Journal of Statistical Physics, Vol. 84, Nos. 5/6, 1996

Markov Chains with Exponentially Small Transition Probabilities: First Exit Problem from a General Domain. II. The General Case

E. Olivieri[1] and E. Scoppola[2]

Received July 18, 1995; final February 19, 1996

In this paper we consider aperiodic ergodic Markov chains with transition probabilities exponentially small in a large parameter β. We extend to the general, not necessarily reversible case the analysis, started in part I of this work, of the first exit problem from a general domain Q containing many stable equilibria (attracting equilibrium points for the $\beta = \infty$ dynamics). In particular we describe the tube of typical trajectories during the first excursion outside Q.

KEY WORDS: Markov chains; first exit problem; large deviations.

1. INTRODUCTION

In this paper we extend to the general, not necessarily reversible case the analysis, started in ref. 11 for the reversible case, of the typical exiting trajectories during the first excursion from a general domain Q.

More precisely, let $\{X_t^{(\beta)}\}_{t=0,1,2,...}$ be a family of Markov chains defined on the finite state space S, with transition probabilities $P^{(\beta)}(x, y)$ depending on a positive parameter β and satisfying the following conditions:

1. *Ergodicity condition*:

$$\forall x, y \in S \quad \exists n \qquad \text{such that} \quad P^{(\beta)n}(x, y) > 0$$

(where $P^{(\beta)n}(x, y)$ is the n-step transition probability).

[1] Dipartimento di Matematica, Università di Roma Tor Vergata, 00173 Rome, Italy; e-mail: Olivieri@mat.utovrm.it.

[2] Dipartimento di Matematica, Terza Università di Roma, 00146 Rome, Italy; e-mail: scoppola@mat.uniroma3.it.

0022-4715/96/0900-0987$09.50/0

2. *Property* $\mathscr{P}$: there exist a function $\Delta(x, y)$, $x, y \in S$, assuming values $\Delta_0 = 0 < \Delta_1 < \Delta_2 < \cdots < \Delta_n$, for some positive integer n, with $\Delta_n < \infty$ and a positive function $\gamma = \gamma(\beta)$, with $\gamma \to 0$ as $\beta \to \infty$, such that if $x \neq y$ and $P^{(\beta)}(x, y) > 0$, then

$$\exp\{-\Delta(x, y)\,\beta - \gamma\beta\} \leqslant P^{(\beta)}(x, y) \leqslant \exp\{-\Delta(x, y)\,\beta + \gamma\beta\} \tag{1.1}$$

We will denote by $X_t(x)$ the Markov chain at time $t \in \mathbb{N}$ starting from x at time 0; we will omit everywhere the index β, for notational simplicity.

We will denote by P_x the probability distribution of the process starting from x at $t = 0$ and by E_x the corresponding expectation. Moreover, given any set of states $Q \subset S$, we will denote by τ_Q the first hitting time to Q:

$$\tau_Q \equiv \min\{t > 0: X_t \in Q\} \tag{1.2}$$

For any set $Q \subset S$ we will denote by $Q^c \equiv S \setminus Q$ the complement of Q.

The aim of this paper is to provide a complete description of the typical behavior of the Markov chain X_t up to the time τ_{Q^c}, for any set $Q \subset S$, and for β sufficiently large.

We refer to ref. 11 for a general discussion of the problem. Here we want only to recall that the results obtained by Freidlin and Wentzell concern only the asymptotics for β large of the first exit time τ_{Q^c} and of the first exit point $X_{\tau_{Q^c}}$. The description of the tube of typical trajectories was given by Freidlin and Wentzell(5) only in the case of a domain Q completely attracted by a unique stable equilibrium point.

It turns out from the analysis of many particular models (see, for instance, refs. 7–10) that for general domains the typical escape involves the permanence of the process in suitable sets during suitable random times, exponentially diverging in β. This sort of "temporal entropy" is essential to provide an efficient mechanism of escape.

In ref. 11, by exploiting reversibility, we were able to reduce the solution of the problem to the analysis of the energy landscape. In particular, the decomposition of the space into special sets called "cycles," which play a crucial role in the theory, was simply obtained in terms of the energy.

In the present paper, to study the general nonreversible case, in order to get the cycle decomposition, we are forced to use graphical methods like those introduced by Freidlin and Wentzell.[5]

Our strategy will be to combine this graphical approach with an analysis in terms of increasing scales of time introduced in ref. 12. In that paper the long-time behavior of the chain X_t was studied by constructing a sequence of *renormalized Markov chains* $X_t^{(1)}, X_t^{(2)}, \ldots, X_t^{(j)}, \ldots$ whose state spaces $S^{(1)}, S^{(2)},..., S^{(j)},...$ were composed by equilibrium points of increasing stability. These chains provide a rougher and rougher description of our stochastic evolution adapted to the analysis of phenomena taking place in increasing scales of time (exponential in β).

If $S \equiv S^{(0)}$ is the original state space, then $S^{(1)}$ is just the set of stable equilibria for the original Markov chain $X_t \equiv X_t^{(0)}$; $X_t^{(1)}$ is suitably defined on $S^{(1)}$. The set $S^{(2)}$ is the set of stable equilibria for $X_t^{(1)}$ and so on. In Section 2 we will recall all the necessary definitions given in ref. 12. In ref. 12 with this construction of renormalized chains, some results on the typical long-time behavior of the original chain X_t were easily obtained. In fact to each exponentially long path of the chain X_t, a short path of a chain $X_t^{(N)}$ was associated, with a suitable renormalization index N.

However, a detailed description of the behavior of the original chain X_t during each interval of time corresponding to each transition of the chain $X^{(N)}$ was missing in ref. 12.

This detailed description turns out to be strictly connected to the problem of the definition of the typical exiting tube. Indeed let $N = N(Q)$ be the level such that the $(N+1)$th renormalized Markov chain does not contain states inside Q: $S^{(N+1)} \cap Q = \emptyset$. This means that the first excursion outside Q for the chain $X_t^{(N)}$ is a sort of "descent" along the drift.

As we already sketched in ref. 11, a first rough approximation of the typical tube of escape from Q is given by the set of typical trajectories followed during the first excursion outside Q by the chain $X_t^{(N)}$.

There remains the question of "reading" the result in terms of the paths followed on the original scale of time by our original chain X_t.

This problem of analyzing the set of trajectories of the original chain X_t corresponding to a given trajectory of the chain $X_t^{(N)}$ is solved in the present paper. As noted above, this not only will give a full characterization of the typical tube of trajectories followed by the original chain X_t during the first excursion from Q; this paper completes the analysis introduced in ref. 12 based on the renormalization procedure. Namely, we will be able to associate to any state x^N of $S^{(N)}$ a suitable set $Q_{x^N} \in S$, a sort of generalized cycle, representing the set where the original process X_t typically remains in the interval of time corresponding to a jump of the chain $X_t^{(N)}$.

