MULTI-SCALE METASTABLE DYNAMICS AND THE ASYMPTOTIC STATIONARY DISTRIBUTION OF PERTURBED MARKOV CHAINS

VOLKER BETZ AND STÉPHANE LE ROUX

ABSTRACT. We assume that the transition matrix of a Markov chain depends on a parameter ε, and converges as $\varepsilon \to 0$. The chain is irreducible for $\varepsilon > 0$ but may have several essential communicating classes when $\varepsilon = 0$. This leads to metastable behavior, possibly on multiple time scales. For each of the relevant time scales, we derive two effective chains. The first one describes the (possibly irreversible) metastable dynamics, while the second one is reversible and describes metastable escape probabilities. Closed probabilistic expressions are given for the asymptotic transition probabilities of these chains. As a consequence, we obtain efficient algorithms for computing the committor function and the limiting stationary distribution.

1. INTRODUCTION

In this paper we give a detailed analysis of the asymptotic dynamics and stationary distribution for a special class of metastable Markov chains. Loosely speaking, a metastable Markov chain is one that, on short time scales, looks like a stationary Markov chain exploring only a small subset of its state space; on longer time scales, however, it performs fast and rare transitions between different such subsets.

The topic of metastability is an old one. Its origins can be traced back at least to the works of Eyring [10] and Kramers [13], who studied it in the context of chemical reaction rates. In the context of perturbed dynamical systems, Freidlin and Wentzell [11] developed a systematic approach based on large deviation theory. This approach was extended by Berglund and Gentz [3] to cover stochastic bifurcation and stochastic resonance, and by Olivieri and Scoppola [20, 21] to study dynamics of Markov chains with exponentially small transition probabilities. Bovier, Eckhoff, Gayrard and Klein [5, 6, 7] developed a systematic approach based on capacities, and gave a precise mathematical definition for metastability. The transition path theory [25, 26] investigates the most probable paths that the Markov chain uses when travelling between different metastable states. Recent books on various aspects of metastability include the monograph [19], and the lecture notes [4]. As we will discuss in Section 4, the chains treated in the present paper are metastable in the sense of Bovier at al. In the case of reversible Markov chains, many of our results can be deduced from the theory of [5, 6, 7], although our proofs are different and do not rely on the variational methods used there. For the non-reversible case, our results are new. Our main new tool in this context is Proposition 2.1, which, while being quite elementary, is surprisingly useful for the situation at hand.

A different type of metastable Markov chains, called nearly decomposable chains, is well known in the computer science community. For these, reversibility is not assumed, but instead the asymptotic essential classes need to be directly accessible from each other, and transient states are not allowed. The seminal paper in this context is by Simon and Ando [22], where they derive the metastable behavior and some information on the asymptotic

stationary measure for such chains. Subsequently, the method was clarified and extended, and Meyer [17] realized that many of the extensions have a common foundation that he called the theory of the stochastic complement. Many further extensions and refinements of the method have been given since. We refer to the recent papers [24, 18, 23] and the references therein. Our contribution in this context is to show that the assumption of almost decomposability is not necessary for extracting the asymptotic information on both the metastable behavior and the stationary measure. We comment more on this at the end of Section 5.

A third, apparently independent, effort to treat metastable Markov chains took place in the context of game theory and mathematical economy. Here the start was made by HP Young [29]. He basically advocated using the Markov chain tree theorem, as given in [1] or [11]. Up to normalization, it gives the stationary measure $\mu(x)$ as the sum of terms $w(t)$ indexed by the directed spanning trees of S rooted in x, where the weight $w(t)$ of a tree t is the product of all transition probabilities along its edges; for details see [1]. As has been pointed out in [9], the problem with this formula is that while it is in principle not numerically unstable, it involves computing all spanning trees, which is exponentially expensive and thus becomes non tractable for large state spaces. Moreover, all of the $w(t)$ are usually tiny, and so we are trying to add an astronomical number of tiny terms, which is not a good idea.

A different approach was taken by Wicks and Greenwald [27, 28] who offer a solution that is closer to the one described in [17], but differs in some important details. At the center of their method is what they call the quotient construction on stochastic matrices, which allows them to recursively simplify the state space and, by keeping track of the various simplifications, to compute $\lim_{\varepsilon \to 0} \mu_\varepsilon$ in the end.

In Section 5, we give a new algorithm for computing the stationary distribution of metastable Markov chains, and comment on the similarities and differences to the algorithms given in [27, 28].

Let us describe our setup and results in some more detail. Consider a family of discrete time Markov chains $X^{(\varepsilon)} = (X_n^{(\varepsilon)})_{n \in \mathbb{N}_0}$ with finite state space S and transition matrices $P_\varepsilon = (p_\varepsilon(x, y))_{x,y \in S}$. We assume that the map $\varepsilon \mapsto p_\varepsilon(x, y)$ is continuous at $\varepsilon = 0$ for all $x, y \in S$, and that the Markov chain $X^{(\varepsilon)}$ is irreducible when $\varepsilon > 0$. For $\varepsilon = 0$ however, the chain may have several essential communicating classes. Such a family of Markov chains is called an irreducible perturbation of $X^{(0)}$, or simply an irreducibly perturbed Markov chain.

The first main result of the paper is a description of the multi-scale metastable behavior of the chain. Let $E_1, \ldots, E_n$ be the essential classes of the chain at parameter $\varepsilon = 0$. We pick $x_i \in E_i$ for all $i \leqslant n$ and define an effective chain $\hat{X}^{(\varepsilon)}$ with state space $\{x_1, \ldots, x_n\}$. We prove that this chain captures the effective dynamics of the original chain on the shortest metastable time scale, in the sense that its escape probabilities and stationary distribution are asymptotically independent of the choice of the representatives $x_1, \ldots, x_n$, and asymptotically equal to those of the original chain. For the stationary distribution, this means that $\lim_{\varepsilon \to 0} \hat{\mu}_\varepsilon(x_i)/\mu_\varepsilon(E_i) = 1$, where $\hat{\mu}_\varepsilon$ and μ_ε are the stationary distributions of the respective chains. A central tool is a natural, *reversible* chain that has the same stationary distribution as $\hat{X}^{(\varepsilon)}$ and is interesting in its own right.

In order to explore longer metastable time scales, we renormalize the effective chain: for $\hat{X}^{(\varepsilon)}$, all transitions between different states will vanish in the limit $\varepsilon \to 0$. By rescaling time under suitable conditions, we obtain a new perturbed Markov chain, where at least one transition probability between distinct states is of order one as $\varepsilon \to 0$. We can now iterate the procedure described above, yielding effective chains on smaller and smaller state

2

spaces and encoding the dynamics of the original chain on longer and longer metastable time scales.

A similar program has been carried out before by Olivieri and Scoppola [20, 21]. The difference to our approach is that [20, 21] relies on (and extends) the theory of Freidlin and Wentzell, while our approach is closer to the potential theoretic methods of Bovier et. al. [4]. This allows us to avoid many of the technical complications found in [20, 21]. Also, Olivieri and Scoppola only consider Markov chains with exponentially small transition probabilities, and study asymptotics on a logarithmic scale. In contrast, our methods allow for much more general families of transition matrices, and our results are asymptotically sharp in the sense that we identify the correct prefactors for all our asymptotic identities. The last fact is particularly useful in practice, since it allows us to devise numerically stable algorithms for computing the asymptotic stationary distribution $\lim_{\varepsilon \to 0} \mu_\varepsilon(x)$ for all states $x \in S$. Alternatively, we can compute the ratio of the stationary distributions for two given states x, y without computing the full stationary distribution, thus potentially decreasing the computational cost considerably. These algorithms are the second main result of our work.

The paper is organized as follows: in Section 2, we collect some results on escape times for irreducible Markov chains that seem hard to find in the literature. In Section 3, we introduce perturbed Markov chains and show how the results from Section 2 can be used to obtain asymptotic expressions of various important quantities. These will be used in Section 4 to describe the multi-scale effective dynamics of the chain. Finally, in Section 5, we present our numerical algorithms and compare them to those present in the computer science and economics literature.

2. Stationary measures, escape probabilities and hitting distributions

Here we collect the main tools that we will use. In this section, X is a general discrete time Markov chain. In contrast to the remainder of the paper, we do not assume the state space S to be finite, but we will assume that X is irreducible and recurrent unless stated otherwise.

All of the results below are more ore less direct consequences of the strong Markov property, and thus likely to be known. However, we did not find direct references for them, so we also give the short proofs.

For a Markov chain X on a state space S, the hitting time of a set $A \subset S$ is denoted by $\tau_A(X) = \inf\{n \geqslant 0 : X_n \in A\}$, and the return time by $\tau_A^+(X) = \inf\{n > 0 : X_n \in A\}$. As usual, we will write τ_x instead of $\tau_{\{x\}}$ for $x \in S$, and similarly for τ_x^+.

Proposition 2.1. *Assume that X is irreducible and positive recurrent, and write μ for the unique stationary distribution. Then for all $x, y \in S$,*

$$(2.1) \qquad \mu(x)\mathbb{P}^x(\tau_y^+ < \tau_x^+) = \mu(y)\mathbb{P}^y(\tau_x^+ < \tau_y^+).$$

Proposition 2.1 looks rather fundamental. Note in particular its formal similarity to the detailed balance equation. We were surprised that we could not find a reference where it is stated explicitly. When the Markov chain X itself is reversible, Proposition 2.1 is a direct consequence of the well established theory of electrical networks, see e.g. [16]. For the non-reversible case, we have not found Proposition 2.1 explicitly anywhere in the literature, but it follows from Corollary 8 of Chapter 2 in the unfinished, but brilliant, monograph by Aldous and Fill [1]. We give a different, short proof for the convenience of the reader, based on the following result:

Lemma 2.2. *Let X be irreducible and positive recurrent. For all states $x, y, z \in S$,*

$$\mathbb{E}^z(\tau_x^+) = \mathbb{E}^z(\min(\tau_x^+, \tau_y^+)) + \mathbb{P}^z(\tau_y^+ < \tau_x^+)\mathbb{E}^y(\tau_x^+).$$

Proof. By our assumptions, $\tau_y^+ < \infty$ almost surely. Thus

$$\mathbb{E}^z(\tau_x^+) = \mathbb{E}^z(\tau_x^+, \tau_x^+ \leqslant \tau_y^+) + \mathbb{E}^z(\tau_y^+, \tau_y^+ < \tau_x^+) + \mathbb{E}^z(\tau_x^+ - \tau_y^+, \tau_y^+ < \tau_x^+),$$

The first two terms of the last line above sum up to $\mathbb{E}^z(\min(\tau_x^+, \tau_y^+))$, and the last term is equal to $\mathbb{P}^z(\tau_y^+ < \tau_x^+)\mathbb{E}^y(\tau_x^+)$ by the strong Markov property. $\qquad\square$

Proof of Proposition 2.1. If $x = y$, the claim boils down to $0 = 0$, so let us assume that $x \neq y$. By using Lemma 2.2 in two different ways we obtain

$$
\begin{aligned}
(2.2) \qquad \mathbb{E}^y(\tau_y^+) &= \mathbb{E}^y(\min(\tau_y^+, \tau_x^+)) + \mathbb{P}^y(\tau_x^+ < \tau_y^+)\mathbb{E}^x(\tau_y^+), \\
\mathbb{E}^y(\tau_x^+) &= \mathbb{E}^y(\min(\tau_x^+, \tau_y^+)) + \mathbb{P}^y(\tau_y^+ < \tau_x^+)\mathbb{E}^y(\tau_x^+).
\end{aligned}
$$

Rearranging gives $\mathbb{E}^y(\min(\tau_x^+, \tau_y^+)) = \mathbb{E}^y(\tau_x^+)\mathbb{P}^y(\tau_x^+ < \tau_y^+)$, and using $\mu(y) = \mathbb{E}^y(\tau_y^+)^{-1}$ leads to

$$(2.3) \qquad \mu(y)\mathbb{P}^y(\tau_x^+ < \tau_y^+) = \frac{1}{\mathbb{E}^x(\tau_y^+) + \mathbb{E}^y(\tau_x^+)},$$

which is essentially Corollary 8 of Chapter 2 of [1]. For our purposes, we note that the right-hand side of (2.3) is invariant under swapping x and y. $\qquad\square$

Remark: In the continuous time setting, the whole proof of Lemma 2.2 and almost all of the proof of Proposition 2.1 goes through unchanged if we define $\tau_x^+ = \inf\{t > 0 : X_t = x, X_s \neq X_0 \text{ for some } 0 \leqslant s \leqslant t\}$. The only difference is that the formula for the stationary measure in that case is given by $\mu(y) = \mathbb{E}^y(\tau_y^+)^{-1}\lambda(y)^{-1}$, where $\lambda(y)$ is the exponential rate with which the process jumps away from y. This gives the formula

$$\mu(x)\lambda(x)\mathbb{P}^x(\tau_y^+ < \tau_x^+) = \mu(y)\lambda(y)\mathbb{P}^y(\tau_x^+ < \tau_y^+),$$

which is a special case of the symmetry result on capacities for non-reversible continuous time Markov chains derived by Gaudilliére and Landim [12], and applied to investigate metastability by Beltrán and Landim [2]. Their proof is much less elementary than the one presented here.

