}}^*) \leq \mu(\mathcal{R}) \phi_{\mathcal{R}, \mathcal{S}}^* \frac{\kappa/\gamma_{\mathcal{R}}}{1 - \epsilon_{\mathcal{R}, \mathcal{S}}^*}.$$

As far as the first term is concerned we write

$$\begin{aligned} \mathcal{D}(h_{\mathcal{R}, \mathcal{S}}^*) &= \frac{1}{2} \sum_{x, y \in \mathcal{R} \setminus \mathcal{S}} c(x, y) [h_{\mathcal{R}, \mathcal{S}}^*(x) - h_{\mathcal{R}, \mathcal{S}}^*(y)]^2 + \sum_{\substack{x \in \mathcal{R} \setminus \mathcal{S}, \\ y \in \mathcal{S}}} c(x, y) h_{\mathcal{R}, \mathcal{S}}^{*2}(x) \\ &= \mu(\mathcal{R}) \langle h_{\mathcal{R}, \mathcal{S}}^*, -\mathcal{L}_{\mathcal{R}, \mathcal{S}}^* h_{\mathcal{R}, \mathcal{S}}^* \rangle_{\mathcal{R}} = \mu(\mathcal{R}) \phi_{\mathcal{R}, \mathcal{S}}^* \|h_{\mathcal{R}, \mathcal{S}}^*\|_{\mathcal{R}}^2 = \mu(\mathcal{R}) \phi_{\mathcal{R}, \mathcal{S}}^* \left(1 + \text{Var}_{\mu_{\mathcal{R}}}(h_{\mathcal{R}, \mathcal{S}}^*)\right) \\ &\leq \mu(\mathcal{R}) \phi_{\mathcal{R}, \mathcal{S}}^* \left(1 + \frac{\epsilon_{\mathcal{R}, \mathcal{S}}^*}{1 - \epsilon_{\mathcal{R}, \mathcal{S}}^*}\right) = \mu(\mathcal{R}) \phi_{\mathcal{R}, \mathcal{S}}^* \frac{1}{1 - \epsilon_{\mathcal{R}, \mathcal{S}}^*}, \end{aligned}$$

where the inequality is again an application Proposition 2.6. Observing that $\mathcal{X} = \mathcal{R} \cup \mathcal{S}$ and $\mu(\mathcal{S}) \geq \mu(\mathcal{R})$ imply $\mu(\mathcal{S}) \geq 1/2$, this leads to (23). The last comment is only based on the fact that hypotheses (H) include $\epsilon_{\mathcal{R}, \mathcal{S}}^* \ll 1$ and $\kappa \ll \gamma_{\mathcal{R}}$ is assumed in Theorem 1. \square

Remark 2.11. *All the estimates of Section 2 are “symmetric” in \mathcal{R} and \mathcal{S} . This means that the same results hold when the role of \mathcal{R} and \mathcal{S} are exchanged as well as those of κ and λ .*

We note however that there is an asymmetry in hypotheses (H). It will start to play a role in the next section.

3 Proof of Theorem 1

3.1 Spectral gap estimates

The statement contained in Theorem 1 concerning the spectral gap γ is a consequence of the quantitative Theorem 4 together with Lemma 2.10, estimates (13) and (18) as well as Remark 2.11. Note also that hypotheses (H) include $\phi_{\mathcal{R}, \mathcal{S}}^* \ll \gamma_{\mathcal{S}}$. This is where the asymmetry of (H) matters.

Theorem 4. *The spectral gap γ satisfies*

$$\gamma \geq \phi_\kappa^\lambda \left\{ 1 + \frac{\kappa + \phi_\kappa^\lambda (1 + \mu(\mathcal{R} \cap \mathcal{S}))}{\gamma_{\mathcal{R}}} + \frac{\lambda + \phi_\kappa^\lambda (1 + \mu(\mathcal{R} \cap \mathcal{S}))}{\gamma_{\mathcal{S}}} \right\}^{-1}. \quad (24)$$

In addition, when $\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^$ and $\epsilon_{\mathcal{S}, \kappa \mathcal{R}}^*$ are smaller than 1, we obtain the converse inequality*

$$\gamma \leq \phi_\kappa^\lambda \left\{ 1 - \frac{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}{\kappa} - \frac{\phi_{\mathcal{S}, \kappa \mathcal{R}}^*}{\lambda} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}{1 - \epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{S}, \kappa \mathcal{R}}^*}{1 - \epsilon_{\mathcal{S}, \kappa \mathcal{R}}^*}} \right\}^{-2} \quad (25)$$

as soon as the braced sum is positive.

Proof of the lower bound: We recall that γ satisfies the variational principle

$$\gamma := \min_{\substack{f: \mathcal{X} \rightarrow \mathbb{R}, \\ \text{Var}_\mu(f) \neq 0}} \frac{\mathcal{D}(f)}{\text{Var}_\mu(f)}, \quad (26)$$

We look for a Poincaré inequality, i.e. an inequality of the form

$$\forall f \in \ell^2(\mu), \quad \text{Var}_\mu(f) \leq k \mathcal{D}(f),$$

which is equivalent, by mean of Equation (26), to a lower bound on the spectral gap. Let Z, Z' two i.i.d. random variables with same law μ . The variance of f can be written as

$$\text{Var}_\mu(f) = \mathbb{E}_\mu[f^2] - \mathbb{E}_\mu^2[f] = \frac{1}{2} \mathbb{E}[(f(Z) - f(Z'))^2]. \quad (27)$$

From the r.h.s. of (27) we get

$$\begin{aligned} \text{Var}_\mu(f) &\leq \frac{1}{2} \mathbb{E}[(f(Z) - f(Z'))^2 \mathbf{1}_{\{Z, Z' \in \mathcal{R}\}}] + \frac{1}{2} \mathbb{E}[(f(Z) - f(Z'))^2 \mathbf{1}_{\{Z, Z' \in \mathcal{S}\}}] \\ &\quad + \mathbb{E}[(f(Z) - f(Z'))^2 \mathbf{1}_{\{Z \in \mathcal{R}, Z' \in \mathcal{S}\}}] \\ &= \mu^2(\mathcal{R}) \text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \mu^2(\mathcal{S}) \text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) \\ &\quad + \mu(\mathcal{R})\mu(\mathcal{S}) \left(\mathbb{E}_{\mu_{\mathcal{R}}}[f|_{\mathcal{R}}^2] + \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}^2] - 2\mathbb{E}_{\mu_{\mathcal{R}}}[f|_{\mathcal{R}}] \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}] \right) \\ &= \mu(\mathcal{R}) \left(\mu(\mathcal{R}) + \mu(\mathcal{S}) \right) \text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \mu(\mathcal{S}) \left(\mu(\mathcal{R}) + \mu(\mathcal{S}) \right) \text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) \\ &\quad + \mu(\mathcal{R})\mu(\mathcal{S}) \left(\mathbb{E}_{\mu_{\mathcal{R}}}^2[f|_{\mathcal{R}}] + \mathbb{E}_{\mu_{\mathcal{S}}}^2[f|_{\mathcal{S}}] - 2\mathbb{E}_{\mu_{\mathcal{R}}}[f|_{\mathcal{R}}] \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}] \right) \\ &= \mu(\mathcal{R}) \left(\mu(\mathcal{R}) + \mu(\mathcal{S}) \right) \text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \mu(\mathcal{S}) \left(\mu(\mathcal{R}) + \mu(\mathcal{S}) \right) \text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) \\ &\quad + \mu(\mathcal{R})\mu(\mathcal{S}) \left(\mathbb{E}_{\mu_{\mathcal{R}}}[f|_{\mathcal{R}}] - \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}] \right)^2. \end{aligned} \quad (28)$$