We will also analyze the typical "descent" to the "bottom" of a generalized basin of attraction of a stable equilibrium $x^N \in S^{(N)}$. In this case we can make a comparison with our previous results in the reversible case. The situation when analyzing the typical "ascent" against the drift is more complicated and, of course, in the nonreversible case typical ascents outside a generalized basin Q and typical descents to the bottom of Q are not related by time reversal.

It turns out that, similar to what happens in the reversible case, the typical descent to the bottom of a domain containing many attractors takes place in a way that can be considered as the natural generalization of the typical descent to the bottom x of a domain Q completely attracted by the unique stable equilibrium point x. The main difference is the following: in the completely attracted domain the system does not "hesitate" and it always follows the drift up to the arrival at x in a finite time, uniformly bounded in β, whereas in a general, not completely attracted domain the process tries to follow the drift in finite times as far as possible but sometimes it has to enter into suitable "permanence sets" Q_i, waiting suitable random times T_i and then getting out from Q_i through suitable optimal points. The fact that the permanence times T_i are close to the typical escape times from Q_i and that this escape takes place in the optimal way is the counterpart of the fact that in the completely attracted case the only permanence sets are trivial in the sense that they reduce to single points and the way of getting out from these single points (after a unitary permanence time!) is optimal in the sense that it is along the drift.

Problems in many respects similar to the ones studied in the present paper have been considered in the framework of the theory of simulated annealing. We refer to refs. 2, 3, 6, 1, and 14–16 for results connected to the tube of exit from a general domain.

The paper is organized as follows. In Section 2 we recall the renormalization procedure and extend previous results. In Section 3 we present the Freidlin–Wentzell (FW) graphical method by extending to a more general case their definition of cycles. In Section 4 we establish some useful properties of the cycles. In Section 5 we state and prove our main theorem, which determines typical trajectories of the original chain X_t corresponding to a single step of the renormalized chain $X_t^{(N)}$. Finally in Section 6 we use the results of previous sections to give a characterization of the tube of typical trajectories during the first excursion outside Q.

2. THE RENORMALIZATION PROCEDURE

In this section we recall the construction of the sequence of renormalized chains introduced in ref. 12 and we prove some new results.

Let $\Phi(S) \equiv \{\{\phi_i\}_{i\in\mathbf{N}}: \phi_i \in S\}$ be the set of paths. Following the theory of large deviations developed in ref. 5, we define, for each $t \in \mathbf{N}$, a functional $I_{[0,t]}$ on $\Phi(S)$ associating to each path $\phi \in \Phi(S)$ the value

$$I_{[0,t]}(\phi) \equiv \sum_{i=0}^{t-1} \Delta(\phi_i, \phi_{i+1}) \tag{2.1}$$

where for $x \neq y$, $P(x, y) > 0$, the function $\Delta(x, y)$ has been defined in (1.1) and we set $\Delta(x, x) = 0$ for each $x \in S$ and $\Delta(x, y) = \infty$ if $P(x, y) = 0$. This functional is the *cost function* of each path ϕ; we briefly recall now the main results and the construction developed in ref. 12.

Lemma 2.1. Let ϕ be a given path starting from x at time 0; then, for $t \in \mathbf{N}$, the following hold:

(i) We have

$$P_x(X_s = \phi_s \ \forall s \in [0, t]) \leqslant e^{-I_{[0,t]}(\phi)\beta + \gamma t \beta}$$

where γ is the quantity introduced in (1.1).

(ii) If ϕ is such that $\phi_s \neq \phi_{s+1}$ for any $s \in [0, t]$, then we have also a lower bound:

$$P_x(X_s = \phi_s \ \forall s \in [0, t]) \geqslant e^{-I_{[0,t]}(\phi)\beta - \gamma t \beta}$$

(iii) For any constant $I_0 > 0$, for any $\alpha > 0$ sufficiently small ($\alpha < \Delta_1$), for any $t < e^{\alpha\beta}$, and for any sufficiently large β

$$\sup_x P_x(I_{[0,t]}(X_s) \geqslant I_0) \leqslant e^{-I_0\beta + \varepsilon\beta}$$

where $\varepsilon \to 0$ as $\beta \to \infty$.

By using the functional $I_{[0,t]}(\phi)$, an equivalence relation, denoted by $\sim$, can be introduced in the state space S: for each pair of states x, y we define

$$V(x, y) \equiv \inf_{t,\phi:\ \phi_0 = x,\ \phi_t = y} I_{[0,t]}(\phi) \tag{2.2}$$

and we set

$$x \sim y \qquad \text{iff} \quad V(x, y) = V(y, x) = 0 \tag{2.3}$$

We denote by $(x)_\sim$ the equivalence class of x, i.e., $(x)_\sim \equiv \{y \in S: y \sim x\}$.

We say that x is a *stable state* if and only if

$$\text{for any} \quad y \not\sim x, \qquad V(x, y) > 0 \tag{2.4}$$

i.e., if each path leaving from $(x)_\sim$ has a positive cost. We will denote by M the set of stable states.

It is immediate to see that if the set M contains a state x, then it contains the whole equivalence class of x, namely $M \supset (x)_\sim$.

An immediate consequence of Lemma 2.1 is the following:

Lemma 2.2. (i) There exist constants $T_0 \in [0, |S|]$ and $\beta_0 > 0$ such that for any $\beta > \beta_0$ and for any $t > T_0$

$$\sup_{x \in S} P_x(\tau_M > t) \leqslant a^{[t/T_0]}$$

where $0 < a = 1 - C^{T_0}$ for some constant $0 < C < 1$ and $[\cdot]$ denotes the integer part.

(ii) For any $\eta > 0$ and for any $t \geqslant e^{\eta\beta}$ and β sufficiently large we have

$$\sup_{x \in S} P_x(\tau_M > t) \leqslant \exp\{-e^{\eta\beta/2}\}$$

This means that the process spends, with large probability, almost all the time in M. This result suggests that, if we look at the process X_t on a sufficiently large time scale, then it can be described in terms of transitions between states in M; in this way only the behavior of the process on small times is neglected.

Indeed we can consider the less stable states in M and we can define a time scale t_1 corresponding to this smallest stability:

$$t_1 \equiv e^{V_1\beta + \delta\beta} \tag{2.5}$$

where

$$V_1 \equiv \min_{x \in M,\, y \in S,\, x \nrightarrow y} V(x, y) \tag{2.6}$$

and $\delta = \delta(\beta)$ goes to zero as β tends to infinity.

We can then construct a new Markov chain $\bar{X}_t$ with state space M, corresponding to the original process with a rescaling time t_1, by defining a sequence of stopping times $\zeta_1, \zeta_2, \ldots, \zeta_n, \ldots$ such that $\zeta_{n+1} - \zeta_n$ is of order t_1 with large probability and X_{ζ_n} belongs to M.