A direct consequence of Proposition 2.1 is

Corollary 2.3. *The stationary distribution μ of X fulfills the set of equations*

$$(2.4) \qquad \frac{1}{\mu(x)} = \sum_{y \in S} \frac{\mathbb{P}^x(\tau_y^+ \leqslant \tau_x^+)}{\mathbb{P}^y(\tau_x^+ \leqslant \tau_y^+)}.$$

To use Corollary 2.3, we need find a way to compute the escape probabilities appearing in (2.4). We will now collect some tools that will help us to do this, asymptotically, in the context of perturbed Markov chains.

Proposition 2.4. *Let X be an irreducible, recurrent Markov chain. For $A \subset S$, $x \in S$ and $y \in A$, we have*

$$(2.5) \qquad \mathbb{P}^x(X_{\tau_A^+} = y) = p(x, y) + \sum_{z \in S \setminus A} \frac{\mathbb{P}^x(\tau_z^+ < \tau_A^+)}{\mathbb{P}^z(\tau_A^+ < \tau_z^+)} p(z, y).$$

When $x \in S \setminus A$, (2.5) simplifies to

$$(2.6) \qquad \mathbb{P}^x(X_{\tau_A^+} = y) = \sum_{z \in S \setminus A} \frac{\mathbb{P}^x(\tau_z < \tau_A^+)}{\mathbb{P}^z(\tau_A^+ < \tau_z^+)} p(z, y).$$

4

Proof. Put $\tau_{z,0}^+ := 0$, and for $k \geqslant 1$ let

$$\tau_{z,k}^+ := \min\{n \in \mathbb{N} : |\{0 < j \leqslant n : X_j = z\}| = k\},$$

be the k-the hitting time of z. With

$$\Omega_{y,k,z} := \{\tau_{z,k}^+ < \tau_A^+, X_{\tau_{z,k}^+ + 1} = y\},$$

we have $\bigcup_{k \geqslant 1} \bigcup_{z \in S \setminus A} \Omega_{y,k,z} = \{1 < \tau_A^+ < \infty, X_{\tau_A^+} = y\}$, and the sets $\Omega_{y,k,z}$ are disjoint. As the chain is irreducible and recurrent, $\mathbb{P}(\tau_A^+ = \infty) = 0$ holds, and thus

$$(2.7) \qquad \mathbb{P}^x(X_{\tau_A^+} = y) = p(x,y) + \sum_{z \in S \setminus A} \sum_{k \geqslant 1} \mathbb{P}^x(\Omega_{y,k,z}).$$

For $k \geqslant 1$, the strong Markov property and induction gives

$$\mathbb{P}^x(\Omega_{y,k,z}) = \mathbb{P}^x(\tau_z^+ < \tau_A^+)\mathbb{P}^z(\Omega_{y,k-1,z}) = \mathbb{P}^x(\tau_z^+ < \tau_A^+)\mathbb{P}^z(\tau_z^+ < \tau_A^+)^{k-1}\mathbb{P}^z(\Omega_{y,0,z}).$$

Summing up the geometric series in k and using $\mathbb{P}^z(\Omega_{y,0,z}) = p(z,y)$ proves (2.5).

To deduce (2.6) from (2.5), note that when $x \notin A$, we have $\mathbb{P}^x(\tau_z^+ < \tau_A^+) = \mathbb{P}^x(\tau_z < \tau_A^+)$. Furthermore, when $z = x$, $\mathbb{P}^x(\tau_x^+ < \tau_A^+) + \mathbb{P}^x(\tau_A^+ < \tau_x^+) = 1 = \mathbb{P}^x(\tau_x < \tau_A^+)$, and (2.6) follows by rearranging. $\qquad \square$

A variant of Proposition 2.4 is well known and is the basis of many algorithms for computing stationary distributions of large Markov chains. It is called the quotient construction by [27, 28], and the stochastic complement by Meyer [17]. While in all those references, it is written in matrix language, we give here the probabilistic formulation, which also has the benefit that we can give a short and transparent proof.

Proposition 2.5. *([17, 28]) Assume that the state space S is finite, $A \subset S$, $x \in S$ and $y \in A$. Then*

$$(2.8) \qquad \mathbb{P}^x(X_{\tau_A^+} = y) = p(x,y) + \sum_{z,w \in A^c} p(x,w)(1 - P|_{A^c})^{-1}(w,z)p(z,y),$$

where $P|_{A^c} = (p(x,y))_{x,y \in A^c}$ is the restriction of the transition matrix P to A^c.

Proof. Clearly, $\mathbb{P}^x(X_{\tau_A^+} = y) = p(x,y) + \sum_{w \in A^c} p(x,w)\mathbb{P}^w(X_{\tau_A} = y)$. Now, standard results [16] state that $h_y(w) := \mathbb{P}^w(X_{\tau_A} = y)$ is the unique harmonic extension of the function $1_{\{y\}}$ from A to S. In other words, h_y is the unique function so that $Ph_y(w) = h_y(w)$ for all $w \in A^c$, and $h_y(w) = 1_{\{y\}}(w)$ on A. This can be rewritten as $(P|_{A^c} - 1)h_y(w) = -P1_{\{y\}}(w) = -p(w,y)$ for all $w \in A^c$. Since X is irreducible, there exists $n \in \mathbb{N}$ with $\|(P|_{A^c})^n\| < 1$, where $\|.\|$ is the operator norm of a matrix. Thus $(1 - P|_{A^c})$ is invertible. The claim follows. $\qquad \square$

For the following result, we do not assume irreducibility of the chain.

Lemma 2.6. *Let (X_n) be an arbitrary Markov chain, $A, B \subset S$. Assume $x \notin A \cup B$ and $\mathbb{P}^x(\tau_B^+ < \infty) > 0$. Then*

$$\mathbb{P}^x(\tau_B^+ < \tau_A^+) = \frac{\mathbb{P}^x(\tau_B^+ < \tau_{A \cup \{x\}}^+)}{\mathbb{P}^x(\tau_B^+ < \tau_x^+)}.$$

Proof. The strong Markov property gives

$$\mathbb{P}^x(\tau_B^+ < \tau_A^+) = \mathbb{P}^x(\tau_B^+ < \tau_{A \cup \{x\}}^+) + \mathbb{P}^x(\tau_x^+ < \tau_B^+)\mathbb{P}^x(\tau_B < \tau_A).$$

5

As $x \notin A \cup B$, we have $\mathbb{P}^x(\tau_B < \tau_A) = \mathbb{P}^x(\tau_B^+ < \tau_A^+)$. Since we assumed $\mathbb{P}^x(\tau_B^+ < \infty) > 0$, we must have $1 - \mathbb{P}^x(\tau_x^+ < \tau_B^+) = \mathbb{P}^x(\tau_B^+ < \tau_x^+) > 0$. Thus we can rearrange and obtain the result. $\qquad\square$

For our next statement, fix a proper subset $C \subsetneq S$, and define for all $x, y \in S$

$$(2.9) \qquad \tilde{p}(x,y) := p(x,y) \text{ if } x \notin C, \qquad \tilde{p}(x,y) := \mathbb{P}^x(X_{\tau_{C^c}} = y) \text{ if } x \in C.$$

Proposition 2.7. *Let X be an irreducible, recurrent Markov chain. Then $\tilde{P} = (\tilde{p}(x,y))_{x,y \in S}$ is the transition matrix of a Markov chain $\tilde{X}$. Denoting its path measure by $\tilde{\mathbb{P}}$, we have*

$$(2.10) \qquad \tilde{\mathbb{P}}^x(\tau_B < \tau_A) = \mathbb{P}^x(\tau_B < \tau_A).$$

for all $A, B \subset S$ with $(A \cup B) \cap C = \emptyset$, and all $x \in S$.

Proof. Since (X_n) is irreducible and recurrent and $C^c \neq \emptyset$, $\mathbb{P}^x(\tau_{C^c} < \infty) = 1$ for all $x \in C$. Thus it is obvious that $\tilde{P}$ is a stochastic matrix. The statement (2.10) is also intuitively obvious, since all we do is replace the motion inside C with the effective motion from C to its exterior. We nevertheless give the short formal proof.

We write σ_m for the m-th time that the chain (X_n) travels between two states that are not both in C, i.e.

$$\sigma_0 := 0, \quad \sigma_m := \min\{n > \sigma_{m-1} : X_n \notin C \text{ or } X_{n-1} \notin C\}.$$

On $\Omega_0 = \{\sigma_m < \infty \ \forall m \in \mathbb{N}\}$, we define $\tilde{X}_m = X_{\sigma_m}$. Then $\mathbb{P}^x(\Omega_0) = 1$ for all $x \in S$ by recurrence and irreducibility of X, and $\tilde{X}$ is a Markov chain by the strong Markov property of X. Since $\mathbb{P}^x(\tilde{X}_1 = y) = \mathbb{P}^x(X_{\sigma_1} = y) = \tilde{p}(x,y)$, the transition probabilities of $\tilde{X}$ are given by (2.9). Since C is disjoint from A and B, we have

$$\{\tau_A(X) < \tau_B(X)\} \cap \Omega_0 = \{\tau_A(\tilde{X}) = \tau_B(\tilde{X})\} \cap \Omega_0,$$

and (2.10) follows by taking expectations. $\qquad\square$

For our final general statement, we introduce the notion of a direct path which will be useful in several places below. Let J, A and B be subsets of S. A tuple $\gamma = (x_1, \ldots, x_n) \in S^n$ is called a *direct J-path of length n from A to B* if $x_1 \in A$, $x_n \in B$, and for all $1 \leqslant i < j \leqslant n$, if $x_i = x_j$ then $i = 1$ and $j = n$. Note that we allow $x_1, x_n \notin J$. The set of all direct J-paths from A to B will be denoted by $\Gamma_J(A, B)$, and the components of $\gamma \in \Gamma_J(A, B)$ will be written γ_i, $i = 1, \ldots, n$. $|\gamma|$ will denote the length of γ. For $A = \{x\}$ or $B = \{y\}$ we will use the notations $\Gamma(A, y)$ instead of $\Gamma(A, \{y\})$ etc, and speak of direct J-paths from A to y, from x to y or from x to B. The probability of a direct J-path is defined by $\mathbb{P}(\gamma) := \prod_{j=1}^{|\gamma|-1} p(\gamma_j, \gamma_{j+1})$.

Proposition 2.8. *Let J be a finite subset of S. Then for all $x \in J$ and $y \in S \setminus J$,*

$$(2.11) \qquad \mathbb{P}^x(X_{\tau_{S \setminus J}} = y) = \sum_{\gamma \in \Gamma_J(x,y)} \prod_{i=1}^{|\gamma|-1} \frac{p(\gamma_i, \gamma_{i+1})}{1 - \mathbb{P}^{\gamma_i}(X_{\tau_{(S \setminus J) \cup \{\gamma_1, \ldots, \gamma_i\}}^+} = \gamma_i)}$$

Proof. The idea of the proof is to start at state x and run the Markov chain until it either hits $S \setminus J$ or returns to x. In the first case we have reduced the problem to computing $\mathbb{P}^z(X_{\tau_{(S \setminus J) \cup \{x\}}} = y)$ and we iterate the argument for the smaller set $J \setminus \{x\}$; in the second case we use the strong Markov property to restart the process.

We proceed by induction on $|J|$. The claim trivially holds for $J = \emptyset$; so now let $x \in J$. The strong Markov property then gives

$$\mathbb{P}^x(X_{\tau_{S \setminus J}} = y) = \mathbb{P}^x(X_{\tau_{(S \setminus J) \cup \{x\}}^+} = y) + \mathbb{P}^x(X_{\tau_{(S \setminus J) \cup \{x\}}^+} = x)\mathbb{P}^x(X_{\tau_{S \setminus J}} = y)$$

6

By recurrence and irreducibility, we have $\mathbb{P}^x(X_{\tau^+_{(S\setminus J)\cup\{x\}}} = x) < 1$, and we rearrange to

$$\mathbb{P}^x(X_{\tau_{S\setminus J}} = y) = \frac{\mathbb{P}^x(X_{\tau^+_{(S\setminus J)\cup\{x\}}} = y)}{1 - \mathbb{P}^x(X_{\tau^+_{(S\setminus J)\cup\{x\}}} = x)}$$

where the numerator may be decomposed as $p(x,y) + \sum_{z\in J\setminus\{x\}} p(x,z)\mathbb{P}^z(X_{\tau_{(S\setminus J)\cup\{x\}}} = y)$.