As a test function in Dirichlet principle (21) we pick

$$\tilde{f} = \frac{f - \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}]}{\mathbb{E}_{\mu_{\mathcal{R}}}[f|_{\mathcal{R}}] - \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}]}, \quad (29)$$

which is such that $\mathbb{E}_{\mu_{\mathcal{R}}}[\tilde{f}|_{\mathcal{R}}] = 1$ and $\mathbb{E}_{\mu_{\mathcal{S}}}[\tilde{f}|_{\mathcal{S}}] = 0$. By plugging (29) in (21) we get

$$\begin{aligned} C_\kappa^\lambda(\mathcal{R}, \mathcal{S}) &\leq \mathcal{D}(\tilde{f}) + \kappa \mu(\mathcal{R}) \text{Var}_{\mu_{\mathcal{R}}}(\tilde{f}|_{\mathcal{R}}) + \lambda \mu(\mathcal{S}) \text{Var}_{\mu_{\mathcal{S}}}(\tilde{f}|_{\mathcal{S}}) \\ &= \left(\mathbb{E}_{\mu_{\mathcal{R}}}[f|_{\mathcal{R}}] - \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}] \right)^{-2} \left(\mathcal{D}(f) + \kappa \mu(\mathcal{R}) \text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \lambda \mu(\mathcal{S}) \text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) \right) \end{aligned}$$

which actually is an upper bound on $\left(\mathbb{E}_{\mu_{\mathcal{R}}}[f|_{\mathcal{R}}] - \mathbb{E}_{\mu_{\mathcal{S}}}[f|_{\mathcal{S}}] \right)^2$. Bounding from above the last term in (28) we then get

$$\text{Var}_\mu(f) \leq \mu(\mathcal{R}) \left(\mu(\mathcal{R}) + \mu(\mathcal{S}) \right) \text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \mu(\mathcal{S}) \left(\mu(\mathcal{R}) + \mu(\mathcal{S}) \right) \text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}})$$

$$\begin{aligned}
& + \frac{\mu(\mathcal{R})\mu(\mathcal{S})}{C_\kappa^\lambda(\mathcal{R}, \mathcal{S})} \left(\mathcal{D}(f) + \kappa\mu(\mathcal{R})\text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \lambda\mu(\mathcal{S})\text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) \right) \\
& = \frac{\mathcal{D}(f)}{\phi_\kappa^\lambda} + \mu(\mathcal{R})\text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) \left(\frac{\kappa}{\phi_\kappa^\lambda} + (1 + \mu(\mathcal{R} \cap \mathcal{S})) \right) \\
& \quad + \mu(\mathcal{S})\text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) \left(\frac{\lambda}{\phi_\kappa^\lambda} + (1 + \mu(\mathcal{R} \cap \mathcal{S})) \right) \\
& \leq \frac{\mathcal{D}(f)}{\phi_\kappa^\lambda} + \frac{\mu(\mathcal{R})\mathcal{D}_{\mathcal{R}}(f|_{\mathcal{R}})}{\gamma_{\mathcal{R}}} \left(\frac{\kappa}{\phi_\kappa^\lambda} + (1 + \mu(\mathcal{R} \cap \mathcal{S})) \right) + \frac{\mu(\mathcal{S})\mathcal{D}_{\mathcal{S}}(f|_{\mathcal{S}})}{\gamma_{\mathcal{S}}} \left(\frac{\lambda}{\phi_\kappa^\lambda} + (1 + \mu(\mathcal{R} \cap \mathcal{S})) \right) \\
& \leq \frac{\mathcal{D}(f)}{\phi_\kappa^\lambda} + \frac{\mathcal{D}(f)}{\gamma_{\mathcal{R}}} \left(\frac{\kappa}{\phi_\kappa^\lambda} + (1 + \mu(\mathcal{R} \cap \mathcal{S})) \right) + \frac{\mathcal{D}(f)}{\gamma_{\mathcal{S}}} \left(\frac{\lambda}{\phi_\kappa^\lambda} + (1 + \mu(\mathcal{R} \cap \mathcal{S})) \right)
\end{aligned}$$

where, in the second inequality, the variational characterization of spectral gaps $\gamma_{\mathcal{R}}$ and $\gamma_{\mathcal{S}}$ given by Equation (3) has been used, whereas in the third one we exploited the following fact:

$$\mathcal{D}(f) \geq \mu(\mathcal{R})\mathcal{D}_{\mathcal{R}}(f|_{\mathcal{R}}), \quad \mathcal{D}(f) \geq \mu(\mathcal{S})\mathcal{D}_{\mathcal{S}}(f|_{\mathcal{S}}).$$

Factorising $\mathcal{D}(f)/\phi_\kappa^\lambda$ we get the Poincaré inequality

$$\text{Var}_\mu(f) \leq \frac{\mathcal{D}(f)}{\phi_\kappa^\lambda} \left(1 + \frac{\kappa + \phi_\kappa^\lambda(1 + \mu(\mathcal{R} \cap \mathcal{S}))}{\gamma_{\mathcal{R}}} + \frac{\lambda + \phi_\kappa^\lambda(1 + \mu(\mathcal{R} \cap \mathcal{S}))}{\gamma_{\mathcal{S}}} \right).$$

□

Before giving the proof of the upper bound, we observe that the proof of this Poincaré inequality carries over to the case of any finite covering of irreducible sets $\mathcal{X} = \cup_{i < m} \mathcal{R}_i$.

Proposition 3.1. *If $\mathcal{R}_0, \mathcal{R}_1, \dots, \mathcal{R}_{m-1}$ are subsets of \mathcal{X} such that $X_{\mathcal{R}_i}$ is irreducible for each $i < m$, and $\kappa_0, \kappa_1, \dots, \kappa_{m-1}$ are any positive numbers then, with*

$$\phi_i^j = \phi_j^i = \frac{C_{\kappa_i}^{\kappa_j}(\mathcal{R}_i, \mathcal{R}_j)}{\mu(\mathcal{R}_i)\mu(\mathcal{R}_j)}, \quad i \neq j, \quad i, j < m,$$

$$\frac{1}{\phi} = \frac{1}{2} \sum_{i,j:i \neq j} \frac{1}{\phi_i^j}$$

and, for each $i < m$,

$$\frac{1}{\phi_i} = \sum_{j \neq i} \frac{1}{\phi_i^j} \leq \frac{1}{\phi},$$

it holds

$$\gamma \geq \phi \left\{ 1 + \sum_{i < m} \frac{\phi}{\gamma_i} \left(\frac{\kappa_i}{\phi_i} + \sum_{j < m} \mu(\mathcal{R}_j) \right) \right\}^{-1} \geq \phi \left\{ 1 + \sum_{i < m} \frac{\kappa_i + \phi \sum_{j < m} \mu(\mathcal{R}_j)}{\gamma_i} \right\}^{-1}.$$