More precisely, we define the sequence of stopping times

$$\zeta_0 \equiv \min\{t \geqslant 0: X_t \in M\}$$

and for each $n \geqslant 1$

$$\begin{aligned}
&\sigma_n \equiv \min\{t > \zeta_{n-1}: X_t \nrightarrow X_{\zeta_{n-1}}\} \\
&\tau_n \equiv \min\{t \geqslant \sigma_n: X_t \in M\} \\
&\zeta_n = \begin{cases} \zeta_{n-1} + t_1 & \text{if } \ \sigma_n - \zeta_{n-1} > t_1 \\ \tau_n & \text{if } \ \sigma_n - \zeta_{n-1} \leqslant t_1 \end{cases}
\end{aligned} \tag{2.7}$$

It is easy to see that the sequence $\bar{X}_n = X_{\zeta_n}$ is a homogeneous Markov chain. For any pair of states $x, y \in M$ we denote by $\bar{P}(x, y)$ the transition probabilities of the chain $\bar{X}_n$; it is possible to prove[12] that these transition probabilities satisfy the same assumption (property $\mathscr{P}$) satisfied by the original chain X_t, provided we identify states which are equivalent with respect to the relation (2.3). More precisely, for any $x, y \in M$, $x \nsim y$,

$$\exp\{-\Delta^{(1)}(x, y)\,\beta - \gamma'\beta\} \leqslant \bar{P}(x, y) \equiv P(X_{\zeta_n} = y \mid X_{\zeta_{n-1}} = x) \\ \leqslant \exp\{-\Delta^{(1)}(x, y)\,\beta + \gamma'\beta\} \tag{2.8}$$

The quantities $\Delta^{(1)}(x, y)$ are defined by

$$\Delta^{(1)}(x, y) = \inf_{\substack{t, \phi:\ \phi_0 = x,\ \phi_t = y, \\ \phi_s \notin M \setminus [(x)_\sim \cup (y)_\sim]}} I_{[0,t]}(\phi) - V_1 \tag{2.9}$$

and $\gamma' \to 0$ as $\beta \to \infty$.

It is easy to show that the quantities $\Delta^{(1)}(x, y)$ are invariant with respect to the equivalence relation, i.e., $\Delta^{(1)}(x, y) = \Delta^{(1)}(x', y')$ if $x \sim x'$ and $y \sim y'$. In other words, equivalent states are not distinguishable in this construction.

More precisely, let $S^{(1)} \equiv M/_\sim \equiv \{\text{equivalent classes in } M\}$ and for any $i \in S^{(1)}$ let m_i be the subset of M given by the states belonging to the equivalence class i, that is, $M = \bigcup_{i \in S^{(1)}} m_i$.

We can define a new chain $X_t^{(1)}$ on $S^{(1)}$ with transition probabilities

$$P^{(1)}(i, j) = \frac{1}{\bar{\mu}(m_i)} \sum_{x \in m_i} \bar{\mu}(x) \sum_{y \in m_j} \bar{P}(x, y)$$

where $\bar{\mu}$ denotes the invariant measure of the chain $\bar{X}_t$. This measure $\bar{\mu}$ is related to the invariant measure μ of the chain X_t (which is strictly positive by the ergodicity condition) by the following relation [see ref. 17, Eq. (4.3), p. 119]:

$$\mu(C) = \frac{1}{Z} \sum_{x \in M} \bar{\mu}(x)\, E_x \sum_{t=0}^{\zeta_1} \chi_C(X_t)$$

where $C \subset S$, χ_C is the characteristic function of C, and Z is a normalization constant.

Property $\mathscr{P}$ obviously holds also for the chain $X_t^{(1)}$ with the same function $\Delta^{(1)}$ and the invariant measure of this new chain is given by $\mu^{(1)}(i) = \bar{\mu}(m_i)$ for each $i \in S^{(1)}$.

Thus we have a new chain $X_t^{(1)}$ on the state space $S^{(1)}$, to which we can apply again the same analysis, by defining new stable states, a time scale T_2, a corresponding chain $X_t^{(2)}$, and so on.

We recall here the iteration scheme introduced in ref. 12.

Notation. The superscript (k) will denote the various quantities referring to the kth chain $X_t^{(k)}$; e.g., $\tau_Q^{(k)} = \min\{t: X_t^{(k)} \in Q\}$, $Q \in S^{(k)}$.

For any $k \geqslant 1$ we define the following quantities. For any $\phi: \mathbf{N} \to S^{(k)}$,

$$I_{[0,t]}^{(k)}(\phi) = \sum_{i=0}^{t-1} \Delta^{(k)}(\phi_i, \phi_{i+1}) \tag{2.10}$$

$$V^{(k)}(x, y) \equiv \min_{t,\phi:\, \phi_0 = x,\, \phi_t = y} I_{[0,t]}^{(k)}(\phi) \qquad \forall x, y \in S^{(k)} \tag{2.11}$$

$$x \sim^{(k)} y \qquad \text{if and only if} \quad V^{(k)}(x, y) = V^{(k)}(y, x) = 0 \tag{2.12}$$

$$M^{(k)} = \{x \in S^{(k)}: \forall y \in S^{(k)},\ y \not\sim^{(k)} x \quad V^{(k)}(x, y) > 0\} \tag{2.13}$$

$$V_{k+1} = \min_{x \in M^{(k)},\, y \in S^{(k)}\, x \not\sim^{(k)} y} V^{(k)}(x, y) \tag{2.14}$$

$$t_{k+1} = e^{V_{k+1}\beta + \delta\beta} \tag{2.15}$$

$$T_1 = t_1$$

$$T_{k+1} = t_1 t_2 \cdots t_k t_{k+1} \tag{2.16}$$

$$S^{(k+1)} = M^{(k)}/_{\sim^{(k)}} \tag{2.17}$$

$$\Delta^{(k+1)}(x, y) = \min_{\substack{t,\phi:\, \phi_0 = x,\, \phi_t = y, \\ \phi_s \notin M^{(k)} \setminus [(x)_{\sim^{(k)}} \cup (y)_{\sim^{(k)}}]\ \forall s \in [0,t]}} I_{[0,t]}^{(k)}(\phi) - V_{k+1} \quad \forall x, y \in S^{(k+1)} \tag{2.18}$$

The main results proved in ref. 12 can be summarized as follows:

Theorem 2.1. Let $W \subset S^{(1)}$ and $B \subset W$, and let $\bar{W}$ and $\bar{B}$ the corresponding sets in M ($\bar{W} = \bigcup_{i \in W} m_i$ and analogously for $\bar{B}$). Then for any sufficiently large β and for any $x \in m_i$, $i \in S^{(1)} \setminus W$, $j \in W$, there is a positive γ' depending on γ and tending to zero as $\beta \to \infty$ such that

$$\exp\{-\gamma'\beta\}\ P_i^{(1)}(X_{\tau_W}^{(1)} = j) \leqslant P_x(X_{\tau_{\bar{W}}} \in m_j) \leqslant \exp\{\gamma'\beta\}\ P_i^{(1)}(X_{\tau_W}^{(1)} = j)$$

$$\exp\{-\gamma'\beta\}\ t_1 E_i^{(1)} \tau_W^{(1)} \leqslant E_x \tau_{\bar{W}} \leqslant \exp\{\gamma'\beta\}\ t_1 E_i^{(1)} \tau_W^{(1)}$$

$$\exp\{-\gamma'\beta\}\ \mu^{(1)}(B) \leqslant \mu(\bar{B}) \leqslant \exp\{\gamma'\beta\}\ \mu^{(1)}(B)$$

for any $B \subset S^{(1)}$. Moreover, for any $A \in S \setminus M$

$$\mu(A) \leqslant \exp\{-V_1\beta + \gamma'\beta\}$$

Since $|S^{(i)}| \leqslant |S^{(i-1)}|$ (actually one can prove that $|S^{(i+1)}| < |S^{(i-1)}|$), the above results provide a useful tool for the evaluation of these quantities when $|S|$ is large; in fact one can consider a time rescaling T_n so large that the corresponding state space $S^{(n)}$ is so small that explicit computations are easy at this level.