Finally, we use the induction hypothesis for the set $J\setminus\{x\}$ to rewrite $\mathbb{P}^z(X_{\tau_{(S\setminus J)\cup\{x\}}} = y)$ for all $z \in J \setminus \{x\}$, and obtain

$$\mathbb{P}^x(X_{\tau_{S\setminus J}} = y) = \frac{p(x,y)}{1 - \mathbb{P}^x(X_{\tau^+_{(S\setminus J)\cup\{x\}}} = x)}$$

$$+ \sum_{z\in J\setminus\{x\}} \sum_{\gamma\in\Gamma_{J\setminus\{x\}}(z,y)} \frac{p(x,z)}{1 - \mathbb{P}^x(X_{\tau^+_{(S\setminus J)\cup\{x\}}} = x)} \prod_{i=1}^{|\gamma|-1} \frac{p(\gamma_i,\gamma_{i+1})}{1 - \mathbb{P}^{\gamma_i}(X_{\tau^+_{(S\setminus J)\cup\{x,\gamma_1,\ldots,\gamma_i\}}} = \gamma_i)}$$

Re-indexing yields the claim. $\qquad\square$

3. Perturbed Markov chains: escape probabilities

Let $X^{(0)} = (X_n^{(0)})_{n\in\mathbb{N}}$ be a Markov chain on a finite state space S. A family $X^{(\varepsilon)} = (X_n^{(\varepsilon)})_{n\in\mathbb{N}}$ of Markov chains on S indexed by $\varepsilon \geqslant 0$ is called a *perturbation* of $X^{(0)}$ if $\lim_{\varepsilon\to 0} p_\varepsilon(x,y) = p_0(x,y)$ for all $x,y \in S$, where $p_\varepsilon(x,y)$ denotes the elements of the transition matrix P_ε of the chain $X^{(\varepsilon)}$, $\varepsilon \geqslant 0$. We will speak of an *irreducible perturbation* of $X^{(0)}$ (or, alternatively, call the family $X^{(\varepsilon)}$ an *irreducibly perturbed Markov chain*) if the chain $X^{(\varepsilon)}$ is irreducible for all $\varepsilon > 0$.

Note that in the definition of irreducibly perturbed Markov chains, we do not require that $X^{(0)}$ be irreducible, and indeed the case where $X^{(0)}$ has several ergodic components is the interesting one. Recall that $x \in S$ is called *accessible* from $y \in S$ under $X^{(\varepsilon)}$ if $\mathbb{P}_\varepsilon^y(X_n = x) > 0$ for some $n \geqslant 0$. We write $x \to y$ if y is accessible from x, and say that two states x and y *communicate* if $x \to y$ and $y \to x$. The property to communicate forms an equivalence relation, and the respective equivalence classes are called *communicating classes*. A state x is called *essential* if $y \to x$ for all $y \in S$ such that $x \to y$, otherwise *transient*. It is easy to see that either all members of a communicating class E are essential, or all are transient. In the first case, E is called an essential (communicating) class, or ergodic component.

S can thus be decomposed into finitely many disjoint essential classes $E_1, \ldots E_n$ and the set $F = S \setminus \bigcup_{i=1}^n E_i$ of transient states. To emphasize that a nontrivial ergodic decomposition only exists for $\varepsilon = 0$, we will always speak of P_0-essential classes and P_0-transient states. $\mathcal{E}$ will denote the set of all P_0-essential classes.

The sets E_i and F can be conveniently described in terms direct paths. The following statement could be taken as a definition of P_0-essential classes and P_0-transient states; the proof of equivalence to the traditional definition of essential classes (see e.g. [16]) is very easy, and omitted here. Here and below, we will say that a direct path γ is P_0-*relevant* if $\mathbb{P}_0(\gamma) = \lim_{\varepsilon\to 0} \mathbb{P}_\varepsilon(\gamma) > 0$, otherwise P_0-*irrelevant*.

Lemma 3.1. *Let $X^{(\varepsilon)}$ be an irreducibly perturbed Markov chain.*
a) $x,y \in S$ are in the same P_0-essential class E if and only if there exists a P_0-relevant direct E-path from x to y, and a P_0-relevant direct E-path from y to x.
b) $x \in S$ is in the transient set F if and only if all direct S-paths from $\bigcup_{j=1}^n E_j$ to x are P_0-irrelevant.

In much of what follows, we will use the following concept of asymptotic equivalence. Two functions $\varepsilon \mapsto a_\varepsilon$ and $\varepsilon \mapsto b_\varepsilon$ from $\mathbb{R}_0^+$ to $\mathbb{R}_0^+$ are *asymptotically equivalent*, if either a_ε and b_ε are identically zero, or $b_\varepsilon > 0$ for all $\varepsilon \in (0, \varepsilon_0)$ with some $\varepsilon_0 > 0$ and $\lim_{\varepsilon \to 0} a_\varepsilon / b_\varepsilon = 1$. Note that in the latter case, we do not assume convergence of a_ε or b_ε. We write $a_\varepsilon \simeq b_\varepsilon$ if a_ε is asymptotically equivalent to b_ε. It is easy to see that $\simeq$ is indeed an equivalence relation, and in particular this implies $1/a_\varepsilon \simeq 1/b_\varepsilon$ whenever $a_\varepsilon \simeq b_\varepsilon$ and a_ε is not identically zero. We will also need to know that $\simeq$ is stable under addition and multiplication in the following sense: if $a_\varepsilon \simeq b_\varepsilon$ and $c_\varepsilon \simeq d_\varepsilon$, then $a_\varepsilon + c_\varepsilon \simeq b_\varepsilon + d_\varepsilon$, and $a_\varepsilon c_\varepsilon \simeq b_\varepsilon d_\varepsilon$. Stability under multiplication is trivial, and stability under addition follows from

$$|\tfrac{a_\varepsilon + c_\varepsilon}{b_\varepsilon + c_\varepsilon} - 1| = |\tfrac{a_\varepsilon/b_\varepsilon - 1}{1 + c_\varepsilon/b_\varepsilon}| \leqslant |\tfrac{a_\varepsilon}{b_\varepsilon} - 1|$$

and transitivity of $\simeq$. Note that we did not assume that $a_\varepsilon \simeq c_\varepsilon$ in either case.

Let $E \in \mathcal{E}$ be a P_0-essential class. The restriction of $X^{(0)}$ to E is the Markov chain with state space E and transition matrix $(p_0(x, y))_{x, y \in E}$. It is irreducible, and thus has a unique strictly positive stationary distribution ν_E. The trivial extension of ν_E to S (by putting $\nu_E(x) := 0$ for $x \notin E$) will be denoted by the same symbol, and is an extremal point of the convex set of stationary distributions for P_0. The following lemma shows that when we focus our attention on a single P_0-essential class, the unperturbed chain gives a faithful asymptotic description of both the dynamics and the stationary distribution. Here and below we will write μ_ε for the unique stationary distribution of $X^{(\varepsilon)}$, when $0 < \varepsilon$.

Lemma 3.2. *Let $E \in \mathcal{E}$ be a P_0-essential class. Then for all $x, y \in E$ and all $z \in S$,*

$$(3.1) \qquad \lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_z^+) = \mathbb{P}_0^x(\tau_y^+ < \tau_z^+), \quad \text{and} \quad \lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^x(\tau_y < \tau_z) = \mathbb{P}_0^x(\tau_y < \tau_z),$$

and

$$(3.2) \qquad \qquad \lim_{\varepsilon \to 0} \frac{\mu_\varepsilon(x)}{\mu_\varepsilon(y)} = \frac{\nu_E(x)}{\nu_E(y)}.$$

In particular, $\mu_\varepsilon(x)/\mu_\varepsilon(y) \simeq \nu_E(x)/\nu_E(y)$.

Proof. We only prove the first equality from (3.1), the proof for the second one is identical. We decompose

$$(3.3) \qquad \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_z^+) = \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{\{z\} \cup E^c}^+) + \sum_{w \in E^c} \mathbb{P}_\varepsilon^x(X_{\tau_{E^c \cup \{y\}}^+} = w) \mathbb{P}_\varepsilon^w(\tau_y^+ < \tau_z^+).$$

By Proposition 2.4, for $w \in E^c$ we have

$$\mathbb{P}_\varepsilon^x(X_{\tau_{E^c \cup \{y\}}^+} = w) = p_\varepsilon(x, w) + \sum_{u \in E \setminus \{y\}} \frac{\mathbb{P}_\varepsilon^x(\tau_u^+ < \tau_{E^c \cup \{y\}})}{\mathbb{P}_\varepsilon^u(\tau_{E^c \cup \{y\}}^+ < \tau_u^+)} p_\varepsilon(u, w),$$

and for each $u \in E$, there is a P_0-relevant direct E-path γ from u to y. Thus we have $\liminf_{\varepsilon \to 0} \mathbb{P}_\varepsilon^u(\tau_{E^c \cup \{y\}}^+ < \tau_u^+) > 0$, therefore $\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^x(X_{\tau_{E^c \cup \{y\}}^+} = w) = 0$, and the second term on the right-hand side of (3.3) vanishes as $\varepsilon \to 0$.

For the first term of (3.3), x and y being in the same P_0-essential class implies that for each small enough $\varepsilon_0 > 0$ and each $\delta > 0$, we can find $n \in \mathbb{N}$ such that for all $\varepsilon < \varepsilon_0$

$$\mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{\{z\} \cup E^c}^+, \tau_y^+ < n) \leqslant \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{\{z\} \cup E^c}^+) \leqslant \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{\{z\} \cup E^c}^+, \tau_y^+ < n) + \delta.$$

Since the elements of the transition matrix converge, we have

$$\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{\{z\} \cup E^c}, \tau_y^+ < n) = \mathbb{P}_0^x(\tau_y^+ < \tau_{\{z\} \cup E^c}, \tau_y^+ < n) = \mathbb{P}_0^x(\tau_y^+ < \tau_z^+, \tau_y^+ < n).$$

As δ was arbitrary, (3.1) follows. For (3.2), we apply (2.1). $\qquad \square$

As an immediate corollary, we obtain some information on the structure of the stationary distribution in the limit $\varepsilon \to 0$. Recall that $\mathcal{E} = \{E_1, \ldots, E_n\}$ is the collection of P_0-essential classes, and F is the set of transient states.

Corollary 3.3. *Let $z \in S$.*
a) If $z \in F$, then $\lim_{\varepsilon \to 0} \mu_\varepsilon(z) = 0$.
b) If $z \in E$ for some $E \in \mathcal{E}$, then $\mu_\varepsilon(z) \simeq \mu_\varepsilon(E)\nu_E(z)$.
c) In particular if $\lim_{\varepsilon \to 0} \mu_\varepsilon(E)$ exists for all $E \in \mathcal{E}$, then $\lim_{\varepsilon \to 0} \mu_\varepsilon(x)$ exists for all $x \in S$, and

$$\lim_{\varepsilon \to 0} \mu_\varepsilon(x) = \sum_{E \in \mathcal{E}} \lim_{\varepsilon \to 0} \mu_\varepsilon(E)\nu_E(x).$$

The practical usefulness of Corollary 3.3 depends on our ability to compute asymptotic expressions for the $\mu_\varepsilon(E)$. We now give two statements that help us achieve this. The first says that hitting probabilities are asymptotically equivalent when the transition matrices are. The second describes how a perturbed Markov chain leaves a P_0-essential class, with or without the additional condition that it cannot return to its starting point.

Theorem 3.4. *Let $X^{(\varepsilon)}$ and $\tilde{X}^{(\varepsilon)}$ be perturbed Markov chains with finite state space S, but not necessarily irreducible. Let us assume that $p_\varepsilon(x,y) \simeq \tilde{p}_\varepsilon(x,y)$ for the elements of the respective transition matrices. Then for all $A, B \subset S$ and all $x \in S$, we have*

$$\mathbb{P}_\varepsilon^x(\tau_B < \tau_A) \simeq \tilde{\mathbb{P}}_\varepsilon^x(\tau_B < \tau_A).$$

Proof. We will first show that the statement holds in the case where P_ε and $\tilde{P}_\varepsilon$ only differ in one row, i.e. where

(3.4) $p_\varepsilon(z,y) \simeq \tilde{p}_\varepsilon(z,y)$ for some $z \in S$, and $p_\varepsilon(x,y) = \tilde{p}_\varepsilon(x,y)$ for all other $x \in S$.

The claim is then proved by induction. For the case where (3.4) holds, first note (e.g. using a coupling argument) that for all $x \in S$,

$$\mathbb{P}_\varepsilon^x(\tau_B < \tau_A, \tau_B \leqslant \tau_z) = \tilde{\mathbb{P}}_\varepsilon^x(\tau_B < \tau_A, \tau_B \leqslant \tau_z).$$

Thus,

$$\mathbb{P}_\varepsilon^x(\tau_B < \tau_A) = \tilde{\mathbb{P}}_\varepsilon^x(\tau_B < \tau_A, \tau_B \leqslant \tau_z) + \mathbb{P}_\varepsilon^x(\tau_B < \tau_A, \tau_z < \tau_B).$$

Since $\{\tau_z < \tau_B, \tau_z = \infty\} = \emptyset$, we can now use the strong Markov property to find

$$\mathbb{P}_\varepsilon^x(\tau_B < \tau_A, \tau_z < \tau_B) = \mathbb{P}_\varepsilon^x(\tau_z < \tau_B)\mathbb{P}_\varepsilon^z(\tau_B < \tau_A).$$

Again $\mathbb{P}_\varepsilon^x(\tau_z < \tau_B) = \tilde{\mathbb{P}}_\varepsilon^x(\tau_z < \tau_B)$, and it remains to show that $\mathbb{P}_\varepsilon^z(\tau_B < \tau_A) \simeq \tilde{\mathbb{P}}_\varepsilon^z(\tau_B < \tau_A)$. If $z \in A \cup B$, this is trivial. For $z \notin A \cup B$, $\mathbb{P}_\varepsilon^z(\tau_B < \tau_A) = \mathbb{P}_\varepsilon^z(\tau_B^+ < \tau_A^+)$, and $\tilde{\mathbb{P}}_\varepsilon^z(\tau_B < \tau_A) = \tilde{\mathbb{P}}_\varepsilon^z(\tau_B^+ < \tau_A^+)$. We are aiming to use Lemma 2.6, and thus need to deal with the possibility that $\mathbb{P}_\varepsilon^z(\tau_B^+ < \infty) = 0$.