Proof of the upper bound in Theorem 4: We want to exploit the variational principle (26), so we look for an upper bound on $\mathcal{D}(f)$ and a lower bound on $\text{Var}_\mu(f)$ for a suitable function f . From (27) we get

$$\begin{aligned}
\text{Var}_\mu(f) & \geq \frac{1}{2} \mathbb{E} \left[(f(Z) - f(Z'))^2 (\mathbb{1}_{\{Z \in \mathcal{R}, Z' \in \mathcal{S}\}} + \mathbb{1}_{\{Z \in \mathcal{S}, Z' \in \mathcal{R}\}} - \mathbb{1}_{\{Z, Z' \in \mathcal{R} \cap \mathcal{S}\}}) \right] \\
& = \mathbb{E} \left[(f(Z) - f(Z'))^2 \mathbb{1}_{\{Z \in \mathcal{R}, Z' \in \mathcal{S}\}} \right] - \frac{1}{2} \mathbb{E} \left[(f(Z) - f(Z'))^2 \mathbb{1}_{\{Z, Z' \in \mathcal{R} \cap \mathcal{S}\}} \right] \\
& = \mu(\mathcal{R})\mu(\mathcal{S}) \left(\mathbb{E}_{\mu_{\mathcal{R}}} [f|_{\mathcal{R}}^2] + \mathbb{E}_{\mu_{\mathcal{S}}} [f|_{\mathcal{S}}^2] \right) - 2\mu(\mathcal{R})\mu(\mathcal{S}) \mathbb{E}_{\mu_{\mathcal{R}}} [f|_{\mathcal{R}}] \mathbb{E}_{\mu_{\mathcal{S}}} [f|_{\mathcal{S}}] \\
& \quad - \mu^2(\mathcal{R} \cap \mathcal{S}) \text{Var}_{\mu_{\mathcal{R} \cap \mathcal{S}}}(f|_{\mathcal{R} \cap \mathcal{S}}) \\
& = \mu(\mathcal{R})\mu(\mathcal{S}) \left(\text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \mathbb{E}_{\mu_{\mathcal{R}}}^2 [f|_{\mathcal{R}}] \right) + \mu(\mathcal{R})\mu(\mathcal{S}) \left(\text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) + \mathbb{E}_{\mu_{\mathcal{S}}}^2 [f|_{\mathcal{S}}] \right) \\
& \quad - 2\mu(\mathcal{R})\mu(\mathcal{S}) \mathbb{E}_{\mu_{\mathcal{R}}} [f|_{\mathcal{R}}] \mathbb{E}_{\mu_{\mathcal{S}}} [f|_{\mathcal{S}}] - \mu^2(\mathcal{R} \cap \mathcal{S}) \text{Var}_{\mu_{\mathcal{R} \cap \mathcal{S}}}(f|_{\mathcal{R} \cap \mathcal{S}}) \\
& = \mu(\mathcal{R})\mu(\mathcal{S}) \left(\text{Var}_{\mu_{\mathcal{R}}}(f|_{\mathcal{R}}) + \text{Var}_{\mu_{\mathcal{S}}}(f|_{\mathcal{S}}) \right) - \mu^2(\mathcal{R} \cap \mathcal{S}) \text{Var}_{\mu_{\mathcal{R} \cap \mathcal{S}}}(f|_{\mathcal{R} \cap \mathcal{S}})
\end{aligned} \tag{30}$$

$$+ \mu(\mathcal{R})\mu(\mathcal{S}) \left(\mathbb{E}_{\mu_{\mathcal{R}}} [f|_{\mathcal{R}}] - \mathbb{E}_{\mu_{\mathcal{S}}} [f|_{\mathcal{S}}] \right)^2$$

To get a suitable lower bound of (30), we write

$$\begin{aligned} \mu^2(\mathcal{R} \cap \mathcal{S}) \text{Var}_{\mu_{\mathcal{R} \cap \mathcal{S}}} (f|_{\mathcal{R} \cap \mathcal{S}}) &= \frac{1}{2} \mathbb{E} \left[(f(Z) - f(Z'))^2 \mathbf{1}_{\{Z, Z' \in \mathcal{R} \cap \mathcal{S}\}} \right] \\ &\leq \frac{1}{2} \mathbb{E} \left[(f(Z) - f(Z'))^2 \mathbf{1}_{\{Z, Z' \in \mathcal{R}\}} \right] \\ &= \mu^2(\mathcal{R}) \text{Var}_{\mu_{\mathcal{R}}} (f|_{\mathcal{R}}) \leq \mu(\mathcal{R})\mu(\mathcal{S}) \text{Var}_{\mu_{\mathcal{R}}} (f|_{\mathcal{R}}), \end{aligned} \quad (31)$$

assuming, without loss of generality, $\mu(\mathcal{R}) \leq \mu(\mathcal{S})$. Then, by using (31) to bound the second term in (30), dropping one term and taking as test function the equilibrium potential V_{κ}^{λ} , we get

$$\text{Var}_{\mu}(V_{\kappa}^{\lambda}) \geq \mu(\mathcal{R})\mu(\mathcal{S}) \left(\mathbb{E}_{\mu_{\mathcal{R}}} [V_{\kappa}^{\lambda}|_{\mathcal{R}}] - \mathbb{E}_{\mu_{\mathcal{S}}} [V_{\kappa}^{\lambda}|_{\mathcal{S}}] \right)^2. \quad (32)$$

Next, we have the two following simples lemmas.

Lemma 3.2. *If $\epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^* < 1$, then*

$$\mathbb{E}_{\mu_{\mathcal{R}}} [V_{\kappa}^{\lambda}|_{\mathcal{R}}] \geq 1 - \frac{\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}{\kappa} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}{1 - \epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}}.$$

Proof: Denoting by σ_{κ} an exponential time of rate κ that is independent of X , it holds

$$\mathbb{E}_{\mu_{\mathcal{R}}} [V_{\kappa}^{\lambda}|_{\mathcal{R}}] = \mathbb{P}_{\mu_{\mathcal{R}}} (T_{\kappa_{\mathcal{R}}} < T_{\lambda_{\mathcal{S}}}) = \mathbb{P}_{\mu_{\mathcal{R}}} (\ell_{\mathcal{R}}(T_{\kappa_{\mathcal{R}}}) < \ell_{\mathcal{R}}(T_{\lambda_{\mathcal{S}}})) = \mathbb{P}_{\mu_{\mathcal{R}}} (\sigma_{\kappa} < T_{\lambda_{\mathcal{S}}}^{\mathcal{R}}).$$

Let \mathbb{P}^{opt} be an optimal coupling at time $t = 0$ between $X^{\mu_{\mathcal{R}}}$ and $X^{\mu_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}$ that start with initial distribution $\mu_{\mathcal{R}}$ and $\mu_{\mathcal{R}, \lambda_{\mathcal{S}}}^*$ and evolve jointly if they have the same starting configuration. By the union bound, Proposition (2.6) and the convexity of $x \mapsto \frac{1}{1+x}$ we get

$$\begin{aligned} \mathbb{P}_{\mu_{\mathcal{R}}} (\sigma_{\kappa} < T_{\lambda_{\mathcal{S}}}^{\mathcal{R}}) &\geq \mathbb{P}_{\mu_{\mathcal{R}, \lambda}^*} (\sigma_{\kappa} < T_{\lambda_{\mathcal{S}}}^{\mathcal{R}}) - \mathbb{P}^{opt} (X^{\mu_{\mathcal{R}}}(0) \neq X^{\mu_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}(0)) \\ &\geq \frac{\kappa}{\kappa + \phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}{1 - \epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}} \geq 1 - \frac{\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}{\kappa} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}{1 - \epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}} \end{aligned}$$

□

Lemma 3.3. *If $\epsilon_{\mathcal{S}, \kappa_{\mathcal{R}}}^* < 1$, then*

$$\mathbb{E}_{\mu_{\mathcal{S}}} [V_{\kappa}^{\lambda}|_{\mathcal{S}}] \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{S}, \kappa_{\mathcal{R}}}^*}{1 - \epsilon_{\mathcal{S}, \kappa_{\mathcal{R}}}^*}} + \frac{\phi_{\mathcal{S}, \kappa_{\mathcal{R}}}^*}{\lambda}.$$