In order to control the large-deviation phenomena, for the Markov chain X_t, taking place during exponentially long times T_k, we collect, in the rest of this section, some general and quite easily proved results related to this renormalization procedure. The main statement on the behavior of the chain X_t on exponentially long time intervals will be given in Section 5.

Let us start by the following remark: We note that in the construction of the previously recalled renormalized chains, we have to identify equivalent states at each step of the iteration. In fact, due to the time rescaling, the renormalized chain does not have enough resolution to distinguish among equivalent states. This fact implies that chains at different steps of the iteration live on different probability spaces. However, we want to define now an application between trajectories of chains at different levels of the iteration.

More precisely, for any integer n, to any path ϕ of the Markov chain X_t, $\phi \in \Phi(S)$, we want to associate a path $\phi^{(n)}$ of the Markov chain $X_t^{(n)}$, $\phi^{(n)} \in \Phi(S^{(n)})$, which is in some sense a projection of the path ϕ in the smaller space $\Phi(S^{(n)})$. On the other hand, to any path $\phi^{(n)} \in \Phi(S^{(n)})$ we want to associate a set, say a tube, of paths $\phi \in \Phi(S)$ having projection $\phi^{(n)}$.

Definition 2.1. For each path $\phi \in \Phi(S)$ we evaluate the sequence of stopping times ζ_n and we define a path $\bar{\phi} \in \Phi(M)$ given by $\bar{\phi}_i = \phi_{\zeta_i}$. To each path $\bar{\phi} \in \Phi(M)$ we can obviously associate a path $\phi^{(1)} \in \Phi(S^{(1)})$ by defining $\phi_s^{(1)} = i$ if $\bar{\phi}_s \in m_i$.

Using the same construction, we can thus define a sequence of trajectories $\{\phi_i^{(n)}\}_{i \in \mathbb{N}}$ in the spaces $\Phi(S^{(n)})$ with $n = 2, 3,\ldots$.

On the other hand, to any given sequence of states in $S^{(n)}$: $\psi_i^{(n)}$, $i \leqslant T$, we can associate a tube of trajectories in $\Phi(S)$ as

$$\mathscr{T}(\psi^{(n)}, T) = \{\phi \in \Phi(S) : \phi_i^{(n)} = \psi_i^{(n)}, \forall i \leqslant T\} \tag{2.19}$$

By construction this application between trajectories in $\Phi(S)$ and trajectories in $\Phi(S^{(n)})$ is such that

$$P^{(n)}(X_s^{(n)} = \phi_s^{(n)}, \forall s \leqslant T) \asymp P(X_t \in \mathscr{T}(\psi^{(n)}, T)) \tag{2.20}$$

where we say that two quantities A and B are *logarithmically equivalent* and we write $A \asymp B$ if they have the same exponential asymptotic behavior in β, namely

$$A \asymp B \qquad \text{if and only if} \quad \lim_{\beta \to \infty} \frac{1}{\beta} \ln A = \lim_{\beta \to \infty} \frac{1}{\beta} \ln B \tag{2.21}$$

We say that A is *logarithmically greater* than B and we write $A \succcurlyeq B$ if and only if

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln A \geqslant \lim_{\beta \to \infty} \frac{1}{\beta} \ln B \tag{2.22}$$

Definition 2.2. With this application we can also define a sequence of random times Z_k^n corresponding, on the original time scale, to the times k on the time scale of the chain $X^{(n)}$. More precisely, given a path $\phi \in \Phi(S)$, we have defined a sequence of times $\zeta_0,..., \zeta_k,...$ and a path $\phi^{(1)} \in \Phi(S^{(1)})$ associated to it, and, iterating, sequences of times $\zeta_0^{(i)},..., \zeta_k^{(i)},...$ and paths $\phi^{(i+1)} \in \Phi(S^{(i+1)})$ for each $i = 0, 1,...$.

We define

$$Z_k^n \equiv \zeta_{\zeta^{(1)}_{\zeta^{(2)}_{\ddots_{\zeta_k^{(n-1)}}}}}, \qquad n = 1, 2,... \tag{2.23}$$

which is a random time with respect to the process X_t.

Definition 2.3. For each $x \in S^{(n)}$ we can define a set $\mathscr{E}_x^n$ belonging to the original state space S, obtained by associating, at each step of the iteration ($\leqslant n$), to each equivalence class the set of its elements. More precisely, if $n = 1$,

$$\mathscr{E}_x^1 = m_x \tag{2.24}$$

and, for $j = 1,..., n$, if $x \in S^{(j)}$ and $m_x^{(j-1)}$ are the elements of $S^{(j-1)}$ corresponding to the equivalence class x, then we set

$$\mathscr{E}_x^j \equiv \bigcup_{x' \in m_x^{(j-1)}} \mathscr{E}_{x'}^{j-1} \tag{2.25}$$

From (2.25) we construct iteratively $\mathscr{E}_x^n \subset S$.

If we define

$$\bar{S}^{(n)} = \bigcup_{x \in S^{(n)}} \mathscr{E}_x^n \tag{2.26}$$

we have immediately

$$\bar{S}^{(1)} \equiv M, \qquad \bar{S}^{(i)} \subseteq \bar{S}^{(i-1)} \tag{2.27}$$

and by iteration it is easy to see that

$$X_{Z_k^n} \in \bar{S}^{(n)}, \qquad \forall n = 0, 1, \dots \tag{2.28}$$

and

$$\tau_{S^{(n)}} = Z_0^n \tag{2.29}$$

Indeed (2.29) immediately follows from a more general result:

Lemma 2.3. For any $A \in S^{(n)}$ let $\mathscr{E}_A^n \equiv \bigcup_{x \in A} \mathscr{E}_x^n$; then

$$\tau_{\mathscr{E}_A^n} = Z_{\tau_A^{(n)}}^n$$

Proof. The straightforward proof comes immediately from definitions (2.24)–(2.26). ∎

The results stated in Theorem 2.1 allow a control on the expectation of hitting times. We give here some results which enable us to control also in probability these random times.