We assumed $p(x,y) \simeq \tilde{p}(x,y)$ for all $x, y \in S$, and so we also have $\mathbb{P}_\varepsilon^x(\gamma) \simeq \tilde{\mathbb{P}}_\varepsilon^x(\gamma)$ for each direct path from x to B. By the definition of $\simeq$, a direct path γ from x to B fulfills $\mathbb{P}_\varepsilon^x(\gamma) > 0$ for all $\varepsilon > 0$ in a neighborhood of $\varepsilon = 0$ if and only if $\tilde{\mathbb{P}}_\varepsilon^x(\gamma) > 0$ in a neighborhood of 0. Let us first assume that no such direct path exists. Then $\mathbb{P}_\varepsilon^z(X_n^{(\varepsilon)} \in B) = \tilde{\mathbb{P}}_\varepsilon^z(X_n^{(\varepsilon)} \in B) = 0$ for all $n \in \mathbb{N}$, and thus $\mathbb{P}^z(\tau_B < \tau_A) = \tilde{\mathbb{P}}_\varepsilon^z(\tau_B < \tau_A) = 0$. Now let us assume that such direct paths do exist. Since $\mathbb{P}_\varepsilon^x(\tau_B < \infty) \geqslant \mathbb{P}_\varepsilon^x(\gamma)$, we can use Lemma 2.6 to get

$$\mathbb{P}_\varepsilon^z(\tau_B^+ < \tau_A^+) = \frac{\mathbb{P}_\varepsilon^z(\tau_B^+ < \tau_{A \cup \{z\}}^+)}{\mathbb{P}_\varepsilon^z(\tau_B^+ < \tau_z^+)}.$$

9

Now,

$$\mathbb{P}_\varepsilon^z(\tau_B^+ < \tau_{A\cup\{z\}}^+) = \sum_{w\in S} p_\varepsilon(z,w)\mathbb{P}_\varepsilon^w(\tau_B < \tau_{A\cup\{z\}})$$

$$\simeq \sum_{w\in S} \tilde{p}_\varepsilon(z,w)\mathbb{P}_\varepsilon^w(\tau_B < \tau_{A\cup\{z\}}) = \sum_{w\in S} \tilde{p}_\varepsilon(z,w)\tilde{\mathbb{P}}_\varepsilon^w(\tau_B < \tau_{A\cup\{z\}}) = \tilde{\mathbb{P}}_\varepsilon^z(\tau_B^+ < \tau_{A\cup\{z\}}^+),$$

and the same argument shows $\mathbb{P}_\varepsilon^z(\tau_B^+ < \tau_z^+) \simeq \tilde{\mathbb{P}}_\varepsilon^z(\tau_B^+ < \tau_z^+)$. The claim follows. $\qquad\square$

The statement of Theorem 3.4 is less obvious than it might appear. The reason is that even though in each step that the chain takes from x on its way to B, the probabilities for the chains X and $\tilde{X}$ differ only by a factor that becomes negligibly close to one as $\varepsilon \to 0$, in the same limit the number of steps needed to reach B can diverge. So one could fear that the errors committed by changing each transition probability to an asymptotically equivalent one will pile up. Theorem 3.4 shows that this is not the case.

Theorem 3.5. *Let E be a P_0-essential class, $x \in E$ and $z \notin E$. Then*

(3.5) $$\mathbb{P}_\varepsilon^x(\tau_{E^c}^+ < \tau_x^+, X_{\tau_{E^c}} = z) \simeq \frac{1}{\nu_E(x)} \sum_{y\in E} \nu_E(y)p_\varepsilon(y,z),$$

and

(3.6) $$\mathbb{P}_\varepsilon^x(X_{\tau_{E^c}} = z) \simeq \frac{1}{Z_\varepsilon(E)} \sum_{y\in E} \nu_E(y)p_\varepsilon(y,z),$$

with normalizing constant

$$Z_\varepsilon(E) = \sum_{\tilde{z}\in E^c} \sum_{\tilde{y}\in E} \nu_E(\tilde{y})p_\varepsilon(\tilde{y},\tilde{z}).$$

Remark: (3.6) is intuitively clear: for small ε, the Markov chain spends a long time in E before exiting, and thus is essentially stationary when it does. Formula (3.5) is less obvious, since a return to x happens in a time of order one.

Proof of Theorem 3.5. To prove (3.5), choose $A = E^c \cup \{x\}$ in Proposition 2.4. Then

$$\mathbb{P}_\varepsilon^x(\tau_{E^c}^+ < \tau_x^+, X_{\tau_{E^c}} = z) = \mathbb{P}_\varepsilon^x(X_{\tau_A^+} = z) = p_\varepsilon(x,z) + \sum_{y\in E\setminus\{x\}} \frac{\mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{E^c\cup\{x\}}^+)}{\mathbb{P}_\varepsilon^y(\tau_{E^c\cup\{x\}}^+ < \tau_y^+)} p_\varepsilon(y,z).$$

We decompose

(3.7) $$\mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{E^c\cup\{x\}}^+) = \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{E^c\cup\{x\}}^+, \tau_{E^c}^+ > \tau_x^+) + \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{E^c\cup\{x\}}^+, \tau_{E^c}^+ < \tau_x^+).$$

An application of Lemma 3.2 yields

$$\lim_{\varepsilon\to 0} \mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{E^c\cup\{x\}}^+) = \mathbb{P}_0^x(\tau_y^+ < \tau_x^+)$$

and, for $y \in E$,

$$\lim_{\varepsilon\to 0} \mathbb{P}_\varepsilon^y(\tau_{E^c\cup\{x\}}^+ < \tau_y^+) = \lim_{\varepsilon\to 0} \mathbb{P}_\varepsilon^y(\tau_x^+ < \tau_y^+) = \mathbb{P}_0^y(\tau_x^+ < \tau_y^+).$$

By Proposition 2.1, we conclude

$$\lim_{\varepsilon\to 0} \frac{\mathbb{P}_\varepsilon^x(\tau_y^+ < \tau_{E^c\cup\{x\}}^+)}{\mathbb{P}_\varepsilon^y(\tau_{E^c\cup\{x\}}^+ < \tau_y^+)} = \frac{\mathbb{P}_0^x(\tau_y^+ < \tau_x^+)}{\mathbb{P}_0^y(\tau_x^+ < \tau_y^+)} = \frac{\nu_E(y)}{\nu_E(x)},$$

and (3.5) is shown.

To see (3.6), we use Proposition 2.4 with $A = E^c$. Since now $x \notin A$, we can use (2.6) and obtain

$$(3.8) \qquad \mathbb{P}_\varepsilon^x(X_{\tau_{E^c}^+} = z) = \sum_{y \in E} \frac{\mathbb{P}^x(\tau_y < \tau_{E^c}^+)}{\mathbb{P}_\varepsilon^y(\tau_{E^c}^+ < \tau_y^+)} p_\varepsilon(y, z).$$

As before, $\mathbb{P}^x(\tau_y < \tau_{E^c}^+) \simeq 1$ for all $y \in E$. By summing (3.5) over all $\tilde{z} \notin E$, we get

$$\mathbb{P}_\varepsilon^y(\tau_{E^c}^+ < \tau_y^+) \simeq \frac{1}{\nu_E(y)} \sum_{\tilde{y} \in E, \tilde{z} \notin E} \nu_E(\tilde{y}) p_\varepsilon(\tilde{y}, \tilde{z}).$$

Plugging these into (3.8), we obtain (3.6). $\qquad\qquad\qquad\qquad\qquad\qquad \square$

4. Perturbed Markov chains: metastable dynamics

Here we describe the metastable dynamics of a perturbed Markov chain. As in the previous section, we will restrict our attention to a finite state space S throughout.

We follow the theory of Bovier et al [5, 6, 7]. In the case of perturbed Markov chains on a finite state space, Definition 2.1 from [5] (see also [4]) is as follows: a set $M \subset S$ is called a *set of metastable points* if for all $x \in M$ and $y \notin M$,

$$(4.1) \qquad \lim_{\varepsilon \to 0} \frac{\mathbb{P}_\varepsilon(\tau_{M \setminus \{x\}}^+ < \tau_x^+)}{\mathbb{P}_\varepsilon^y(\tau_M^+ < \tau_y^+)} = 0.$$

In words, this means that reaching M from the outside of M is much easier than traveling between different points of M, in both cases with the restriction not to return to one's starting point first.

Using Lemma 3.1 and Lemma 3.2, it is easy to see that if we choose precisely one point from each of the P_0-essential classes $E_1, \ldots, E_n$, then the set $S_0 = \{x_1, \ldots, x_n\}$ is a set of metastable points. Also, S_0 is maximal in the sense that adding a further point to S_0 will result in a set no longer fulfilling (4.1). On the other hand, removing points from S_0 or replacing them with points from F may in certain cases still result in a metastable set, depending on the structure of the Markov chain and the points in question. We will not pursue this further since S_0 is the most natural choice. Of course, when some of the E_i contain more than one point, the choice of S_0 is not unique. One of our main results is that when defining the effective chain by the transition matrix

$$(4.2) \qquad \begin{aligned} \hat{p}_\varepsilon(x_i, x_j) &:= \nu_{E_i}(x_i) \mathbb{P}_\varepsilon^{x_i}(X_{\tau_{S_0}^+} = x_j) \quad \text{for } i \neq j, \\ \hat{p}_\varepsilon(x_i, x_i) &:= \nu_{E_i}(x_i) \mathbb{P}_\varepsilon^{x_i}(X_{\tau_{S_0}^+} = x_i) + 1 - \nu_{E_i}(x_i), \end{aligned}$$

then the relevant dynamical quantities will be asymptotically independent of the choice of the representatives x_i.

The occurrence of the expression $\mathbb{P}_\varepsilon^{x_i}(X_{\tau_{S_0}^+} = x_j)$ in (4.2) is intuitively obvious, since it means that we just monitor the chain when it hits one of our reference points x_j. The factor $\nu_{E_i}(x_i)$ may be less obvious. To motivate it, note that by (3.5),

$$(4.3) \qquad \begin{aligned} \mathbb{P}_\varepsilon^{x_i}(X_{\tau_{S_0}^+} = x_j) &= \sum_{z \in S \setminus E_i} \mathbb{P}_\varepsilon^x(\tau_{E^c} < \tau_{x_i}^+, X_{\tau_{E^c}} = z) \mathbb{P}_\varepsilon^z(X_{\tau_{S_0}} = x_j) \\ &\simeq \frac{1}{\nu_{E_i}(x_i)} \sum_{w \in E_i, z \notin E_i} \nu_{E_i}(w) p_\varepsilon(w, z) \mathbb{P}_\varepsilon^z(X_{\tau_{S_0}} = x_j). \end{aligned}$$

This shows that the factor $\nu_{E_i}(x_i)$ in (4.2) cancels one of the dependencies of $\mathbb{P}_\varepsilon^{x_i}(X_{\tau_{S_0}^+} = x_j)$ on the choice of our set S_0. While the terms $\mathbb{P}_\varepsilon^z(X_{\tau_{S_0}^+} = x_j)$ still do depend on the choice of S_0, we will see below that including the factor $\nu_{E_i}(x_i)$ in the definition is enough to obtain the asymptotically correct stationary distribution and escape probabilities. This justifies the following definition:

Definition 1. *Let $X^{(\varepsilon)}$ be an irreducibly perturbed Markov chain on a finite state space. The Markov chain $\hat{X}^{(\varepsilon)}$ with state space S_0 and transition matrix (4.2) is called the **effective metastable representation** of $X^{(\varepsilon)}$ corresponding to S_0.*

In order to show the properties of the chain $\hat{X}^{(\varepsilon)}$ announced above, we define a second effective Markov chain, this time without reference to a set of representatives. For $E, E' \in \mathcal{E}$ with $E \neq E'$ we put

$$(4.4) \qquad \hat{q}_\varepsilon(E, E') := \sum_{x \in E} \nu_E(x)^2 \, \mathbb{P}_\varepsilon^x(\tau_{E'}^+ < \tau_x^+),$$

and $\hat{q}_\varepsilon(E, E) := 1 - \sum_{E' \in \mathcal{E} \setminus \{E\}} \hat{q}_\varepsilon(E, E')$. The $\hat{q}_\varepsilon$ are the elements of a transition matrix when ε is sufficiently small. As the following Proposition shows, this chain is reversible and the reversible measure of $E \in \mathcal{E}$ is $\mu(E)$:

Proposition 4.1. *The quantities $\hat{q}_\varepsilon$ satisfy the asymptotic detailed balance equation*

$$\mu_\varepsilon(E)\hat{q}_\varepsilon(E, E') \simeq \mu_\varepsilon(E')\hat{q}_\varepsilon(E', E).$$