The proof of this lemma is similar to that of the previous one and we omit it. By plugging the two bounds provided by Lemma 3.2 and Lemma 3.3 into (32) we obtain

$$\text{Var}_{\mu}(V_{\kappa}^{\lambda}) \geq \mu(\mathcal{R})\mu(\mathcal{S}) \left\{ 1 - \frac{\phi_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}{\kappa} - \frac{\phi_{\mathcal{S}, \kappa_{\mathcal{R}}}^*}{\lambda} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}{1 - \epsilon_{\mathcal{R}, \lambda_{\mathcal{S}}}^*}} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{S}, \kappa_{\mathcal{R}}}^*}{1 - \epsilon_{\mathcal{S}, \kappa_{\mathcal{R}}}^*}} \right\}^2 \quad (33)$$

as soon as the braced quantity is positive. The upper bound on $\mathcal{D}(V_{\kappa}^{\lambda})$ is straightforward: since the functional associated with the Dirichlet principle, which reaches its minimum $C_{\kappa}^{\lambda}(\mathcal{R}, \mathcal{S})$ in V_{κ}^{λ} , is larger than \mathcal{D} , we get

$$\mathcal{D}(V_{\kappa}^{\lambda}) \leq C_{\kappa}^{\lambda}(\mathcal{R}, \mathcal{S}). \quad (34)$$

Inequality (25) follows from Equation (26) together with (33) and (34). □

3.2 Exit rate estimates

To prove the last part of Theorem 1 we first give sharp upper and lower bounds for $\phi_{\mathcal{R},\lambda_S}^*$. These are sharp bounds in the sense that they have the same asymptotic behaviour as soon as κ and λ are chosen in the right window.

Theorem 5. *The exit rate $\phi_{\mathcal{R},\lambda_S}^*$ satisfies*

$$\phi_{\mathcal{R},\lambda_S}^* \geq \frac{C_\kappa^\lambda(\mathcal{R},\mathcal{S})}{\mu(\mathcal{R})} \left\{ \frac{1 - 2\mu(\mathcal{S})^{-1}\phi_{\mathcal{R},\lambda_S}^*/\lambda}{(1 - \phi_{\mathcal{R},\lambda_S}^*/\lambda)^2} \right\} \left\{ 1 + \frac{\kappa + \phi_\kappa^\lambda(1 + \mu(\mathcal{R} \cap \mathcal{S}))}{\gamma_{\mathcal{R}}} + \frac{\lambda + \phi_\kappa^\lambda(1 + \mu(\mathcal{R} \cap \mathcal{S}))}{\gamma_{\mathcal{S}}} \right\}^{-1}.$$

In addition, when $\epsilon_{\mathcal{R},\lambda_S}^* < 1$, we get the converse inequality

$$\phi_{\mathcal{R},\lambda_S}^* \leq \frac{C_\kappa^\lambda(\mathcal{R},\mathcal{S})}{\mu(\mathcal{R})} \left\{ 1 + \frac{\phi_{\mathcal{R},\lambda_S}^*}{\lambda} \right\} \left\{ 1 - \frac{\phi_{\mathcal{R},\lambda_S}^*}{\kappa} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} \right\}^{-2}$$

as soon as the last braced sum is positive.

Proof of the lower bound: We consider the process \tilde{X} introduced in Section 2.3 in the limiting case $\tilde{\kappa} = +\infty$. In this case $\tilde{\mathcal{X}}$ can be redefined as $\tilde{\mathcal{X}} = \mathcal{X} \cup \mathcal{S}$ and we call $\tilde{\mathcal{L}}$ the generator of this process \tilde{X} on $\tilde{\mathcal{X}}$. According with the notation of Section 2.1, we denote by $\tilde{\phi}_{\mathcal{X},\lambda_S}^*$ and $\tilde{\mu}_{\mathcal{X},\lambda_S}^*$ respectively the smallest eigenvalue and the corresponding normalized left eigenvector (i.e., the quasi-stationary distribution associated with \mathcal{X} and λ_S) of $-\tilde{\mathcal{L}}_{\mathcal{X},\lambda_S}^*$, which is defined by

$$(\tilde{\mathcal{L}}_{\mathcal{X},\lambda_S}^* f)(x) = -\lambda_S(x)f(x) + \sum_{y \in \mathcal{X}} w(x,y)[f(y) - f(x)], \quad f : \mathcal{X} \rightarrow \mathbb{R}, \quad x \in \mathcal{X}.$$

For any probability ν on \mathcal{X} , $\tilde{\phi}_{\mathcal{X},\lambda_S}^*$ is the decay rate of the survival probability $\mathbb{P}_\nu(T_{\lambda_S} > t)$:

$$\lim_{t \rightarrow \infty} -\frac{1}{t} \ln \mathbb{P}_\nu(T_{\lambda_S} > t) = \tilde{\phi}_{\mathcal{X},\lambda_S}^*.$$

Since

$$\mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*}(T_{\lambda_S} > t) \geq \mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*}(\ell_{\mathcal{R}}(T_{\lambda_S}) > t) = \mathbb{P}_{\mu_{\mathcal{R},\lambda_S}^*}(T_{\lambda_S}^{\mathcal{R}} > t) = e^{-t\phi_{\mathcal{R},\lambda_S}^*},$$

it follows that

$$\phi_{\mathcal{R},\lambda_S}^* \geq \tilde{\phi}_{\mathcal{X},\lambda_S}^*. \quad (35)$$

To give a lower bound on $\tilde{\phi}_{\mathcal{X},\lambda_S}^*$, we use its variational representation

$$\tilde{\phi}_{\mathcal{X},\lambda_S}^* = \min_{f: \mathcal{X} \rightarrow \mathbb{R}} \frac{\langle f, -\tilde{\mathcal{L}}_{\mathcal{X},\lambda_S}^* f \rangle}{\|f\|^2} = \min_{f: \mathcal{X} \rightarrow \mathbb{R}} \frac{\mathcal{D}(f) + \|\sqrt{\lambda_S} f\|^2}{\|f\|^2}. \quad (36)$$

Indeed, for any $f : \mathcal{X} \rightarrow \mathbb{R}$,

$$\langle f, -\tilde{\mathcal{L}}_{\mathcal{X},\lambda_S}^* f \rangle = \frac{1}{2} \sum_{x,y \in \mathcal{X}} \mu(x)w(x,y)[f(x) - f(y)]^2 + \sum_{x \in \mathcal{X}} \mu(x)\lambda_S(x)f^2(x).$$

Since the minimum is reached in

$$\tilde{h}_{\mathcal{X},\lambda_S}^* : x \in \mathcal{X} \mapsto \tilde{\mu}_{\mathcal{X},\lambda_S}^*(x)/\mu(x).$$

it holds

$$\tilde{\phi}_{\mathcal{X},\lambda_S}^* \geq \frac{\mathcal{D}(\tilde{h}_{\mathcal{X},\lambda_S}^*)}{\|\tilde{h}_{\mathcal{X},\lambda_S}^*\|^2} = \frac{\|\tilde{h}_{\mathcal{X},\lambda_S}^*\|^2 - 1}{\|\tilde{h}_{\mathcal{X},\lambda_S}^*\|^2} \frac{\mathcal{D}(\tilde{h}_{\mathcal{X},\lambda_S}^*)}{\text{Var}_\mu(\tilde{h}_{\mathcal{X},\lambda_S}^*)} \geq \left(1 - \frac{1}{\|\tilde{h}_{\mathcal{X},\lambda_S}^*\|^2}\right) \gamma. \quad (37)$$