We first notice that for any set $Q \subset S$ we can define a chain X_t^Q with almost absorbing states in Q^c as follows:

$$\begin{aligned} P^Q(x, y) &= P(x, y) && \text{if} \quad x \in Q \\ P^Q(x, y) &= P(x, y)\, e^{-\beta \Delta(Q)} && \text{if} \quad x \in Q^c, \quad x \neq y \end{aligned} \tag{2.30}$$

where $\Delta(Q) > \sum_{y,z \in Q} \Delta(y, z)$ and $P^Q(x, x)$ is defined by normalization.

As far as the first exit from the domain Q is concerned, the two chains X_t^Q and X_t are completely equivalent (the superscript Q denotes all the quantities related to the chain X_t^Q).

Proposition 2.1. For any $Q \subset S$ and any integer n let $(\bar{S}^{(n)})^Q$ denote the state space of the nth renormalization of the above-defined chain X_t^Q; let

$$N = N(Q) \equiv \max\{N\colon Q \cap (\bar{S}^{(N)})^Q \neq \emptyset\} \tag{2.31}$$

Then: (i) For any $x \in Q \cap (\bar{S}^{(N)})^Q$

$$E_x \tau_{Q^c} \asymp T_N$$

[see (2.21) for the definition of $\asymp$].

(ii) For any $\alpha > 0$ there exists $k = k(\alpha) > 0$ such that for any $x \in Q \cap (\bar{S}^{(N)})^Q$

$$P_x(T_N e^{-\alpha\beta} \leqslant \tau_{Q^c} \leqslant T_N e^{\alpha\beta}) \geqslant 1 - e^{-k\beta} \tag{2.32}$$

Proof. Point (i) is an immediate consequence of Theorem 2.1. Indeed, since the events considered in this proposition belong to the σ-algebra $\mathscr{F}_{\tau_{Q^c}}$ generated by the events depending on the process up to time τ_{Q^c}, we can consider the chain X_t^Q instead of the chain X_t, and thus $Q^c \subset (\bar{S}^{(N+1)})^Q$. On the other hand, since $Q \cap (\bar{S}^{(N+1)})^Q = \varnothing$, this implies $Q^c = (\bar{S}^{(N+1)})^Q$. For notational convenience we will omit from now on in this proof the superscript Q: only the chain X_t^Q is considered here; as we noticed above, the same estimates hold for the chain X_t.

By iterating N times Theorem 2.1, we obtain

$$E_x \tau_{Q^c} \asymp t_1 \cdot t_2 \cdots t_N \cdot E_x^{(N)} \tau_{Q^c}^{(N)} \asymp T_N$$

since $Q \cap \bar{S}^{(N+1)} = \varnothing$ and thus $E_x^{(N)} \tau_{Q^c}^{(N)} \asymp 1$.

To prove (ii) we first notice that by using Chebyshev's inequality we have immediately that for every $\alpha > 0$ there exists $k > 0$ such that

$$P_x(\tau_{Q^c} > T_N e^{\alpha\beta}) \leqslant e^{-k\beta}$$

So we need only to prove that $\forall \alpha > 0$

$$P_x(\tau_{Q^c} < T_N e^{-\alpha\beta}) \leqslant e^{-k\beta} \qquad \text{for some} \quad k > 0 \tag{2.33}$$

Since $Q^c = \bar{S}^{(N+1)}$, by (2.29) the proposition is thus proved if (2.33) holds when we replace τ_{Q^c} by

$$Z_0^{N+1} = Z_{\zeta_0^{(N)}}^{N}$$

Since $x \in \bar{S}^{(N)} \backslash \bar{S}^{(N+1)}$, we have $\zeta_0^{(N)} \geqslant 1$ and thus the proposition is proved if we prove the following more general result:

Lemma 2.4. For any $N \geqslant 1$ and for any sufficiently small positive constant α (i.e., such that $t_n e^{-\alpha\beta} > 1$ for all n) there exists a positive constant $k = k(\alpha, N)$ such that for any $x \in \bar{S}^{(N)}$

$$P_x(T_N e^{-\alpha\beta} \leqslant Z_1^N \leqslant T_N e^{\alpha\beta}) \geqslant 1 - e^{-k\beta}$$

for any β sufficiently large.

Remark. In the application of this lemma we are not interested in making the best choice of the constant $k(\alpha, N)$. We will prove the lemma with the crude choice $k(\alpha, N) = \alpha/2^N$.

Proof of Lemma 2.4. The proof is by induction. We first prove the initial step:

$$P_x(t_1 e^{-\alpha\beta} \leqslant \zeta_1 \leqslant t_1 e^{\alpha\beta}) \geqslant 1 - e^{-k\beta}$$

By applying Lemma 2.1 we have immediately

$$\begin{aligned}
&P_x(\zeta_1 \leqslant t_1 e^{-\alpha\beta}) \\
&\quad \leqslant P_x(\sigma_1 < t_1 e^{-\alpha\beta}) \\
&\quad \leqslant P_x(\exists t < t_1 e^{-\alpha\beta}\colon \Delta(X_t, X_{t+1}) \geqslant V_1) \\
&\quad \leqslant t_1 e^{-\alpha\beta} e^{-V_1\beta + \gamma\beta} = e^{-(\alpha - \delta - \gamma)\beta}
\end{aligned}$$

and by Lemma 2.2 the probability $P_x(\zeta_1 \geqslant t_1 e^{+\alpha\beta})$ is superexponentially small. So, for β sufficiently large, $\delta + \gamma < \alpha$, the validity of the first step of the induction is proved with $k(\alpha, 1) = \alpha - \delta - \gamma$.

Now assuming that there exists $k = k(\alpha, n-1)$ such that

$$P_x(T_{n-1} e^{-\alpha\beta} \leqslant Z_1^{n-1} \leqslant T_{n-1} e^{\alpha\beta}) \geqslant 1 - e^{-k(\alpha, n-1)\beta} \tag{2.34}$$

we want to prove the same as $n-1$ is replaced by n. We use the fact that by definition the time

$$Z_1^n = Z^{n-1}_{\zeta_1^{(n-1)}}$$

where $\zeta_1^{(n-1)}$ is a random time defined like ζ_1 for the chain $X^{(n-1)}$. In other words, Z_1^n is a sum of $\zeta_1^{(n-1)}$ random intervals of time distributed like Z_1^{n-1}. We have

$$\begin{aligned}
&P_x(Z_1^n < T_n e^{-\alpha\beta}) \\
&\quad \leqslant P_x(\zeta_1^{(n-1)} < t_n e^{-\alpha\beta/2}) \\
&\qquad + P_x(\{\zeta_1^{(n-1)} \geqslant t_n e^{-\alpha\beta/2}\} \\
&\qquad\qquad \cap \{\text{there are at least } \tfrac{1}{2}\zeta_1^{(n-1)} \text{ intervals} \\
&\qquad\qquad \text{of length } < 2 \cdot T_{n-1} e^{\alpha\beta/2}\})
\end{aligned}$$

By using (2.34) at step $n-1$ and the strong Markov property, we can estimate the last probability by means of a Bernoulli distribution $\xi_i = 0, 1$ with

$$\begin{aligned}
p &\equiv P(\xi = +1) = 1 - P(\xi = 0) \\
&= \max_{x \in \bar{S}^{(n)}} P_x(Z_1^{n-1} < 2 \cdot T_{n-1} e^{\alpha\beta/2}) \leqslant e^{-k(\alpha/2, n-1) - \varepsilon\beta}
\end{aligned}$$

with $\varepsilon \to 0$ as $\beta \to \infty$. Indeed,

$$P\left(\xi_1 + \cdots + \xi_l - lp > l\left(\frac{1}{2} - p\right)\right) \leqslant \frac{lp(1-p)}{l^2(1/2-p)^2}$$

and if α is sufficiently small (but independent of β) and $l \geqslant t_n e^{-\alpha\beta/2}$, we obtain the estimate

$$P_x(Z_1^n < T_n e^{-\alpha\beta}) \leqslant e^{-k\beta}$$

with $k = \alpha/2^n$.