The proof of Proposition 4.1 rests on the following simple lemma:

Lemma 4.2. *Let E, E' be P_0-essential classes, $E \neq E'$, $x \in E$, $y \in E'$, and $z \in S$. Then*
$$(4.5) \qquad \mathbb{P}_\varepsilon^z(\tau_y^+ < \tau_x^+) \simeq \mathbb{P}_\varepsilon^z(\tau_{E'}^+ < \tau_x^+).$$

Proof. From Lemma 3.2, we have $\mathbb{P}^{\tilde{y}}(\tau_y < \tau_x) \simeq 1$ for all $\tilde{y} \in E'$. Since $\{\tau_y^+ < \tau_x^+\} \subset \{\tau_{E'}^+ < \tau_x^+\}$ for all $x \in E$, the strong Markov property gives

$$\mathbb{P}_\varepsilon^z(\tau_y^+ < \tau_x^+) = \sum_{\tilde{y} \in E'} \mathbb{P}_\varepsilon^z(\tau_{E'}^+ < \tau_x^+, X_{\tau_{E'}^+} = \tilde{y})\mathbb{P}_\varepsilon^{\tilde{y}}(\tau_y < \tau_x)$$

$$\simeq \sum_{\tilde{y} \in E'} \mathbb{P}_\varepsilon^z(\tau_{E'}^+ < \tau_x^+, X_{\tau_{E'}^+} = \tilde{y}) = \mathbb{P}_\varepsilon^z(\tau_{E'}^+ < \tau_x^+)$$

$\square$

Proof of Proposition 4.1. When $E = E'$, the claim holds trivially. For $E \neq E'$, pick $x \in E$ and $y \in E'$. We use Corollary 3.3 b), Proposition 2.1, and (4.5) to find

$$(4.6) \qquad \mu_\varepsilon(E)\nu_E(x)\mathbb{P}_\varepsilon^x(\tau_{E'}^+ < \tau_x^+) \simeq \mu_\varepsilon(E')\nu_{E'}(y)\mathbb{P}_\varepsilon^y(\tau_E^+ < \tau_y^+),$$

for all $x \in E$. Since the right-hand side is independent of x, and the left-hand side is independent of y, we find

$$(4.7) \qquad \nu_E(x)\mathbb{P}_\varepsilon^x(\tau_{E'}^+ < \tau_x^+) \simeq \nu_E(\tilde{x})\mathbb{P}_\varepsilon^{\tilde{x}}(\tau_{E'}^+ < \tau_{\tilde{x}}^+)$$

for all $x, \tilde{x} \in E$, and similarly for E'. Thus when we multiply (4.6) with $\nu_E(x)\nu_{E'}(y)$ and sum over $x \in E$ and $y \in E'$, we obtain the claim. $\square$

The next result shows that the effective metastable representation $\hat{X}^{(\varepsilon)}$ indeed describes the metastable dynamics of $X^{(\varepsilon)}$ correctly, in the sense that asymptotically it has the right escape probabilities and thus the right stationary distribution. Let us write $\hat{\mu}_\varepsilon$ for the stationary distribution and $\hat{\mathbb{P}}_\varepsilon$ for the path measure of $\hat{X}^{(\varepsilon)}$.

12

Theorem 4.3. *For $i \neq j$, we have $\hat{\mathbb{P}}_\varepsilon^{x_i}(\tau_{x_j}^+ < \tau_{x_i}^+) \simeq \hat{q}_\varepsilon(E_i, E_j)$. In particular $\hat{\mu}_\varepsilon(x_i) \simeq \mu_\varepsilon(E_i)$.*

Proof. From (4.7) and Lemma 4.2, we see that

$$\hat{q}_\varepsilon(E_i, E_j) \simeq \nu_{E_i}(x_i)\mathbb{P}_\varepsilon^{x_i}(\tau_{E_j}^+ < \tau_{x_i}^+) \simeq \nu_{E_i}(x_i)\mathbb{P}_\varepsilon^{x_i}(\tau_{x_j}^+ < \tau_{x_i}^+).$$

The Markov property and the definition of $\hat{P}_\varepsilon$ then gives

$$\hat{q}_\varepsilon(E_i, E_j) \simeq \hat{p}_\varepsilon(x_i, x_j) + \sum_{k \neq i,j} \hat{p}_\varepsilon(x_i, x_k)\mathbb{P}_\varepsilon^{x_k}(\tau_{x_j}^+ < \tau_{x_i}^+).$$

We will show below that for $k \neq i, j$,

(4.8) $$\mathbb{P}_\varepsilon^{x_k}(\tau_{x_j}^+ < \tau_{x_i}^+) = \hat{\mathbb{P}}_\varepsilon^{x_k}(\tau_{x_j}^+ < \tau_{x_i}^+).$$

Once this is done, the Markov property for $\hat{\mathbb{P}}_\varepsilon$ shows the first claim, and from Proposition 4.1 we get

$$\mu_\varepsilon(E_i)\hat{\mathbb{P}}^{x_i}(\tau_{x_j} < \tau_{x_i}) \simeq \mu_\varepsilon(E_i)\hat{q}(E_i, E_j) \simeq \mu_\varepsilon(E_j)\hat{q}(E_j, E_i) \simeq \mu_\varepsilon(E_j)\hat{\mathbb{P}}^{x_j}(\tau_{x_j} < \tau_{x_i}).$$

Since $\hat{\mu}_\varepsilon(x_i)\hat{\mathbb{P}}^{x_i}(\tau_{x_j} < \tau_{x_i}) = \hat{\mu}_\varepsilon(x_j)\hat{\mathbb{P}}^{x_j}(\tau_{x_i} < \tau_{x_j})$ by Proposition 2.1, we get

$$\frac{\hat{\mu}_\varepsilon(x_i)}{\hat{\mu}_\varepsilon(x_j)} \simeq \frac{\mu_\varepsilon(E_i)}{\mu_\varepsilon(E_j)}.$$

Since $\sum_j \mu_\varepsilon(E_j) \simeq 1$ by Corollary 3.3 a), we can sum over i and obtain $1/\hat{\mu}_\varepsilon(x_i) \simeq 1/\mu_\varepsilon(E_i)$, and thus $\hat{\mu}_\varepsilon(x_i) \simeq \mu_\varepsilon(E_i)$.

To show (4.8), we introduce the shorthand

$$\nu_k = \nu_{E_k}(x_k), \quad p(k,l) = \mathbb{P}^{x_k}(X_{\tau_{S_0}^+} = x_l), \quad \hat{p}(k,l) = \hat{p}_\varepsilon(x_k, x_l) = \hat{\mathbb{P}}^{x_k}(\hat{X}_{\tau_{S_0}^+} = x_l).$$

From (4.2), we get $\hat{p}(k,l) = \nu_k p(k,l) + (1 - \nu_k)\delta_{k,l}$. Now a standard application of the Markov property with the stopping time $\tau_{S_0}^+$ shows that for $k \neq i, j$, $k \mapsto h(k) = \mathbb{P}_\varepsilon^{x_k}(\tau_{x_j}^+ < \tau_{x_i}^+)$ is the unique solution of the harmonic equation $\sum_{l=1}^n p(k,l)h(l) = h(k)$ for all $k \neq i, j$ with boundary conditions $h(i) = 0$, $h(j) = 1$. Likewise, $k \mapsto \hat{h}(k) = \hat{\mathbb{P}}_\varepsilon^{x_k}(\tau_{x_j}^+ < \tau_{x_i}^+)$ is the unique solution of the harmonic equation $\sum_{l=1}^n \hat{p}(k,l)\hat{h}(l) = \hat{h}(k)$ for all $k \neq i, j$ with boundary conditions $\hat{h}(i) = 0$, $\hat{h}(j) = 1$. But since

$$\sum_{l=1}^n \hat{p}(k,l)h(l) = \nu_k \sum_{l=1}^n p(k,l)h(l) + (1 - \nu_k)h(k) = \nu_k h(k) + (1 - \nu_k)h(k) = h(k),$$

we must have $\hat{h}(k) = h(k)$, and the claim follows. $\qquad\square$

The advantage of the chain $\hat{X}^{(\varepsilon)}$ is that its transition matrix is almost diagonal in the sense that $\lim_{\varepsilon \to 0} \hat{p}_\varepsilon(x_i, x_j) = \delta_{i,j}$. In particular, $\hat{X}^{(\varepsilon)}$ is an irreducible perturbation of the trivial (identity) Markov chain. It is now natural to rescale time so that the most likely transition between two different states becomes of order one. More precisely, we set

(4.9) $$\check{p}_\varepsilon(x_i, x_j) := \frac{\hat{p}_\varepsilon(x_i, x_j)}{\sum_{k,l:k\neq l} \hat{p}_\varepsilon(x_k, x_l)}, \qquad \check{p}_\varepsilon(x_i, x_i) := 1 - \sum_{j:j\neq i} \check{p}_\varepsilon(x_i, x_j).$$

Since $\sum_{k,l:k\neq l} \check{p}_\varepsilon(x_k, x_l) = 1$, for each $\varepsilon > 0$ at least one of the terms in the finite sum must be large. The problem is that at this point we cannot guarantee that the quantities $\check{p}_\varepsilon(x_i, x_j)$ converge. To see what could happen, consider the example $S = \{x, y\}$, $p_\varepsilon(x, y) = \varepsilon(2 + \sin(1/\varepsilon))$, $p_\varepsilon(y, x) = \varepsilon$. Then $\hat{P} = P$, but $\check{p}_\varepsilon(y, x) = \frac{1}{3+\sin(1/\varepsilon)}$ does not converge. Of course, this also implies that $\lim_{\varepsilon \to 0} \mu_\varepsilon$ does not exist.

13

So far, we did not have to pay attention to that type of problem - all of our results above are valid as asymptotic equivalences, whether or not the quantities in question converge. Now however, we need proper convergence to carry on, and will give a sufficient criterion. Let $\varepsilon \mapsto a_\varepsilon$, $\varepsilon \mapsto b_\varepsilon$ be two functions of $\varepsilon > 0$. We say that a_ε and b_ε are *asymptotically comparable*, and write $a_\varepsilon \sim b_\varepsilon$, if either both of them are strictly positive and $\lim_{\varepsilon \to 0} a_\varepsilon / b_\varepsilon$ exists in $[0, \infty]$, or if one or both of them are identically zero. Note that we allow 0 and ∞ as possible limits. We caution the reader that unlike asymptotic equivalence, asymptotic comparability is not transitive, and is not stable under multiplications. On the other hand, it is obviously symmetric, and we have the following summability property: If $a_\varepsilon, b_\varepsilon$, and c_ε are mutually asymptotically comparable, and if $\alpha_\varepsilon, \beta_\varepsilon$ and γ_ε have strictly positive, finite limits as $\varepsilon \to 0$, then

$$(4.10) \qquad \alpha_\varepsilon a_\varepsilon + \beta_\varepsilon b_\varepsilon \sim \gamma_\varepsilon c_\varepsilon.$$

We say that an irreducibly perturbed Markov chain $X^{(\varepsilon)}$ is *regular* if for all $m, n \in \mathbb{N}$ and all sequences of pairs $(x_i, y_i)_{i \leqslant n}$, $(z_i, w_i)_{i \leqslant m}$ with $x_i, y_i, z_i, w_i \in S$, we have

$$(4.11) \qquad \prod_{i=1}^{n} p_\varepsilon(x_i, y_i) \sim \prod_{i=1}^{m} p_\varepsilon(z_i, w_i).$$

We will call a transition matrix P regular if the generated Markov chain is regular.

Examples of regular perturbed Markov chains include those treated in [28], where the transition elements are of the form $c_\varepsilon(x, y)\varepsilon^{k(x,y)}$ with c_ε either converging to a strictly positive limit or identically zero, and $k(x, y)$ independent of ε. They also include those with property $\mathcal{P}$ introduced in [21].