Theorem 4 provides a lower bound on γ and it remains to give a lower bound on $\|\tilde{h}_{\mathcal{X},\lambda_S}^*\|$. To estimate this norm we restrict the sum to $\mathcal{X} \setminus \mathcal{S}$ and we use Jensen's inequality:

$$\|\tilde{h}_{\mathcal{X},\lambda_S}^*\|^2 \geq \mu(\mathcal{X} \setminus \mathcal{S}) \sum_{x \in \mathcal{X} \setminus \mathcal{S}} \mu_{\mathcal{X} \setminus \mathcal{S}}(x) \left(\frac{\tilde{\mu}_{\mathcal{X},\lambda_S}^*(x)}{\mu(x)} \right)^2$$

$$\geq \mu(\mathcal{X} \setminus \mathcal{S}) \left(\sum_{x \in \mathcal{X} \setminus \mathcal{S}} \mu_{\mathcal{X} \setminus \mathcal{S}}(x) \frac{\tilde{\mu}_{\mathcal{X}, \lambda \mathcal{S}}^*(x)}{\mu(x)} \right)^2 = \frac{(\tilde{\mu}_{\mathcal{X}, \lambda \mathcal{S}}^*(\mathcal{X} \setminus \mathcal{S}))^2}{\mu(\mathcal{X} \setminus \mathcal{S})}. \quad (38)$$

Now, the last identity of Lemma 2.1 reads in this case $\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^* = \lambda \tilde{\mu}_{\mathcal{X}, \lambda \mathcal{S}}^*(\mathcal{S})$. Therefore, using Inequality (35) again,

$$\tilde{\mu}_{\mathcal{X}, \lambda \mathcal{S}}^*(\mathcal{X} \setminus \mathcal{S}) = 1 - \frac{\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^*}{\lambda} \geq 1 - \frac{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}{\lambda}. \quad (39)$$

It follows from (38) and (39) that

$$\begin{aligned} 1 - \frac{1}{\|\tilde{h}_{\mathcal{X}, \lambda \mathcal{S}}^*\|^2} &\geq 1 - \frac{1 - \mu(\mathcal{S})}{(1 - \phi_{\mathcal{R}, \lambda \mathcal{S}}^*/\lambda)^2} = \frac{\mu(\mathcal{S}) - 2\phi_{\mathcal{R}, \lambda \mathcal{S}}^*/\lambda + (\phi_{\mathcal{R}, \lambda \mathcal{S}}^*/\lambda)^2}{(1 - \phi_{\mathcal{R}, \lambda \mathcal{S}}^*/\lambda)^2} \\ &\geq \mu(\mathcal{S}) \frac{1 - 2\mu(\mathcal{S})^{-1}\phi_{\mathcal{R}, \lambda \mathcal{S}}^*/\lambda}{(1 - \phi_{\mathcal{R}, \lambda \mathcal{S}}^*/\lambda)^2}. \end{aligned} \quad (40)$$

Finally from (35), (37) and (40) together with (24) we get the desired result. \square

Proof of the upper bound: The proof is made of two steps: first we estimate $\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^*$ via a variational principle, then we look for an upper bound of $\phi_{\mathcal{R}, \lambda \mathcal{S}}^*$ in terms of $\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^*$. By taking as test function in (36) the equilibrium potential V_κ^λ we get

$$\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^* \leq \frac{\mathcal{D}(V_\kappa^\lambda) + \|\sqrt{\lambda \mathcal{S}} V_\kappa^\lambda\|^2}{\|V_\kappa^\lambda\|^2} \leq \frac{C_\kappa^\lambda(\mathcal{R}, \mathcal{S})}{\|V_\kappa^\lambda\|^2}. \quad (41)$$

The denominator in (41) is lower bounded by

$$\|V_\kappa^\lambda\|^2 \geq \mu(\mathcal{R}) \sum_{x \in \mathcal{R}} \mu_{\mathcal{R}}(x) V_\kappa^\lambda(x)^2 \geq \mu(\mathcal{R}) \mu_{\mathcal{R}}(V_\kappa^\lambda | \mathcal{R})^2 \geq \mu(\mathcal{R}) \left\{ 1 - \frac{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}{\kappa} - \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}{1 - \epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}} \right\}^2, \quad (42)$$

where we used Jensen inequality, Lemma 3.2 and the assumption that the braced sum was positive.

To bound $\phi_{\mathcal{R}, \lambda \mathcal{S}}^*$ with $\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^*$, we write, for any $0 \leq \beta \leq 1$ and denoting again by σ_λ an exponential time of rate λ that is independent of X ,

$$\begin{aligned} \mathbb{P}_{\mu_{\mathcal{R}, \lambda \mathcal{S}}^*}(T_{\lambda \mathcal{S}} > t) &\leq \mathbb{P}_{\mu_{\mathcal{R}, \lambda}^*}(T_{\lambda \mathcal{S}}^{\mathcal{R}} + \sigma_\lambda > \beta t + (1 - \beta)t) \\ &\leq \mathbb{P}_{\mu_{\mathcal{R}, \lambda}^*}(T_{\lambda \mathcal{S}}^{\mathcal{R}} > \beta t) + \mathbb{P}_{\mu_{\mathcal{R}, \lambda}^*}(\sigma_\lambda > (1 - \beta)t) \\ &= \exp(-\phi_{\mathcal{R}, \lambda \mathcal{S}}^* \beta t) + \exp(-\lambda(1 - \beta)t). \end{aligned}$$

Recall that $\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^*$ is the decay rate of the survival probability $\mathbb{P}_{\mu_{\mathcal{R}, \lambda \mathcal{S}}^*}(T_{\lambda \mathcal{S}} > t)$. If we choose β such that $\phi_{\mathcal{R}, \lambda \mathcal{S}}^* \beta \leq \lambda(1 - \beta)$, then $\phi_{\mathcal{R}, \lambda \mathcal{S}}^* \beta$ is the decay rate of the right-hand side in the previous inequality and we can conclude $\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^* \geq \beta \phi_{\mathcal{R}, \lambda \mathcal{S}}^*$. We then choose $\beta = \lambda/(\phi_{\mathcal{R}, \lambda \mathcal{S}}^* + \lambda)$, which gives us

$$\phi_{\mathcal{R}, \lambda \mathcal{S}}^* \leq \frac{\tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^*}{\beta} = \tilde{\phi}_{\mathcal{X}, \lambda \mathcal{S}}^* \left(1 + \frac{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}{\lambda} \right). \quad (43)$$

The thesis follows from (43), (41) and (42). \square

By Lemma 2.10, estimates (13) and (18) as well as Remark 2.11, we only have to prove that $\phi_{\mathcal{R}, \lambda \mathcal{S}}^* \mathbb{E}_{\mu_{\mathcal{R}}} [T_{\lambda \mathcal{S}}]$ goes to one to conclude the proof of Theorem 1. Since

$$\mathbb{E}_{\mu_{\mathcal{R}}} [T_{\lambda \mathcal{S}}] = \mathbb{E}_{\mu_{\mathcal{R}}} [\ell_{\mathcal{R}}(T_{\lambda \mathcal{S}}) + \ell_{\mathcal{S} \setminus \mathcal{R}}(T_{\lambda \mathcal{S}})] \leq \mathbb{E}_{\mu_{\mathcal{R}}} [\ell_{\mathcal{R}}(T_{\lambda \mathcal{S}})] + \mathbb{E}_{\mu_{\mathcal{R}}} [\ell_{\mathcal{S}}(T_{\lambda \mathcal{S}})] = \mathbb{E}_{\mu_{\mathcal{R}}} [T_{\lambda \mathcal{S}}^{\mathcal{R}}] + \frac{1}{\lambda},$$

we get

$$\phi_{\mathcal{R}, \lambda \mathcal{S}}^* \mathbb{E}_{\mu_{\mathcal{R}}} [T_{\lambda \mathcal{S}}] \leq 1 + \frac{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}{\lambda}$$

from estimate (15). This upper bound goes to one, as a consequence of estimate (13).