We can analogously estimate the probability $P_x(Z_1^n > T_n e^{+\alpha\beta})$. ∎

We conclude this section with a result referring to the existence of a set of trajectories in $\Phi(S)$ associated to a given path in $\Phi(S^{(n)})$ on which the probability is concentrated.

Proposition 2.2. Given $\phi^{(n)} \in \Phi(S^{(n)})$ and a time T (independent of β), let $\mathscr{T}_0(\phi^{(n)}, T) \subset \mathscr{T}(\phi^{(n)}, T)$ be defined by

$$\mathscr{T}_0(\phi^{(n)}, T) \equiv \{\phi \in \mathscr{T}(\phi^{(n)}, T)\colon T_i e^{-\alpha\beta} \leqslant Z_k^i \leqslant T_i e^{\alpha\beta}\ \forall i, k$$
$$\text{such that } Z_k^i \leqslant Z_T^n \text{ and } \forall\alpha \text{ sufficiently small}\}$$

Then for any $\phi \in \mathscr{T}(\phi^{(n)}, T) \backslash \mathscr{T}_0(\phi^{(n)}, T)$ we have

$$P(X_t = \phi, \forall t \leqslant T) \leqslant e^{-k\beta}$$

for some positive k independent of β.

Proof. The proof of this proposition is an easy consequence of Lemma 2.4. ∎

3. GRAPHS AND CYCLES

In this section we recall the main results on the first exit problem proved by Freidlin and Wentzell[5] and we give a more general definition of cycles. We note that the results in ref. 5 are mostly formulated in the continuous case, i.e., for the study of a diffusion process defined by a stochastic differential equation in $\mathbf{R}^d$:

$$dX_t = b(X_t)\,dt + \varepsilon\,dW_t \tag{3.1}$$

where W_t is the d-dimensional Wiener process at time t and ε is very small ($\varepsilon^2 = 1/\beta$). The drift term $b(\cdot)$ is such that the deterministic flow

$$dX_t = b(X_t)\,dt \tag{3.2}$$

has ω-limit sets contained in a finite number of compact sets $K_1, K_2,..., K_l$.

We summarize here their results in the simpler case of a Markov chain satisfying the ergodicity property and property $\mathscr{P}$ given in the introduction.

Their analysis is based on the following:

Definition 3.1 (*B-graphs*). For any set of states $B \subset S$, a B-graph is a graph consisting of arrows $m \to n$ with $m \in B$ and $n \in S$, $m \neq n$, and satisfying the following properties:

1. Every state $m \in B$ is the initial point of exactly one arrow.
2. There are no closed cycles in the graph.

Condition 2 can be replaced by:

2′. For any state $m \in B$ there exists a sequence of arrows leading from it to some point $n \in B^c$.

In other words, a graph is a forest of trees with roots in B^c and with branches given by arrows directed to the root (i.e., the set B^c is the target of the sequences of arrows).

The set of B-graphs is denoted by $G(B)$; for any graph $g \in G(B)$ we define $\pi(g) = \prod_{m \to n \in g} P(m, n)$.

By using this graphic formalism, Freidlin and Wentzel[5] prove the following lemmas.

Warning. Our notation is opposite that used in ref. 5, where a W-graph was a graph with target W, i.e., a W^c-graph in our notation.

Lemma 3.1 (FW). The invariant measure of the Markov chain X_t: $\mu(i)$, $i \in S$, is given by

$$\mu(i) = \frac{q_i}{\sum_{j \in S} q_j}$$

where

$$q_i = \sum_{g \in G\{S \backslash i\}} \pi(g) \tag{3.3}$$

Lemma 3.2 (FW). For any $B \subset S$ let τ_B be the first hitting time to B; then for any $j \in B$

$$P_i(X_{\tau_B} = j) = \frac{\sum_{g \in G_{ij}(B^c)} \pi(g)}{\sum_{g \in G(B^c)} \pi(g)} \tag{3.4}$$

where for any $i \in B^c$ and $j \in B$ we denote by $G_{ij}(B^c)$ the set of B^c-graphs in which the sequence of arrows leading from i into B ends at the point j (i.e., i belongs to the tree with root j).

Lemma 3.3 (FW). We have

$$\begin{aligned} E_i \tau_B &= \frac{\sum_{g \in G(B^c \setminus \{i\})} \pi(g) + \sum_{j \in B^c, j \neq i} \sum_{g \in G_{ij}(B \setminus \{j\})} \pi(g)}{\sum_{g \in G(B^c)} \pi(g)} \\ &= \frac{\sum_{g \in G(i \not\to B)} \pi(g)}{\sum_{g \in G(B^c)} \pi(g)} \end{aligned} \tag{3.5}$$

where $G(i \not\to B)$ is the set of graphs (without closed loops) containing $|B^c| - 1$ arrows $m \to n$ each one emerging from a different point $m \in B^c$ and with $n \in S$, $m \neq n$, and not containing chains of arrows leading from i into B.

Lemma 3.1 is easily proved by showing that the quantities q_i satisfy the stationarity equation; Lemmas 3.2 and 3.3 can be proved by induction over the number of states contained in W^c (see ref. 5, pp. 179, 182).

By using property $\mathscr{P}$, these results can be reformulated as follows.[5] Consider the problem of the first exit of our chain X_t from a domain $Q \subset S$ and let τ_{Q^c} be the first exit time from the domain Q.

Proposition 3.1 (FW). For any $\delta > 0$ and for any sufficiently large β there exists $\delta > 0$ such that

$$e^{-\beta(W_Q(x,y) - W_Q + \delta)} \leqslant P_x(X_{\tau_{Q^c}} = y) \leqslant e^{-\beta(W_Q(x,y) - W_Q - \delta)} \tag{3.6}$$

for any $x \in Q$ and $y \in Q^c$, with

$$W_Q(x, y) = \min_{g \in G_{xy}(Q)} W(g) \tag{3.7}$$

$$W_Q = \min_{g \in G(Q)} W(g) \tag{3.8}$$

$$W(g) \equiv \sum_{m \to n \in g} \Delta(m, n) \tag{3.9}$$

and $\delta \to 0$ as $\beta \to \infty$.