Theorem 4.4. *For a regular perturbed Markov chain with transition matrix P_ε, define $\hat{P}_\varepsilon$ as in (4.2), and $\check{P}_\varepsilon$ as in (4.9). Then $\hat{P}_\varepsilon$ and $\check{P}_\varepsilon$ are transition matrices of regular perturbed Markov chains.*

Proof. By (4.3), for $i \neq j$

$$(4.12) \qquad \hat{p}_\varepsilon(x_i, x_j) \simeq \sum_{w \in E_i, z \notin E_i} \nu_{E_i}(w) p_\varepsilon(w, z) \mathbb{P}_\varepsilon^z(X_{\tau_{S_0}} = x_j),$$

and Proposition 2.8 gives

$$(4.13) \qquad \mathbb{P}_\varepsilon^z(X_{\tau_{S_0}} = x_j) = \sum_{\gamma \in \Gamma_{S_0^c}(z, x_j)} \prod_{i=1}^{|\gamma|-1} \frac{p_\varepsilon(\gamma_i, \gamma_{i+1})}{1 - \mathbb{P}_\varepsilon^{\gamma_i}(X_{\tau_{S_0 \cup \{\gamma_1, \dots, \gamma_i\}}^+} = \gamma_i)}$$

In Lemma 4.5 below we will show that if S_0 contains one representative of each P_0-essential class then $\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^{\gamma_i}(X_{\tau_{S_0 \cup \{\gamma_1, \dots, \gamma_i\}}^+} = \gamma_i)$ exists and is strictly smaller than one for all γ. Thus each $\lim_{\varepsilon \to 0} 1/(1 - \mathbb{P}_\varepsilon^{\gamma_i}(X_{\tau_{S_0 \cup \{\gamma_1, \dots, \gamma_i\}}^+} = \gamma_i)) \geqslant 1$ exists. In other words, $\mathbb{P}_\varepsilon^z(X_{\tau_{S_0}} = x_j)$ is given as a sum of terms of the form $c_\varepsilon(z_1, \dots, z_{n+1}) \prod_{i=1}^{n} p_\varepsilon(z_i, z_{i+1})$ with $z_i \in S_0^c \cup \{z, x_j\}$, where $\lim_{\varepsilon \to 0} c_\varepsilon(z_1, \dots, z_{n_1}) \geqslant 1$ exists for all $(z_1, \dots, z_n)$. When plugging this into (4.12), we can apply the extension of (4.10) to finite sums to show that $\hat{P}_\varepsilon$ is the transition matrix of a regular Markov chain. By (4.9), this immediately implies convergence of the transition probabilities $\check{p}_\varepsilon(x_i, x_j)$. Rewriting the second equation in (4.9) in the form

$$\check{p}_\varepsilon(x_i, x_i) = \frac{\sum_{k,l: k \notin \{l, i\}} \hat{p}_\varepsilon(x_k, x_l)}{\sum_{k,l: k \neq l} \hat{p}_\varepsilon(x_k, x_l)},$$

we see in addition that the chain $\hat{X}^{(\varepsilon)}$ is a regular perturbed Markov chain. $\qquad \square$

It remains to prove the claim used in the proof above.

Lemma 4.5. *Let $X^{(\varepsilon)}$ be a perturbed Markov chain. Assume that a set S_0 contains one element of each P_0-essential class. Let $A \subset S$ with $S_0 \subset A$. Then for all $x \in A \setminus S_0$, $\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^x(X_{\tau_A^+} = x)$ exists and is strictly smaller than 1.*

Proof. As S_0 contains a representative of each P_0-essential class, there must be a P_0-relevant direct path γ from x to some $y \in S_0$. So, $\limsup_{\varepsilon \to 0} \mathbb{P}_\varepsilon^x(X_{\tau_A^+} = x) < 1 - \lim_{\varepsilon \to 0} P_\varepsilon(\gamma) < 1$.

For the existence of the limit, let first $A := S$. For $x \notin S_0$, we have $\mathbb{P}_\varepsilon(X_{\tau_A^+} = x) = p_\varepsilon(x, x) \to p_0(x, x)$ as $\varepsilon \to 0$. Let us now assume that the claim holds for all $\bar{A}$ such that $|\bar{A}| \geqslant |S| - k + 1$ with some $k \in \mathbb{N}$. Let A be such that $|A| = |S| - k$. Then,

$$\mathbb{P}_\varepsilon^x(X_{\tau_A^+} = x) = p_\varepsilon(x, x) + \sum_{y \notin A} p_\varepsilon(x, y) \mathbb{P}_\varepsilon^y(X_{\tau_A} = x)$$

$$= p_\varepsilon(x, x) + \sum_{y \notin A} p_\varepsilon(x, y) \sum_{\gamma \in \Gamma_{S \setminus A}(y, x)} \prod_{i=1}^{|\gamma|-1} \frac{p_\varepsilon(\gamma_i, \gamma_{i+1})}{1 - \mathbb{P}_\varepsilon^{\gamma_i}(X_{\tau_{A \cup \{\gamma_1, \dots, \gamma_i\}}} = \gamma_i)}.$$

By the induction hypothesis, $\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^{\gamma_i}(X_{\tau_{A \cup \{\gamma_1, \dots, \gamma_i\}}} = \gamma_i)$ exists and is strictly smaller than 1. Thus also $\lim_{\varepsilon \to 0} \mathbb{P}_\varepsilon^x(X_{\tau_A^+} = x)$ exists and is strictly smaller than 1. The claim follows by induction. $\qquad\square$

We have thus found a way to successively describe the multi-scale metastable dynamics of regular perturbed Markov chains: starting with the original chain $X^{(\varepsilon)}$, we derive $\hat{X}^{(\varepsilon)}$ and then $\check{X}^{(\varepsilon)}$. By Theorem 4.4, $\hat{X}^{(\varepsilon)}$ and $\check{X}^{(\varepsilon)}$ are again regular perturbed Markov chains. Moreover, all of the $\hat{P}_0$-essential classes consist of exactly one element, and $\lim_{\varepsilon \to 0} \hat{p}_\varepsilon(x_i, x_j) = 0$ whenever $i \neq j$. So, $\hat{P}_\varepsilon$ describes the effective metastable dynamics, but still in the original time scale.

The transformation from $\hat{P}_\varepsilon$ to $\check{P}_\varepsilon$ means that we go to a time scale where the most likely transitions between different states become of order one. In other words, there exist $i \neq j$ with $\lim_{\varepsilon \to 0} \check{p}(x_i, x_j) > 0$. By Lemma 3.1, this implies that $\{x_i\}$ will no longer be a $\check{P}_0$-essential class on its own: it will either form a larger $\check{P}_0$-essential class together with some $\{x_j\}$, $j \neq i$, or it will have become $\check{P}_0$-transient. In any case, the number of $\check{P}_0$-essential classes will be smaller than the number of P_0-essential classes. Thus by applying the transformations $P_\varepsilon \to \hat{P}_\varepsilon \to \check{P}_\varepsilon$ to the matrix $\check{P}_\varepsilon$, and iterating the procedure, we can recursively explore longer and longer time scales of the dynamics.

On a purely theoretical level, our theory of multi-scale metastable dynamics for regular perturbed Markov chains is thus complete. However, if one attempts to (numerically) compute the transition probabilities at the different time scales, the problem arises that all relevant expressions in our theory still contain terms of the form $\mathbb{P}_\varepsilon^z(X_{\tau_{S_0}} = x_j)$. In the next section, we will show why naive attempts to compute this quantity numerically are likely to fail, and present a numerically stable algorithm for computing them. A byproduct of our algorithm is a numerically stable method to compute the matrix elements of the transition matrix $\hat{Q}_\varepsilon$, and thus the stationary weights $\mu_\varepsilon(E)$ for all P_0-essential classes E.

5. COMPUTING HITTING PROBABILITIES AND THE ASYMPTOTIC STATIONARY DISTRIBUTION

This section deals with aspects of the numerical computation of the transition probabilities $\hat{p}_\varepsilon$ and $\hat{q}_\varepsilon$ given in (4.2) and (4.4), respectively. The starting point of our considerations

are the formulae

$$(5.1) \qquad \hat{p}_\varepsilon(x_i, x_j) = \nu_{E_i}(x_i)\Big(\sum_{j \neq i} p_\varepsilon(x_i, x_j) + \sum_{z \notin S_0} p_\varepsilon(x_i, z)\mathbb{P}_\varepsilon^z(X_{\tau_{S_0}} = x_j)\Big),$$

and

$$(5.2) \qquad \hat{q}_\varepsilon(E, E') = \sum_{x \in E} \nu_E(x)^2 \Big(\sum_{y \in E'} p_\varepsilon(x, y) + \sum_{z \notin \{x\} \cup E'} p_\varepsilon(x, z)\mathbb{P}_\varepsilon^z(\tau_{E'}^+ < \tau_x^+)\Big),$$

both of which are obtained from the definition of the respective quantities using the strong Markov property. In both cases, the task is to compute a hitting probability of the form

$$(5.3) \qquad\qquad\qquad h_{A,B}(z) := \mathbb{P}_\varepsilon^z(X_{\tau_{A \cup B}} \in A),$$

where $A, B \subset S$ and $z \in S \setminus (A \cup B)$. In the case of $\hat{q}_\varepsilon(E, E')$, $A = E'$ and $B = \{x\}$. Such hitting probabilities are well understood: $h_{A,B}$ is called the committor function in [25, 26] and the equilibrium potential of the capacitator (B, A) in [4], and is the unique harmonic continuation from $C := A \cup B$ to S of the indicator function 1_A of A. This means that $h_{A,B}$ is the unique solution of the linear system

$$(5.4) \qquad\qquad \sum_{z \in S \setminus C} (P_\varepsilon(x, z) - \delta_{x,z})h_{A,B}(z) = -r(x), \qquad x \in C^c = S \setminus C,$$

where $r(x) := P_\varepsilon 1_A(x)$. Let us write $\bar{P}_\varepsilon = (p_\varepsilon(x, y))_{x,y \in C^c}$ for the restriction of P_ε to C^c. If $C \neq \emptyset$ and P_ε is irreducible, we have seen in the proof of Proposition 2.5 that $1 - \bar{P}_\varepsilon$ is invertible. We thus find the committor function by matrix inversion:

$$(5.5) \qquad\qquad\qquad h_{A,B}(x) = [(1 - \bar{P}_\varepsilon)^{-1}r](x), \qquad x \in C^c.$$

The problem with this formula is that as $\varepsilon \to 0$, the matrix $(1 - \bar{P}_\varepsilon)$ may converge to a non-invertible matrix. In that case, some matrix elements of $(1 - \bar{P}_\varepsilon)^{-1}$ will diverge, and even though the quantities $h_{A,B}(z)$ themselves are bounded by 1 for all ε, computing them numerically becomes unreliable as $\varepsilon \to 0$. Our first result will identify situations where this cannot happen.

We call a state $y \in S$ an *asymptotic dynamical trap* (or simply a trap) with respect to C if

$$\liminf_{\varepsilon \to 0} \mathbb{P}_\varepsilon^y(\tau_C < M) = 0 \qquad \text{for all } M \in \mathbb{N}.$$

A necessary condition for y to be a trap is that there exists no direct P_0-relevant path from y into C. On the other hand, for all $z \in S$ there is at least one P_0-relevant path from y to some P_0-essential class. Thus if C intersects all P_0-essential classes, no traps will exist.

Recall that for a matrix P, the *condition number* is given by $\kappa(P) = \|P\|\|P^{-1}\|$, where $\|.\|$ is the operator norm with respect to any norm on the underlying vector space. In our case, it is convenient to use the supremum norm on the vector space.

Proposition 5.1. *Assume that $C \subset S$ is such that there are no asymptotic dynamical traps with respect to C. Then $\limsup_{\varepsilon \to 0} \kappa(1 - \bar{P}_\varepsilon) < \infty$.*

Proof. Since $\bar{P}_\varepsilon$ is substochastic, clearly $\|\bar{P}_\varepsilon\| \leqslant 1$. On the other hand, the absence of traps with respect to C allows us to find $k_0 \in \mathbb{N}$ and $c < 1$, both of them independent of ε, and so that $\mathbb{P}_\varepsilon^x(\tau_C > k_0) \leqslant c$ for all $x \in S$. The strong Markov property then implies $\mathbb{P}_\varepsilon^x(\tau_C > k) \leqslant c^{\lfloor k/k_0 \rfloor}$ for all $k \in \mathbb{N}$ and all $x \in S$. Thus for $x \in C^c$ and bounded $f : C^c \to \mathbb{R}$, we find

$$\big|(\bar{P}_\varepsilon)^k f(x)\big| = \big|\mathbb{E}_\varepsilon^x(f(X_k)1_{\{\tau_C > k\}})\big| \leqslant \|f\|_\infty \mathbb{P}_\varepsilon^x(\tau_C > k) \leqslant \|f\|_\infty c^{\lfloor k/k_0 \rfloor}.$$

16

Consequently, the left-hand side above is absolutely summable, and

$$\left|(1 - \bar{P}_\varepsilon)^{-1} f(x)\right| = \left|\sum_{k=0}^{\infty} (\bar{P}_\varepsilon)^k f(x)\right| \leqslant \|f\|_\infty (c - c^{(1+k_0^{-1})})^{-1}$$

for all $x \in C^c$. Taking the supremum over x, the claim follows. $\qquad\square$

By construction, S_0 contains precisely one point of each P_0-essential class, and thus there are no asymptotic dynamical traps with respect to S_0. By Proposition 5.1 and (5.1) we can thus compute $\hat{p}_\varepsilon(x_i, x_j)$ in a numerically stable way. In fact, the perturbative nature of the problem makes the following Newton scheme particularly useful.