To get a lower bound, we write

$$\phi_{\mathcal{R}, \lambda_S}^* \mathbb{E}_{\mu_{\mathcal{R}}} [T_{\lambda_S}] = \int_0^{+\infty} \mathbb{P}_{\mu_{\mathcal{R}}} (\phi_{\mathcal{R}, \lambda_S}^* T_{\lambda_S} > t) dt$$

and recall Proposition 2.7: the integrand goes to e^{-t} . Fatou's lemma gives the desired lower bound and this concludes the proof. \square

4 Proof of Theorem 3

We will use the stopping time

$$\tau = T_{\kappa_{\mathcal{R}}} \wedge T_{\lambda_S}$$

to build T^* . The laws of $X(T_{\kappa_{\mathcal{R}}})$ and $X(T_{\lambda_S})$ are indeed close to $\mu_{\mathcal{R}}$ and μ_S . But conditioning on $\{\tau = T_{\kappa_{\mathcal{R}}}\}$ or $\{\tau = T_{\lambda_S}\}$ introduces correlations that can make the law of $X(\tau)$ in general delicate to control. The law of $X(T_{\kappa_{\mathcal{R}}})$ can be written as a convex combination of those of $X(\tau)$ conditioned on $\{\tau = T_{\kappa_{\mathcal{R}}}\}$ and $X(T_{\kappa_{\mathcal{R}}})$ conditioned on $\{T_{\kappa_{\mathcal{R}}} > T_{\lambda_S}\}$: it holds, for all x in \mathcal{X} , and with

$$\alpha(x) = \mathbb{P}_x (T_{\kappa_{\mathcal{R}}} < T_{\lambda_S}),$$

$$\mathbb{P}_x (X(T_{\kappa_{\mathcal{R}}}) = \cdot) = \alpha(x) \mathbb{P}_x (X(\tau) = \cdot \mid \tau = T_{\kappa_{\mathcal{R}}}) + (1 - \alpha(x)) \mathbb{P}_x (X(T_{\kappa_{\mathcal{R}}}) = \cdot \mid T_{\kappa_{\mathcal{R}}} > T_{\lambda_S}).$$

By Proposition 2.8 the first and last of these three distributions are close to $\mu_{\mathcal{R}}$ for a small enough κ . We can then estimate the total variation distance between $\mu_{\mathcal{R}}$ and the law of $X(\tau)$ conditioned on $\{\tau = T_{\kappa_{\mathcal{R}}}\}$:

$$d_{\text{TV}}(\mu_{\mathcal{R}}, \mathbb{P}_x (X(\tau) = \cdot \mid \tau = T_{\kappa_{\mathcal{R}}})) \leq \frac{2 - \alpha(x)}{\alpha(x)} \frac{\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\chi_{\mathcal{R}}}}{2\kappa} \right]_+ \right),$$

so that, if $\alpha(x) \geq 1/2$,

$$d_{\text{TV}}(\mu_{\mathcal{R}}, \mathbb{P}_x (X(\tau) = \cdot \mid \tau = T_{\kappa_{\mathcal{R}}})) \leq \frac{3\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\chi_{\mathcal{R}}}}{2\kappa} \right]_+ \right).$$

This suggests the following construction. We set $\tau^0 = 0$ and, for all $i \geq 0$, we define

$$\tau_{\mathcal{R}}^{i+1} = \tau^i + T_{\kappa_{\mathcal{R}}} \circ \Theta_{\tau^i}, \quad \tau_S^{i+1} = \tau^i + T_{\lambda_S} \circ \Theta_{\tau^i}, \quad \tau^{i+1} = \tau_{\mathcal{R}}^i \wedge \tau_S^i.$$

We call I the first index $i > 0$ such that either

$$\tau^i = \tau_{\mathcal{R}}^i \quad \text{and} \quad \alpha(X(\tau^{i-1})) \geq \frac{1}{2}$$

or

$$\tau^i = \tau_S^i \quad \text{and} \quad \alpha(X(\tau^{i-1})) \leq \frac{1}{2}.$$

We finally define

$$T_{\mathcal{R}}^* = \tau_{\mathcal{R}}^I, \quad T_S^* = \tau_S^I, \quad T^* = T_{\mathcal{R}}^* \wedge T_S^*.$$

The random variable I is stochastically dominated by a geometric random variable of mean 2, each random variable $(\tau^i - \tau^{i-1})$ is stochastically dominated by an exponential random variable of rate $\kappa \wedge \lambda$ and it holds

$$T^* = \sum_{i>0} (\tau^i - \tau^{i-1}) \mathbf{1}_{\{I \geq i\}},$$

so that, observing that for $(\tau^i - \tau^{i-1})$ and $\{I \geq i\}$ are independent for all $i > 0$,

$$\mathbb{E}_x [T^*] \leq \frac{2}{\kappa \wedge \lambda},$$

for all x in \mathcal{X} . By construction it holds

$$d_{\text{TV}}(\mu_{\mathcal{R}}, \mathbb{P}_x (X(\tau) = \cdot \mid T^* = T_{\mathcal{R}}^*)) \leq \frac{3\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\chi_{\mathcal{R}}}}{2\kappa} \right]_+ \right).$$

With

$$\eta = \frac{3\kappa}{\gamma_{\mathcal{R}}} \left(1 + \left[\ln \frac{\gamma_{\mathcal{R}} \sqrt{\lambda \mathcal{R}}}{2\kappa} \right]_+ \right) + \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}{1 - \epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}},$$

the triangular inequality gives

$$d_{TV}(\mu_{\mathcal{R}, \lambda \mathcal{S}}^*, \mathbb{P}_x(X(T^*) = \cdot \mid T^* = T_{\mathcal{R}}^*)) \leq \eta.$$

Using Proposition 2.7, for all $\lambda \geq \phi_{\mathcal{R} \setminus \mathcal{S}}^*$, or any $\lambda \geq \phi_{\mathcal{R}, \lambda \mathcal{S}}^*$, we get

$$\mathbb{P}_x \left(T_{\lambda \mathcal{S}} \circ \Theta_{T^*} > \frac{t}{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*} \mid T^* = T_{\mathcal{R}}^* \right) \geq e^{-t} \{1 - \eta e^t\}$$

and

$$\mathbb{P}_x \left(T_{\lambda \mathcal{S}} \circ \Theta_{T^*} > \frac{t}{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*} \mid T^* = T_{\mathcal{R}}^* \right) \leq e^{-t} \left\{ \exp \left(\sqrt{\frac{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}{\lambda}} \right) + \left[\eta + \exp \left(-\sqrt{\frac{\lambda}{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}} \right) \right] e^t \right\}.$$