Proposition 3.2 (FW). For any $x \in Q$ let Y_x be the set of states in Q^c such that there exists at least a graph minimizing (3.8) and containing the sequence of arrows $x \to \cdots \to y$. Then with probability converging to one as β tends to infinity the first exit from the domain Q of the process starting at x take place in Y_x. (See ref. 5, Theorem 5.2, Chapter 6.)

Proposition 3.3 (FW). For any $x \in Q$

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln E_x \tau_{Q^c} = W_Q - M_Q(x) \tag{3.10}$$

where

$$M_Q(x) = \min_{g \in G(x \nrightarrow Q^c)} W(g) \tag{3.11}$$

Let us now suppose that the set Q contains a unique stable state x_0 completely attracting this set, i.e., for each $y \in Q$ there exists a path $y_0 = y$, $y_1, \ldots, y_n = x_0$ such that $\Delta(y_i, y_{i+1}) = 0$, $\forall i < n$, while $\Delta(y, z) > 0$ for each $y \in Q$, $z \in Q^c$. Then in this case Freidlin and Wentzell can describe in complete detail the exit from Q.

First, in this case the quantities (3.4) and (3.5) can be easily estimated as follows (ref. 5, Theorems 2.1 and 4.1, Chapter 4): for all $x \in Q$

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln E_x \tau_{Q^c} = \min_{y \in Q^c} V(x_0, y)$$

where $V(x_0, y)$ is defined by (2.2), and if there exists a unique state $y_0 \in Q^c$ such that $V(x_0, y_0) = \min_{y \in Q^c} V(x_0, y)$, then

$$\lim_{\beta \to \infty} P_x(X_{\tau_{Q^c}} = y_0) = 1$$

Moreover, in this case of a domain Q containing a unique stable state, the last escape is described quite precisely and completely in ref. 5.

We state now their result in our discrete case of Markov chains (see ref. 5, Theorem 2.3, Chapter 4 for the continuous version of this result).

We want to notice here that in the continuous case of diffusion processes discussed in ref. 5 the dynamics corresponding to zero random noise was given by a dynamical system, that is, the unperturbed system was completely deterministic and for each starting point there was a unique deterministic path emerging from it. The tube of typical exiting trajectories was given, in that case, as a neighborhood, in the uniform topology, of such a deterministic path.

Here, in our present case with finite state space, the situation is different and, even for $\beta=\infty$ the system can still be random. This means that there is not a unique deterministic path, but several possible paths emerging from the same starting point. Moreover, we do not have to consider a neighborhood, since the space is discrete. So the typical exiting tube in this case is a finite set of individual paths.

Proposition 3.4. Let Q be a set of states containing a unique stable state x_0, and for each given α and β define

$$\Phi_{\alpha\beta} \equiv \{\{\phi_s\}_{s\in\mathbb{N}}: \exists T_\phi < e^{\alpha\beta}, \phi_0 = x_0, \phi_{T_\phi} \in Q^c, \phi_s \in Q, \forall s < T_\phi\} \quad (3.12)$$

and

$$\bar{\Phi}_{\alpha\beta} \equiv \{\phi_s \in \Phi_{\alpha\beta}: I_{[0,T_\phi]}(\phi) = \min_{y\in Q^c} V(x_0, y)\} \quad (3.13)$$

Then there exists α_0 (see Lemmas 2.1, 2.2) such that if $\alpha<\alpha_0$ we have

$$\lim_{\beta\to\infty} P_x(X_{\theta_{x_0}+t} = \phi_t, \forall t = 0,..., \tau_{Q^c} - \theta_{x_0} \text{ for some } \phi \in \bar{\Phi}_{\alpha\beta}) = 1 \quad (3.14)$$

where $\theta_{x_0} \equiv \max\{t < \tau_{Q^c}: X_t = x_0\}$.

Proof. A proof of this proposition can be easily obtained by applying Lemmas 2.1 and 2.2 (see ref. 11, Proposition 2.1, for a complete proof). ∎

For each $Q\subset S$ let $g_Q^* \in G(Q)$ be a Q-graph minimizing the quantity W_Q defined in (3.8).

Freidlin and Wentzel, assuming that g_Q^* is unique for every $Q\subset S$, introduced the definition of a sequence of sets, called cycles, which are useful to describe the behavior of the process X_t in time intervals exponentially long in β.

However, this uniqueness hypothesis is very restrictive. Indeed, if we assume it, for instance, in the case of a reversible Metropolis dynamics, we end up with a trivial problem in the sense that there is a unique minimum of the energy function.

We will drop the uniqueness hypothesis and so we will need a more general notion of cycle. In the literature, especially in connection with simulated annealing, extensions of Freidlin–Wentzel cycles have been introduced. See, for instance, refs. 6, 2, 14, and 15 (and references therein), where the basic results on cycles have been proved in the general case. In what follows, to be self-contained, we will give the basic definitions on cycles and discuss their main properties; for the proofs we refer to refs. 2, 3, and 14.[3]

[3] See also the extended version of the present work in the Texas archive mp-arc 95-423.

Definition 3.3 (*General Cycles*). Given a set of states Ω and a graph $\mathscr{G}$ of arrows connecting pairs of states with at least one arrow emerging from each state, we define the following partition of the space Ω.

For any given state $x_0 \in \Omega$ we define the set of its descendants as follows:

$$A_{x_0} \equiv \{x_0\} \cup \{x \in \Omega\colon \exists \text{ a sequence of arrows contained in } \mathscr{G}\colon x_0 \to x_1 \to \cdots \to x\} \tag{3.15}$$

By definition, the graph $\mathscr{G}$ has no arrows exiting from the set A_{x_0} (possibly $A_{x_0} = \Omega$) and $A_x \subseteq A_{x_0}$ for any $x \in A_{x_0}$. We say that A satisfies the cyclic property if

$$A_x = A \qquad \text{for any} \quad x \in A \tag{3.16}$$

If A_{x_0} satisfies property (3.16), we call it a cycle; otherwise its decomposition into cycles goes as follows. We call the singleton $\{x_0\}$ a cycle. Moreover, let $x_1 \in A_{x_0}$ be such that the set of its descendants A_{x_1} is strictly contained in A_{x_0}. Such an x_1 exists if A_{x_0} is not a cycle. If the set A_{x_1} satisfies property (3.16), then it is a cycle. If A_{x_1} is not a cycle, we define $\{x_1\}$ to be a cycle and we choose $x_2 \in A_{x_1}$ such that the set of its descendants A_{x_2} is strictly contained in A_{x_1}. It is easy to prove that this procedure stops at a certain x_n such that A_{x_n} is a cycle. In fact $|A_{x_i}| < |A_{x_{i-1}}|$ and for any set of descendants A we have $|A| \geqslant 2$.

We have now to start again this procedure for all the states which have not been touched by the previous construction, i.e., outside the set $\{x_0\} \cup \{x_1\} \cup \cdots \cup \{x_{n-1}\} \cup A_{x_n}$.

In this way we obtain a partition of Ω into cycles.