Let $\bar{P}_\varepsilon = (p_\varepsilon(x, y))_{x,y \in S_0}$ denote the restriction of P_ε to S_0^c, and set $A_\varepsilon = 1 - \bar{P}_\varepsilon$. By (5.5), we need to find A_ε^{-1}. We use $B_0 = A_0^{-1}$ as a seed for the Newton iteration, and employ the usual recursion $B_{k+1} = 2B_k - B_k A_\varepsilon B_k$. By putting $\tilde{B}_k = B_k A_0$, we find $\tilde{B}_0 = 1$ and $\tilde{B}_{k+1} = 2\tilde{B}_k - \tilde{B}_k A_0^{-1} A_\varepsilon \tilde{B}_k$. So, $\tilde{B}_k$ is a polynomial in $A_0^{-1} A_\varepsilon$, and we can use the resulting commutativity to obtain

$$(5.6) \qquad\qquad B_{k+1} - A_\varepsilon^{-1} = -A_0^{-1} A_\varepsilon (B_k - A_\varepsilon^{-1}) A_0 (B_k - A_\varepsilon^{-1})$$

for all k. In the special case $k = 0$, this can be transformed to

$$(5.7) \qquad B_1 - A_\varepsilon^{-1} = A_0^{-1}(A_\varepsilon - A_0)(A_\varepsilon^{-1} - A_0^{-1}) = A_0^{-1}(\bar{P}_0 - \bar{P}_\varepsilon)(A_\varepsilon^{-1} - A_0^{-1}).$$

Thus

$$\|B_{k+1} - A_\varepsilon^{-1}\| \leqslant 2\kappa(A_0)\|\|B_k - A_\varepsilon^{-1}\|^2$$

and

$$\|B_1 - A_\varepsilon^{-1}\| \leqslant \|A_0^{-1}\|(\|A_0^{-1}\| + \|A_\varepsilon^{-1}\|)\|\bar{P}_\varepsilon - \bar{P}_0\|.$$

Proposition 5.1 guarantees that we can choose ε sufficiently small so that B_k converges to A_ε^{-1} very quickly. To illustrate this, we restrict ourselves to the special case where $P_\varepsilon = P_0 + \varepsilon R_\varepsilon$ with the matrix R_ε bounded uniformly in $\varepsilon > 0$. Then, $\|B_1 - A_\varepsilon^{-1}\| \leqslant c\varepsilon$ for some $c > 0$, and

$$\|B_{k+1} - A_\varepsilon^{-1}\| \leqslant (2\kappa(A_0))^{2^{k-1}+1}(c\varepsilon)^{2^k}.$$

This means that when we are interested only in transitions of size ε^n or bigger, we only have to calculate logarithmically (in n) many B_k. Therefore, it might seem that all is well, but this is not entirely so.

The reason is that a subtle problem arises from the multi-scale structure of the dynamics: at a given metastable time scale, it is in general not obvious what computational accuracy we need to achieve in order to obtain the asymptotically correct dynamics on longer metastable time scales. This phenomenon can best be explained by an example.

Figure 1 shows a graphical representation of a couple of metastable Markov chains. For both of them, $S = \{x, y, z, w\}$, and both of them have transition probabilities corresponding to the solid arrows: $p_\varepsilon(x, w) = p_\varepsilon(w, z) = p_\varepsilon(y, x) = \varepsilon$, $p_\varepsilon(w, y) = 1 - \varepsilon$, and $p_\varepsilon(z, w) = \varepsilon^2$. Only one of them has the dashed arrows, i.e. $p(z, y) = p(y, z) = \varepsilon$. All other transition probabilities are zero except those mapping a point to itself, which are adjusted to give a stochastic matrix. With or without the dashed arrows, $\{x\}$, $\{y\}$ and $\{z\}$ are the P_0-essential classes, while w is P_0-transient. Also in both cases, $\hat{p}_\varepsilon(x, z) = \varepsilon^2$, while $\hat{p}_\varepsilon(x, y) = \varepsilon$. So on the first metastable time scale, transitions from x to z play no role. But whether or not we can stop our computation of $\hat{p}_\varepsilon(x, z)$ after reaching order ε depends on the presence of the dashed arrows.

If the dashed arrows are present, we can stop the computation of $\hat{p}_\varepsilon$ after reaching order ε: on the next (and final) metastable time scale, we will have $\check{p}_\varepsilon(x, y) = \check{p}_\varepsilon(z, y) = 1$ and
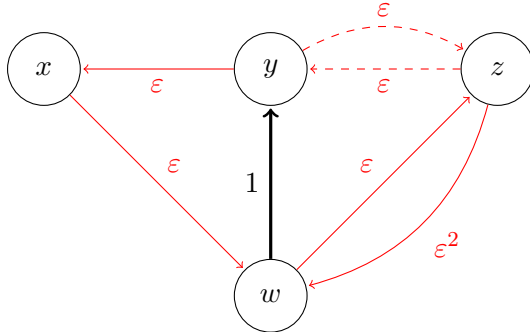
FIGURE 1. Schematic drawing of a perturbed Markov chain. Leading order transition probabilities are written on the arrows.

$\check{p}_\varepsilon(y, x) = \check{p}_\varepsilon(y, z) = 1/2$. z will be connected to x via y, by transition probabilities of order one.

However, if the dashed arrows are absent, stopping the calculation at order ε leads to an effective Markov chain where z cannot be reached from x, and thus to wrong results on the next metastable time scale. In the correct dynamics on that time scale x and y form a new effective metastable state, and transitions between it and z are (after rescaling) of order ε. For this to be resolved correctly, the transition from x to z of order ε^2 needs to be present already in the effective dynamics on the first metastable time scale.

In the simple example at hand it is easy to directly figure out what is going on, but to decide when a given approximation of $\hat{p}_\varepsilon$ is good enough to give correct dynamical results on all further metastable time scales for general chains on large state spaces is a subtle problem. Here we only give a necessary condition, about which we conjecture that it is also sufficient, and which is accessible to numerical validation. Let us write $\hat{\mathbb{P}}_{\varepsilon,a}$ for path measure of a given approximation to the chain $\hat{X}^{(\varepsilon)}$. By Theorem 4.3, $\hat{q}_\varepsilon(E_i, E_j) \simeq \hat{\mathbb{P}}_\varepsilon^{x_i}(\tau_{x_j} < \tau_{x_i})$ when x_i is the representative from E_i and x_j the representative from E_j, and thus $\hat{\mu}(x_i) \simeq \mu(E_i)$ for all i. So in order to obtain the correct asymptotic stationary distribution for our approximate chain, we have to increase the accuracy at least until

$$(5.8) \qquad \hat{q}_\varepsilon(E_i, E_j) \simeq \hat{\mathbb{P}}_{\varepsilon,a}^{x_i}(\tau_{x_j} < \tau_{x_i}).$$

It would not be surprising if this were already sufficient for some sort of agreement of the metastable dynamics on all further metastable time scales. Since in general the escape probabilities do not characterize the transition probabilities of a Markov chain, a proof of this conjecture is not immediate, and we do not pursue this any further here. Instead, we discuss how to check (5.8) numerically.

By (5.2), the numerically tricky part in computing $\hat{q}_\varepsilon(E_i, E_j)$ is $\mathbb{P}_\varepsilon^x(\tau_{E'}^+ < \tau_x^+)$. Since $\{x\} \cup E'$ will not intersect all P_0-essential classes unless there are only two of them, we cannot use Proposition 5.1 this time, and indeed in most situations a direct calculation of (5.5) will be numerically unreliable. However, for the very same reason, namely since C intersects only two P_0-essential classes, we can successively lift these traps and arrive at a simplified chain without traps for which the probability of hitting E' before x is asymptotically equivalent to the original one.

The basic step in this procedure is the following. Assume that E is a P_0-essential class of a perturbed Markov chain $X^{(\varepsilon)}$, and that $E \neq S$. We define a new Markov chain $\tilde{X}^{(\varepsilon)}$ on the state space $\tilde{S} = (S \setminus E) \cup \{E\}$ by its transition probabilities $\tilde{p}_\varepsilon(x, y)$, where

18

$\tilde{p}_\varepsilon(x, y) = p_\varepsilon(x, y)$ whenever $x, y \in S \setminus E$, and

$$(5.9) \quad \tilde{p}_\varepsilon(x, E) := \sum_{z \in E} p_\varepsilon(x, z), \qquad \tilde{p}_\varepsilon(E, x) := \frac{1}{Z_\varepsilon(E)} \sum_{z \in E} \nu_E(z) p_\varepsilon(z, x), \qquad \tilde{p}(E, E) := 0$$

for all $x \in S \setminus E$. Here, $Z_\varepsilon(E) = \sum_{z \in E, y \notin E} \nu_E(z) p_\varepsilon(z, y)$ is the normalization that ensures that $\tilde{P}$ is a stochastic matrix. We say that the traps in E (with respect to $\cup(\mathcal{E} \setminus \{E\})$) have been lifted in $\tilde{X}^{(\varepsilon)}$. This terminology is justified by

Theorem 5.2. *Let $X^{(\varepsilon)}$ be a perturbed Markov chain, E a P_0-essential class of $X^{(\varepsilon)}$, and $\tilde{X}^{(\varepsilon)}$ the Markov chain where E has been lifted.*
a) Let $A, B \subset S \setminus E$. Then for all $z \in S \setminus E$, $\mathbb{P}_\varepsilon^z(\tau_B < \tau_A) \simeq \tilde{\mathbb{P}}_\varepsilon^z(\tau_B < \tau_A)$, while for $z \in E$, $\mathbb{P}_\varepsilon^z(\tau_B < \tau_A) \simeq \tilde{\mathbb{P}}_\varepsilon^E(\tau_B < \tau_A)$.
b) If $X^{(\varepsilon)}$ is regular, then $\tilde{X}^{(\varepsilon)}$ is a regular perturbed Markov chain.
c) If $X^{(\varepsilon)}$ is regular, then either E is a $\tilde{P}_0$-transient state, or E is an element of a $\tilde{P}_0$-essential class that contains at least one further element $z \in F$. In the latter case, the number of $\tilde{P}_0$-transient states is strictly smaller than the number of P_0-transient states.

Proof. Consider the chain $Y^{(\varepsilon)}$ with state space S and transition matrix R_ε, where $r_\varepsilon(x, y) = p_\varepsilon(x, y)$ when $x \notin E$ and $r_\varepsilon(x, y) = \mathbb{P}^x(\tau_{E^c} = y)$ when $x \in E$. Denoting its path measure by $\mathbb{P}_{Y,\varepsilon}$, Proposition 2.7 gives $\mathbb{P}_{Y,\varepsilon}^z(\tau_B < \tau_A) = \tilde{\mathbb{P}}_\varepsilon^z(\tau_B < \tau_A)$ for all $z \in S$. We now define $\tilde{Y}$ by replacing $r_\varepsilon(x, y)$ with $\frac{1}{Z_\varepsilon(E)} \sum_{x \in E} \nu_E(x) p_\varepsilon(x, y)$ for $x \in E$, and keeping them the same if $x \notin E$. Then (3.6) implies that $r_\varepsilon(x, y) \simeq \tilde{r}_\varepsilon(x, y)$ for all $x, y \in S$, and thus Theorem 3.4 gives $\mathbb{P}_{\tilde{Y},\varepsilon}^z(\tau_B < \tau_A) \simeq \mathbb{P}_{Y,\varepsilon}^z(\tau_B < \tau_A)$. Finally, noting that $\tilde{r}_\varepsilon(z, w)$ does not depend on z whenever $z \in E$, we can replace all $z \in E$ by a single state $\{E\}$, and claim a) follows.

For b), note that by regularity of the chain, $Z_\varepsilon(E) \sim \sum_{z \in E} \nu_E(z) p_\varepsilon(z, x)$ for all $x \notin E$. So the quotient in (5.9) either converges or diverges to infinity as $\varepsilon \to 0$. Since it is bounded by 1 by construction, the latter is not an option, and the $\tilde{p}_\varepsilon$ converge. So the Markov chain defined by them is a perturbed Markov chain. Finally, this Markov chain is again regular, since products of its elements can be written as weighted sums of products of the p_ε with nonnegative weights. We have shown b).

For c), note that by b) $\lim_{\varepsilon \to 0} \tilde{P}_\varepsilon$ exists, and since $\sum_{y \notin E} \tilde{p}_\varepsilon(E, y) = 1$, there must be at least one state $y \in S \setminus E$ with $\lim_{\varepsilon \to 0} \tilde{p}_\varepsilon(E, y) > 0$. Lemma 3.1 implies that if E' is a P_0-essential class with $E \neq E'$, then all direct paths from E' to E are P_0-irrelevant. So if one of the elements y with $\lim_{\varepsilon \to 0} \tilde{p}_\varepsilon(E, y) > 0$ is connected to a different P_0-essential class via a P_0-relevant direct path, then E is $\tilde{P}_0$-transient. On the other hand, if no y is connected to any $E' \neq E$ by such a direct path, then each such y must be an element of F, and must be connected to E by a P_0-relevant direct path. It follows that y is in the same $\tilde{P}_0$-essential class as E, and thus not a $\tilde{P}_0$-transient state. The claim follows. $\square$

Using Theorem 5.2, we can now give a general recursive algorithm for numerically computing expressions $h_{B,A}(z)$ of the form given in (5.3) simultaneously for all $z \in S$, up to asymptotic equivalence:

(1) Determine the set $\mathcal{E}_0$ of all P_0-essential classes not intersecting $A \cup B$.
(2) If $\mathcal{E}_0 = \emptyset$, compute $h_{A,B}$ by solving the well-conditioned linear system (5.4). Finish the algorithm.
(3) Compute the P_0-stationary measures ν_E for each $E \in \mathcal{E}_0$.
(4) Lift all the traps in $E \in \mathcal{E}_0$ by (5.9). This results in a new state space, where all elements of E are replaced by a single state E. Keep track of the elements of the original state space that become lumped into E.

(5) Return to (1) with the new state space.