Finally, if

$$\delta = \mathbb{P}_{\nu}(T_{\lambda \mathcal{S}} < T_{\kappa \mathcal{R}}),$$

then, by definition of α ,

$$\frac{1}{2} \mathbb{P}_{\nu} \left(\alpha(X(0)) \leq \frac{1}{2} \right) \leq \mathbb{P}_{\nu} \left(T_{\lambda \mathcal{S}} < T_{\kappa \mathcal{R}} \mid \alpha(X(0)) \leq \frac{1}{2} \right) \mathbb{P}_{\nu} \left(\alpha(X(0)) \leq \frac{1}{2} \right) \leq \delta$$

and

$$\mathbb{P}_{\nu}(T^* = T_{\mathcal{R}}^* = T_{\kappa \mathcal{R}} < T_{\lambda \mathcal{S}}) \geq 1 - \mathbb{P}_{\nu} \left(\alpha(X(0)) \leq \frac{1}{2} \text{ or } T_{\lambda \mathcal{S}} < T_{\kappa \mathcal{R}} \right) \geq 1 - (2\delta + \delta) = 1 - 3\delta.$$

Since the symmetrical statements hold when exchanging the roles of \mathcal{R} and \mathcal{S} as well as those of κ and λ , this concludes the proof. \square

5 Proof of Theorem 2

By Proposition 2.8 and Remark 2.11, the law of $X(T_1) = X(T_{\lambda \mathcal{S}})$ is close to $\mu_{\mathcal{S}}$, as the law of $X(T_2)$ is close to $\mu_{\mathcal{R}}$. Hence, it will be sufficient to find θ such that our time averages on time scale θ are close to the expected values computed with $\mu_{\mathcal{R}}$ before T_1 . Theorem 2 follows then from the quantitative

Proposition 5.1. *If $\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^* < 1$ and $\eta < 1$ is such that*

$$\eta^3 \geq \frac{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*}{\lambda} \quad \text{and} \quad \eta^4 \geq \sqrt{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}, \quad (44)$$

then, setting

$$\theta = \frac{1}{\eta^2} \left(\frac{1}{\lambda} \vee \frac{\sqrt{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}}{\eta \phi_{\mathcal{R}, \lambda \mathcal{S}}^*} \right) \leq \frac{\eta}{\phi_{\mathcal{R}, \lambda \mathcal{S}}^*},$$

it holds

$$\mathbb{P}_{\mu_{\mathcal{R}}} \left(\theta < T_1, \sup_{t < T_1 - \theta} |A_{\theta}(t, f) - \mu_{\mathcal{R}}(f)| \leq 4\eta \|f\|_{\infty} \right) \geq \left\{ 1 - 4\eta - \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}{1 - \epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}} \right\} \quad (45)$$

for all $f : \mathcal{X} \rightarrow \mathbb{R}$.

Proof: We can assume, without loss of generality, that $|f|$ is bounded by 1 and we will consider four events, each of them with a small probability, such that when none of them occurs $T_1 > \theta$ and our time averages are indeed close to $\mu_{\mathcal{R}}(f)$. The first of these events is $A = \{T_{\lambda \mathcal{S}}^{\mathcal{R}} \leq \theta\}$. When its complementary occurs $\{\theta < T_1\}$ is implied. By Proposition 2.6, it holds

$$\mathbb{P}_{\mu_{\mathcal{R}}}(A) \leq \mathbb{P}_{\mu_{\mathcal{R}}}(T_{\lambda \mathcal{S}}^{\mathcal{R}} \leq \theta) \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}{1 - \epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}} + 1 - e^{-\phi_{\mathcal{R}, \lambda \mathcal{S}}^* \theta} \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}{1 - \epsilon_{\mathcal{R}, \lambda \mathcal{S}}^*}} + \phi_{\mathcal{R}, \lambda \mathcal{S}}^* \theta. \quad (46)$$

Our next event is $B = \{\ell_S(T_1) > \theta'\}$ with

$$\theta' = \eta\theta$$

being a small fraction of θ . When its complementary occurs the excursions of X in $\mathcal{R} \setminus \mathcal{S}$ up to T_1 , which cannot be longer than $\ell_S(T_1)$, will have a negligible contribution to time averages on time scale θ . Since $\ell_S(T_1)$ is an exponential variable of rate λ , it holds

$$\mathbb{P}_{\mu_{\mathcal{R}}}(B) = \mathbb{P}_{\mu_{\mathcal{R}}}(\ell_S(T_1) > \theta') = e^{-\lambda\theta'}. \quad (47)$$

To control the supremum appearing in (45), we will divide the local time interval $[0, T_{\lambda_S}^{\mathcal{R}}]$ into intervals of length θ' , or smaller as far as the last one is concerned. For k_1 a parameter that we will fix later, our third event is $C = \{T_{\lambda_S}^{\mathcal{R}} > k_1\theta'\}$. When its complementary occurs it is sufficient to control the time averages on time scale θ' for at most k_1 intervals of this length to estimate time averages on time scale θ . By Proposition 2.6, it holds

$$\mathbb{P}_{\mu_{\mathcal{R}}}(C) = \mathbb{P}_{\mu_{\mathcal{R}}}(T_{\lambda_S}^{\mathcal{R}} > k_1\theta') \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + e^{-\phi_{\mathcal{R},\lambda_S}^* k_1\theta'}. \quad (48)$$

Our last event, which we will call D , is that there is a local time interval $[k\theta', (k+1)\theta']$ with $k < k_1$ for which the time average associated with the trace $X^{\mathcal{R}} = X_{\mathcal{R},0_S}$ of X on \mathcal{R} ,

$$A_{\theta'}^{\mathcal{R}}(k\theta') = \frac{1}{\theta'} \int_{k\theta'}^{(k+1)\theta'} f(X^{\mathcal{R}}(s)) ds,$$

differs from $\mu_{\mathcal{R}}(f)$ of more than η . Any time interval $[t, t+\theta]$ with $t+\theta < T_1$ can be associated “by removal of the excursions outside \mathcal{R} ” with a local time interval $[\ell_{\mathcal{R}}(t), \ell_{\mathcal{R}}(t+\theta)]$, which, in turn, can be divided into at most θ/θ' fully covered and two partially covered local time intervals of the form $[k\theta', (k+1)\theta']$. Hence, if none of our last three events B – D occurs, then, for each $t < T_1 - \theta$ and using $|f| \leq 1$, it holds

$$\left| A_{\theta}(t, f) - \mu_{\mathcal{R}}(f) \right| \leq \frac{\theta' + \frac{\theta}{\theta'}\theta'\eta + 2\theta'}{\theta} = 4\eta.$$

As far as the probability of D is concerned, since $X^{\mathcal{R}}$ starts from equilibrium, the expected value of $A_{\theta'}^{\mathcal{R}}(k\theta')$ is equal to $\mu_{\mathcal{R}}(f)$ and we can make an exact computation of its variance by writing the decomposition f in the eigenfunctions basis associated with the generator of $X^{\mathcal{R}} = X_{\mathcal{R},0_S}$. By Proposition 2.4 its spectral gap, which is associated with a zero value of λ , is larger than $\gamma_{\mathcal{R},\lambda_S}$ and we get the upper bound

$$\mathbb{E}_{\mu_{\mathcal{R}}} \left[\left(A_{\theta'}^{\mathcal{R}}(k\theta') - \mu_{\mathcal{R}}(f) \right)^2 \right] \leq \frac{2\text{Var}_{\mu_{\mathcal{R}}}(f)}{\gamma_{\mathcal{R},\lambda_S}\theta'} \leq \frac{2}{\gamma_{\mathcal{R},\lambda_S}\theta'}.$$

We then get

$$\mathbb{P}_{\mu_{\mathcal{R}}}(D) \leq \frac{2k_1}{\gamma_{\mathcal{R},\lambda_S}\theta'\eta^2}. \quad (49)$$

