Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 126 (2016) 3499–3526

stochastic processes and their applications

www.elsevier.com/locate/spa

Multi-scale metastable dynamics and the asymptotic stationary distribution of perturbed Markov chains

Volker Betz[a,*], Stéphane Le Roux[b]

[a] Fachbereich Mathematik, TU Darmstadt, Schlossgartenstrasse 7, 64289 Darmstadt, Germany
[b] Département d'Informatique, Université Libre Du Bruxelles, ULB CP212, Boulevard du Triomphe, 1050 Bruxelles, Belgium

Received 23 July 2015; received in revised form 3 May 2016; accepted 6 May 2016
Available online 13 May 2016

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.

MSC: 60J10; 60J22

Keywords: Escape times; Non-reversible Markov chains; Asymptotics

* Corresponding author.
E-mail addresses: betz@mathematik.tu-darmstadt.de (V. Betz), stephane.le.roux@ulb.ac.be (S. Le Roux).
URLs: http://www.mathematik.tu-darmstadt.de/~betz/ (V. Betz), http://www.ulb.ac.be/di/verif/sleroux/ (S. Le Roux).

http://dx.doi.org/10.1016/j.spa.2016.05.003
0304-4149/

1. Introduction

Loosely speaking, a metastable Markov chain looks like a stationary Markov chain on short time scales, exploring only a small subset of its state space; on longer time scales, however, it performs fast and rare transitions between different such subsets.

In this article, we consider a special class of metastable Markov chains, for which a rather complete understanding of both the dynamics and the stationary distribution are possible. We assume that the state space S is fixed and finite, that the transition matrices P_ε depend continuously on a parameter $\varepsilon \geqslant 0$, and that the chain is irreducible for $\varepsilon > 0$ but has several essential communicating classes when $\varepsilon = 0$. We do not assume reversibility. Such a family $X^{(\varepsilon)}$ of Markov chains will be called an irreducible perturbation of the $\varepsilon = 0$ chain, or simply an irreducibly perturbed Markov chain.

The dynamics of an irreducibly perturbed chain has a multi-scale structure. The general picture is that for small $\varepsilon > 0$, and on times of order 1, the chain simply explores one of the essential classes of the chain $X^{(0)}$. The next relevant time scale is determined by the inverse of the transition probabilities between the different essential classes. More precisely, let us assume that $X^{(0)}$ has the essential communicating classes $E_1, \dots, E_k$, and pick a representative $x_j \in E_j$ from each class. The effective transition probability from x_i to x_j is given by $p_{i,j}(\varepsilon) := \mathbb{P}^{x_i}(\text{the first hit of } (X_n^{(\varepsilon)})_{n\geqslant 1} \text{ on the set } \{x_1, \dots, x_k\} \text{ occurs in } x_j)$. By a geometric trials argument it will take the Markov chain a time of order $1/p_{i,j}(\varepsilon)$ to travel from x_i to x_j. By the continuity of P_ε, $p_{i,j}(\varepsilon)$ converges to zero as $\varepsilon \to 0$ for different i, j, and the pair (i, j) with the slowest speed of convergence determines the most frequent transitions between different essential classes of $X^{(0)}$. Its inverse determines the first metastable time scale. On that time scale, the motion inside each of the E_j is too fast to be resolved, and the essential classes E_j of $X^{(0)}$ act as the states of the new Markov chain. This justifies (a posteriori) the choice of one representative per essential class.

The aim of our paper is to make the above picture rigorous and quantitatively precise. In particular, the above consideration is only correct to the extent that the $p_{i,j}$ give the correct orders of magnitude, but they will in general fail to describe important details (such as the asymptotic stationary distribution) of the metastable effective chain correctly. In particular, these asymptotic properties will in general depend on the choice of representative from each E_j. The first main result of the paper is that this defect is cured when modifying the definition of the $p_{i,j}$ by multiplying each $p_{i,j}$ with the 'local' asymptotic stationary weight of the starting point x_i, see (5.2) and Definition 1. We then obtain an effective chain $\hat{X}^{(\varepsilon)}$ with state space $\{x_1, \dots, x_k\}$ whose relevant asymptotic properties are independent of the choice of representatives. In particular, it turns out that its asymptotic escape probabilities, i.e. the quantities $\mathbb{P}^{x_i}(\hat{X}^{(\varepsilon)} \text{ hits } x_j \text{ before returning to } x_i)$ are asymptotically independent of the choice of representative, and asymptotically equivalent to the transition probabilities $\hat{q}_\varepsilon(E_i, E_j)$ of a natural, *reversible* Markov chain running on the state space composed of the essential classes $E_1, \dots, E_n$; see (5.4) and Theorem 5.3. We thus see that $\hat{X}^{(\varepsilon)}$ describes the behavior of $X^{(\varepsilon)}$ on the first metastable time scale.

In order to explore longer metastable time scales, we want to renormalize the effective chain: for $\hat{X}^{(\varepsilon)}$, all transitions between different states vanish in the limit $\varepsilon \to 0$, so we rescale time such that the fastest transitions of $\hat{X}^{(\varepsilon)}$ happen on a time scale of order one after rescaling. We would then like to repeat the procedure described above by introducing effective hitting probabilities for the rescaled chain, but here we face the problem that we do not have any guarantee that the rescaled transition matrices will still converge as $\varepsilon \to 0$. In the second main result of the paper, Theorem 5.4, we show that a rather weak and natural assumption on the original chain, which