We note that steps (3) and (4) are trivial to parallelize. By Theorem 5.2 c), each step either decreases the number of P_0-essential classes in the chain, or leaves it unchanged and decreases the number of transient states. We thus see that the algorithm terminates. Once it does (in step 2), we know $h_{A,B}(\tilde{z})$ for all $\tilde{z}$ in the final state space $\tilde{S}$. Theorem 5.2 a) now guarantees that $h_{A,B}(z) \simeq h_{A,B}(\tilde{z})$ for all states z of the chain from the previous step that were collapsed into $\tilde{z}$. Thus we can recursively go backwards until we reach the original state space, where we now know all $h_{A,B}(z)$ up to asymptotic equivalence. In particular, this gives a stable algorithm for the asymptotic numerical approximation of the coefficients $\hat{q}_\varepsilon$. Since the expressions $\hat{\mathbb{P}}^{x_i}_{\varepsilon,a}(\tau_{x_j} < \tau_{x_i})$ are also escape probabilities (for a different Markov chain), we can compute them by the same algorithm. If they agree with $\hat{q}_\varepsilon(E_i, E_j)$ to leading order in ε, our necessary criterion is met and the approximate chain has the same asymptotic stationary measure as the true one.

Another useful aspect of our algorithm is that the $\hat{q}_\varepsilon$ determine the limiting stationary distribution of the chain through the formula

$$\frac{1}{\mu_\varepsilon(E)} \simeq \sum_{E' \in \mathcal{E}} \frac{\hat{q}_\varepsilon(E, E')}{\hat{q}_\varepsilon(E', E)},$$

which is derived in analogy to (2.4), using Proposition 4.1. Computing the stationary distribution of a large Markov chain with many metastable sets is a very important problem in practice. For example, it is how internet search engines compute page importance ranks. As a consequence, there has been tremendous activity in the computer science community on the topic. Most of the developments seem to be based on a seminal paper by Simon and Ando [22]. Seemingly independently, the problem has been treated by a much smaller group of people in mathematical economy, starting with [29] and with significant recent progress by Wicks and Greenwald [27, 28].

Both approaches are based on formula (2.8), which itself is closely related to (5.4). In the literature following [22] and [17], this leads to what is known as the method of the stochastic complement. For a finite Markov chain X on a state space S, the first step of the method is to decompose S into disjoint sets $S_1, \ldots, S_n$. Equation (2.8) with $A = S_j$ then allows to compute

$$(5.10) \qquad \qquad \hat{p}(x, y) := \mathbb{P}^x(X_{\tau^+_{S_j}} = y)$$

for $x, y \in S$ by using matrix multiplications and by computing the inverse of the matrix $(1 - P|_{S_j^c})$. The $\hat{p}(x, y)$ are the transition probabilities of an effective Markov chain only running inside S_j. Writing ν_j for the stationary distribution of the effective chain, and μ for the full stationary distribution, it can be shown that

$$(5.11) \qquad \qquad \mu(x) = \xi_j \nu_j(x)$$

for all $x \in S_j$, where $(\xi_j)_{j \leqslant n}$ is the stationary distribution of the Markov chain with state space $\{S_1, \ldots, S_n\}$ and transition probabilities

$$(5.12) \qquad \qquad q(S_i, S_j) = \sum_{x \in S_i, y \in S_j} \nu_i(x) p(x, y).$$

Equation (5.11) is similar to the statements of our Corollary 3.3, with the ξ_j taking the role of $\mu_\varepsilon(E)$, and the $\nu_j(x)$ the role of $\nu_E(x)$. Equation (5.12) is in analogy to the expression

$$(5.13) \qquad \hat{p}_\varepsilon(x_i, x_j) \simeq \sum_{w \in E_i, z \notin E_i} \nu_{E_i}(w) p_\varepsilon(w, z) \mathbb{P}^z_\varepsilon(X_{\tau_{S_0}} = x_j)$$

that we get for $\hat{p}_\varepsilon(x_i, x_j)$ when combining (4.2) and (4.3). The drawback of the method is that a priori, we have no control over the numerical difficulty of computing $(1 - P|_{S_j^c})^{-1}$. For example, let S_1 consist of two elements x, y. Then $\hat{p}(x, y) = \mathbb{P}^x(\tau_y^+ < \tau_x^+)$, and thus the computation of $\hat{p}(x, y)$ is no easier than the problem we have treated in the present paper; in particular, if X is a perturbed Markov chain and x and y are in different P_0-essential classes, the matrix $(1 - P|_{S_j^c})$ will become singular as $\varepsilon \to 0$. Therefore without any further assumptions, the theory of Simon and Ando as it stands gives no numerically feasible way of computing μ.

A suitable such further assumption is to choose the decomposition in a way that makes all transitions between different S_j small. The situation where this is possible has been treated already in [22], and is nowadays known as a the theory of nearly reducible (or nearly decomposable) Markov chains. In the framework of the present paper, a perturbed Markov chain is nearly reducible if for each $y \in S$ there exists a unique P_0-essential class $E(y)$ so that all P_0-relevant paths from y to $S \setminus F$ end in $E(y)$. In the terminology of [5], this means that the local valleys corresponding to the maximal metastable set $S_0 = \{x_1, \ldots, x_n\}$ from Section 4 do not intersect. When a Markov chain is nearly reducible, it is known (and follows from (3.2) in our case) that we can ignore transitions between different S_j for the approximate computation of the ν_j; in the case of perturbed Markov chains and when each S_j contains exactly one P_0-essential class E_j, this means $\nu_j \approx \nu_{E_j}$. The reduced chain (5.12) is then similar to our $\hat{X}_\varepsilon$, and by a recursive algorithm similar to the one given in the present section, the stationary measure μ can be computed.

So in the context of nearly reducible Markov chains, the contribution of our work is on the one hand a systematic, rigorous asymptotic theory, and on the other hand an extension to the case where the Markov chain no longer needs to be nearly reducible: in the latter case, the E_j take the role of the S_j, and the presence of the transient set is accounted for by replacing (5.12) by (5.13), together with a recipe to compute the escape probabilities contained in the latter equation.

The second approach that we are aware of which uses (2.8) is the recent work by Wicks and Greenwald [27, 28], who call their approach the method of the stochastic quotient. They work in the situation where $P_\varepsilon = P_0 + \varepsilon R_\varepsilon$ with bounded corrector matrix R_ε, and they do not need to assume almost decomposability. As we do, they pick a representative x from each P_0-essential class E. Then they apply (2.8) with $A = \{x\} \cup S \setminus E$, i.e. they compute the probabilities to either leave E at a given $y \notin E$, or to return to x. The leading order of this quantity can be computed efficiently by a matrix calculation, since the matrices $(1 - P_\varepsilon|_{A^c})^{-1}$ remain bounded as $\varepsilon \to 0$ thanks to the absence of x from A^c. Indeed, as Wicks and Greenwald note, it suffices to invert $(1 - P_0|_{A^c})$. This construction leads to an effective chain where the class E is replaced by a P_0-essential class containing just the one element x. They do this construction for all P_0-essential classes, and indeed also for transient communicating classes. After that, they rescale transition probabilities out of each of the (now trivial) P_0-essential classes much like we do in (5.9), keeping track of the factors by which they speed up each individual trap. This results in a Markov chain with fewer P_0-essential states, or fewer transient states. Recursively iterating the procedure while always keeping track of the rescaling factors, they arrive at a stable algorithm for computing the stationary distribution.

It is obvious that the algorithm of Wicks and Greenwald and ours share quite similar ideas. The difference is that while our algorithm lifts metastable traps completely, the Wicks-Greenwald algorithm keeps one point in each trap. The advantage of the Wicks-Greenwald algorithm is that the whole stationary distribution can be computed at once,

while in our algorithm one has to compute $\hat{q}(E, E')$ separately for each pair E, E'. The advantage of our approach is that it is local: if we are only interested in the relative importance of two given states $x \in E$ and $y \in E'$, we need only compute the ratio $\hat{q}(E, E')/\hat{q}(E', E)$. Depending on the structure of the chain, this can be done by lifting only a tiny fraction of the traps present in the state space. An additional advantage of our approach is of course that we also obtain information about the metastable dynamics, information which is not contained in the stationary distribution alone.

REFERENCES

[1] D. Aldous, J.A. Fill. *Reversible Markov Chains and Random Walks on Graphs.* Unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/~aldous/RWG/book.html

[2] J. Beltran, C. Landim. Tunneling and Metastability of Continuous Time Markov Chains II, the Non-reversible Case. *Journal of Statistical Physics* 149, 598-618, 2012.

[3] N. Berglund, B. Gentz. *Noise-induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach*, Springer (2006)

[4] A. Bovier: Metastability. In: R. Kotecky (ed.), *Methods of Contemporary Mathematical Statistical Physics*, Lecture Notes in Mathematics 2009, Springer, 2006.

[5] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and low-lying spectra in reversible Markov chains. *Commun. Math. Phys.*, 228:219255, 2002

[6] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes 1. sharp asymptotics for capacities and exit times. *J. Europ. Math. Soc. (JEMS)*, 6:399424, 2004.

[7] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes 2. precise asymptotics for small eigenvalues. *J. Europ. Math. Soc. (JEMS)*, 7:6999, 2005

[8] P. Deuflhard, W. Huisinga, A. Fischer, Ch. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Linear Algebra and its Applications* 315, 39-59, 2000.

[9] G. Ellison: Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. The Review of Economic Studies 67, 17-45, 2000,

[10] H. Eyring. The activated complex in chemical reactions. *J. Chem. Phys.*, 3:107115, 1935.

[11] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems,* second edition. Springer-Verlag, New York, 1998.

[12] A. Gaudilliere, C. Landim. *A Dirichlet principle for non reversible Markov chains and some recurrence theorems*, *Probability Theory and Related Fields* 158, 55-89, 2014.

[13] H.A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica,* 7:284304, 1940.

[14] A.N. Langville, C.D. Meyer. *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press 2006.

[15] G. Louchard, G. Latouche. Geometric bounds on iterative approximations for nearly completely decomposable Markov chains. *J Appl Probability* 27, 521-529 (1990).

[16] D.A. Levin, Y. Peres, E.L. Wilmer: *Markov Chains and Mixing Times.* AMS Publishing 2008.

[17] C.D. Meyer. Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems. *SIAM Review,* 31(2), 240272, 1989.

[18] A. Milias-Argeitis, J. Lygeros. Efficient stochastic simulation of metastable Markov chains, *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC) Orlando, FL, USA, December 12-15,* 2011

[19] E. Olivieri, M.E. Vares: *Large Deviations and Metastability*, Cambridge University, 2005.

[20] E. Olivieri, E. Scoppola. Markov chains with exponentially small transition probabilities: First exit problem from a general domain I. The reversible case. *J. Stat. Phys* 79, 613647, 1995.

[21] E. Olivieri, E. Scoppola. Markov chains with exponentially small transition probabilities: First exit problem from a general domain II. The general case. *J. Stat. Phys.* 84, 9871041, 1996.

[22] H.A. Simon, A. Ando. Aggregation of variables in dynamic systems. *Econometrica: Journal of the Econometric Society,* 111138, 1961.

[23] H. De Sterck, K. Miller, E. Treister, I. Yavneh. Fast multilevel methods for Markov chains. *Numerical Linear Algebra with Applications* 18, 961980, 2011.

[24] R.M. Tifenbach. A combinatorial approach to nearly uncoupled Markov chains 1: reversible Markov chains. *Electronic Transactions on Numerical Analysis* 40, 120-147, 2013.

[25] E. Vanden-Eijnden: Transition path theory. In: Ferrario, M., Ciccotti, G., Binder, K. (eds.) *Computer Simulations in Condensed Matter: From Materials to Chemical Biology*, pp. 439478. Springer, 2006

[26] M. Cameron, E. Vanden-Eijnden: Flows in Complex Networks: Theory, Algorithms, and Application to LennardJones Cluster Rearrangement. *J. Stat. Phys.* 156, 427-454, 2014.

[27] J.R. Wicks, A. Greenwald: An Algorithm for Computing Stochastically Stable Distributions with Applications to Multiagent Learning in Repeated Games. *Proceedings of the 21st Conference in Uncertainty in Artificial Intelligence, 2005.* AUAI Press 2005

[28] J.R. Wicks, A. Greenwald: A Quotient Construction on Markov Chains with Applications to the Theory of Generalized Simulated Annealing. *International Symposium on Artificial Intelligence and Mathematics (ISAIM 2006), Fort Lauderdale, Florida, USA.*

[29] H.P. Young. The evolution of conventions. *Econometrica* 61, 57-84, 1993.

VOLKER BETZ
FACHBEREICH MATHEMATIK
TU DARMSTADT
SCHLOSSGARTENSTRASSE 7, 64289 DARMSTADT
http://www.mathematik.tu-darmstadt.de/∼betz/
E-mail address: betz@mathematik.tu-darmstadt.de

STÉPHANE LE ROUX
DÉPARTEMENT D'INFORMATIQUE
UNIVERSITÉ LIBRE DU BRUXELLES
ULB CP212, BOULEVARD DU TRIOMPHE, 1050 BRUXELLES
http://www.ulb.ac.be/di/verif/sleroux/
E-mail address: stephane.le.roux@ulb.ac.be