For (48) and (49) to be useful we need to choose k_1 in such a way that, for $\eta \ll 1$,

$$k_1 \gg \frac{1}{\phi_{\mathcal{R},\lambda_S}^*\theta'} \quad \text{and} \quad k_1 \ll \gamma_{\mathcal{R},\lambda_S}\theta'\eta^2,$$

which is possible if

$$\theta'^2 \gg \frac{1}{\eta^2\gamma_{\mathcal{R},\lambda_S}\phi_{\mathcal{R},\lambda_S}^*}. \quad (50)$$

We will choose θ' and η to ensure this condition and we will set

$$k_1 = \sqrt{\frac{1}{\phi_{\mathcal{R},\lambda_S}^*\theta'}\gamma_{\mathcal{R},\lambda_S}\theta'\eta^2} = \frac{\eta}{\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}. \quad (51)$$

For (46) and (47) to be useful we need to choose θ' in such a way that, for $\eta \ll 1$, (recall that $\theta = \theta'/\eta$)

$$\theta' \gg \frac{1}{\lambda} \quad \text{and} \quad \theta' \ll \frac{\eta}{\phi_{\mathcal{R},\lambda_S}^*}.$$

Once combined with (50), this reads

$$\frac{1}{\lambda} \vee \frac{\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\eta\phi_{\mathcal{R},\lambda_S}^*} \ll \theta' \ll \frac{\eta}{\phi_{\mathcal{R},\lambda_S}^*}. \quad (52)$$

We choose

$$\theta' = \frac{1}{\eta} \left(\frac{1}{\lambda} \vee \frac{\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\eta\phi_{\mathcal{R},\lambda_S}^*} \right)$$

to satisfy the first inequality in (52); Condition (44) will ensure the last one.

Estimates (46)–(49) give then, recalling (51) and using (44),

$$\begin{aligned} & \mathbb{P}_{\mu_{\mathcal{R}}} \left(T_1 \leq \theta \text{ or } \exists t < T_1 - \theta, \left| A_\theta(t, f) - \mu_{\mathcal{R}}(f) \right| > 4\eta \right) \\ & \leq \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + \frac{\phi_{\mathcal{R},\lambda_S}^*}{\eta^2} \left(\frac{1}{\lambda} \vee \frac{\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\eta\phi_{\mathcal{R},\lambda_S}^*} \right) + \exp \left\{ -\frac{\lambda}{\eta} \left(\frac{1}{\lambda} \vee \frac{\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\eta\phi_{\mathcal{R},\lambda_S}^*} \right) \right\} \\ & \quad + \frac{1}{2} \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + \exp \left\{ -\phi_{\mathcal{R},\lambda_S}^* \frac{\eta/\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\eta} \left(\frac{1}{\lambda} \vee \frac{\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\eta\phi_{\mathcal{R},\lambda_S}^*} \right) \right\} + \frac{2\eta/\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\gamma_{\mathcal{R},\lambda_S} \frac{1}{\eta} \left(\frac{1}{\lambda} \vee \frac{\sqrt{\epsilon_{\mathcal{R},\lambda_S}^*}}{\eta\phi_{\mathcal{R},\lambda_S}^*} \right) \eta^2} \\ & \leq \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + \eta + 2e^{-1/\eta} + 2\eta \leq \sqrt{\frac{\epsilon_{\mathcal{R},\lambda_S}^*}{1 - \epsilon_{\mathcal{R},\lambda_S}^*}} + 4\eta. \end{aligned}$$

□

References

- [ACGM] L. Avena, F. Castell, A. Gaudillière, and C. Mélot. Approximate and exact solutions of intertwining equations through random spanning forests. <https://arxiv.org/abs/1702.05992>.
- [ACGM20] L. Avena, F. Castell, A. Gaudillière, and C. Mélot. Intertwining wavelets or multiresolution analysis on graphs through random forests. *Applied and Computational Harmonic Analysis*, 48:949–992, 2020.
- [AGL17] I. Armendáriz, S. Grosskinsky, and M. Loulakis. Metastability in a condensing zero-range process in the thermodynamic limit. *Probab. Theory Relat. Fields*, 169:105–175, 2017.
- [BdH15] A. Bovier and F. den Hollander. *Metastability, A Potential-Theoretic approach*. Springer, Cham, 2015.
- [BEGK02] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and low lying spectra in reversible Markov chains. *Commun. Math. Phys.*, 228:219–255, 2002.
- [BEGK04] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes 1. Sharp estimates for capacities and exit times. *J. Eur. Math. Soc.*, 6:399–424, 2004.
- [BEGK05] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes 2. Precise estimates for smalleigenvalues. *J. Eur. Math. Soc.*, 7:69–99, 2005.
- [BG16] A. Bianchi and A. Gaudillière. Metastable states, quasi-stationary distributions and soft measures. *Stochastic Processes and their Applications*, 126(6):1622–1680, 2016.
- [BL10] J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains. *J. Stat. Phys.*, 140:1065–1114, 2010.
- [BL11] J. Beltrán and C. Landim. Metastability of reversible finite state Markov processes. *Stoch. Proc. and Appl.*, 121:1633–1677, 2011.
- [BL15] J. Beltrán and C. Landim. A Martingale approach to metastability. *Prob. Theory and Related Fields*, 161:267–307, 2015.

- [CGOV84] M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares. Metastable behaviour of stochastic dynamics: a pathwise approach. *Jour. Stat. Phys.*, 35:603–634, 1984.
- [CNS15] E. N. M. Cirillo, F. R. Nardi, and J. Sohler. Metastability for general dynamics with rare transitions: escape time and critical configurations. *Journal of Statistical Physics*, 162:365–403, 2015.
- [FMN⁺16] R. Fernandez, F. Manzo, F. R. Nardi, E. Scoppola, and J. Sohler. Conditioned, quasi-stationary, restricted measures and metastability. *The Annals of Applied Probability*, 26:760–793, 2016.
- [FMNS15] R. Fernandez, F. Manzo, F. R. Nardi, and E. Scoppola. Asymptotically exponential hitting times and metastability: a pathwise approach without reversibility. *Electronic Journal of Probability*, 20:1–37, 2015.
- [FW84] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag Berlin Heidelberg, 1984.
- [GL11] A. Gaudillière and C. Landim. A Dirichlet principle for non reversible Markov chains and some recurrence theorems. *Probability Theory and Related Fields*, 158:55–89, 2011.
- [GMV20] A. Gaudillière, P. Milanesi, and M. E. Vares. Asymptotic exponential law for the transition time to equilibrium of the metastable kinetic Ising model with vanishing magnetic field. *Jour. Stat. Phys.* (*in press*), 2020.
- [Lan14] C. Landim. Metastability for a Non-reversible Dynamics: The Evolution of the Condensate in Totally Asymmetric Zero Range Processes. *Commun. Math. Phys.*, 330:1–32, 2014.
- [LMS19] C. Landim, M. Mariani, and I. Seo. Dirichlet’s and Thomson’s Principles for Non-selfadjoint Elliptic Operators with Application to Non-reversible Metastable Diffusion Processes. *Arch. Rational Mech. Anal.*, 231:887–938, 2019.
- [LP71] J. L. Lebowitz and O. Penrose. Rigorous treatment of metastable states in the van der Waals-Maxwell Theory. *J. Stat. Phys.*, 3:211–241, 1971.
- [Max75] J. C. Maxwell. On the dynamical evidence of the molecular constitution of bodies. *Nature*, 11:357–359, 1875.
- [OV05] E. Olivieri and M. E. Vares. *Large deviations and metastability*. Cambridge University Press, 2005.
- [SS98] R.H. Schonmann and S.B. Shlosman. Wulff droplets and the metastable relaxation time of the kinetic Ising model. *Comm. Math. Phys.*, 194(2):389–462, 1998.