Now we will apply this abstract definition of cycles to our case. We consider first $\Omega = S$ and we define a graph $\mathscr{G}$ as follows. For any $Q \subset S$ let $R_Q(\cdot)$ be a function from Q to the parts of $S \backslash Q$ defined as follows:

$$R_Q(x) \equiv \{y \in Q^c\colon \text{there exists a graph } g_Q^* \text{ minimizing } W_Q \text{ and a chain of arrows } x \to \cdots \to y \in g_Q^*\} \tag{3.17}$$

Given a pair of states $x_i, x_j \in S$ we say that x_j is a *successor* of x_i, and we write: $x_i \to x_j$ iff $x_j \in R_{\{x_i\}}(x_i)$.

We obtain in this way a graph $\mathscr{G}^0$.

We want to stress that this graph $\mathscr{G}^0$ is not, in general, a B-graph satisfying Definition 3.1, and generally several arrows emerge from a single state in $\mathscr{G}^0$.

Cycles of rank zero: The 0-cycles are defined as the single states. We denote by $\mathscr{C}^0 = S$ the set of 0-cycles.

Cycles of rank one: If $\Omega = S$ and $\mathscr{G} = \mathscr{G}^0$, the previously defined cycles are called 1-cycles. We denote by $\mathscr{C}^1$ the set of 1-cycles.

Cycles of rank k: We procede by iteration. We consider as states the $(k-1)$-cycles, i.e., at each $(k-1)$-cycle C_i^{k-1} we associate a point and $C_i^{k-1} \to C_j^{k-1}$ iff $R_{C_i^{k-1}}(x) \in C_j^{k-1}$ for all $x \in C_i^{k-1}$. As we show in the next section, dedicated to cycle properties, actually the set $R_{C_i^{k-1}}(x)$ does not depend on x, and so the definition of successor is well posed.

In this way we define a graph $\mathscr{G}^{k-1}$ of arrows between $(k-1)$-cycles. The cycles in the case $\Omega = \mathscr{C}^{k-1}$ and $\mathscr{G} = \mathscr{G}^{k-1}$ are called k-cycles.

Remarks. At each step a k-cycle, with $k \geqslant 1$, turns out to be either a $(k-1)$-cycle or a union of $(k-1)$-cycles which is the minimal descendant set, where minimal is respect to the relation (3.16).

This means that each $(k-1)$-cycle contained in C^k is a descendant of each other $(k-1)$-cycle contained in the same C^k.

We note that for each k the set of cycles of rank k define a partition of the state space S.

We conclude this section with a final remark. The rank of the cycles, here as in the definition by Freidlin and Wentzell, under the uniqueness hypothesis, is a parameter necessary to the construction, but it does not have an intrinsic meaning. The iteration of this cycle construction is completely different from the renormalization procedure, where the iterative parameter has an immediate interpretation in terms of time rescaling.

4. PROPERTIES OF THE CYCLES

We collect in this section the main properties of the cycles.

Proposition 4.1. If C and C' are respectively a k-cycle and a k'-cycle, then

$$C \cap C' \neq \varnothing \qquad \text{implies} \quad C \subseteq C' \quad \text{or} \quad C \supseteq C' \tag{4.1}$$

For any set $B \subset S$ we will denote, as before, a B-graph minimizing W_B by g_B^* [see (3.8)].

Moreover, we recall that the cycles of rank 0 are identified with single states and we set $g_\varnothing^* = \varnothing$.

Proposition 4.2. For any k-cycle C^k, with $k \geqslant 1$, we have:

(a) Given $x \in C^k$ for any $z \in R^{C^k}(x)$, and for any C^k-graph $g^*_{C^k}$ with an arrow ending in z, there exists $y \in C^k$ such that

$$g^*_{C^k} = g^*_{C^k \setminus \{y\}} \cup (y \to z) \tag{4.2}$$

(b) For any set $A \neq \emptyset$, $A \subset C^k$, each $(C^k \setminus A)$-graph $g^*_{C^k \setminus A}$ minimizing $W_{C^k \setminus A}$ does not have arrows exiting from C^k (then at least one of its arrows ends in A) and can be written in the form

$$g^*_{C^k \setminus A} = \bigcup_{m:\, C^{k-1}_m \subseteq C^k} g^*_{C^{k-1}_m \setminus A} \tag{4.3}$$

(c) Each C^k-graph $g^*_{C^k}$ minimizing W_{C^k} has a unique exiting arrow and this implies, in particular, that the successor set of the cycle $R_{C^k}(x)$ does not depend on x, i.e., $R_{C^k}(x) =: R_{C^k}$.

Proposition 4.3. For any set C with $|C| > 1$, the following are equivalent:

(i) C is a cycle.

(ii) For any cycle C' contained in C the successor set [see (3.17)] satisfies

$$R_{C'} \subset C$$

(iii) There exists $K > 0$ such that for every $x \in C$, $B \subset C$, $x \notin B$, and β sufficiently large we have

$$P_x(X_{\tau_{C^c \cup B}} \notin B) < e^{-K\beta} \tag{4.4}$$

(iv) $E_x \tau_{C^c}$ is independent of $x \in C$ in the sense of logarithmic equivalence and for any $B \subset C$ and for any $x \in B$ and $z \in C$

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln E_x \tau_{B^c} < \lim_{\beta \to \infty} \frac{1}{\beta} \ln E_z \tau_{C^c} \tag{4.5}$$

Remark. Statement (iii) describes what we call the *recurrence property*, which is one of the main features of cycles.

We note that from (iv) of this proposition and from Proposition 4.1 the quantities $V_{C_i} := \lim_{\beta \to \infty}(1/\beta) \ln E_x \tau_{C_i^c}$ provide a natural ordering for cycles C_i containing a given state x_0 alternative to the rank.

As proved in ref. 11 for the reversible case, a consequence of the recurrence property of cycles is the exponential distribution of their first exit times.

Proposition 4.4. If C is a cycle, then for any $x, y \in C$, $\delta > 0$, we have

$$\lim_{\beta \to \infty} P_x(E_y \tau_{C^c} e^{-\delta\beta} \leqslant \tau_{C^c} \leqslant E_y \tau_{C^c} e^{+\delta\beta}) = 1 \tag{4.6}$$

and the probability distribution of τ_{C^c}, when suitably renormalized, is asymptotically exponential as $\beta \to \infty$. More precisely, let T_β be such that

$$\sup_{x \in C} P_x(\tau_{C^c} > T_\beta) = e^{-1} \tag{4.7}$$

Then, $\forall x \in C$

$$\lim_{\beta \to \infty} \frac{E_x \tau_{C^c}}{T_\beta} = 1 \tag{4.8}$$

and $\forall x \in C$, $\forall s \in \mathbf{R}^+$

$$\lim_{\beta \to \infty} P_x\left(\frac{\tau_{C^c}}{T_\beta} > s\right) = e^{-s} \tag{4.9}$$

We conclude this section by giving some properties of cycles connecting them to the renormalization procedure of Section 2.

More precisely, the aim of the remaining part of this section is to associate to each state $x^n \in S^{(n)}$ a cycle C_{x_n} for the original chain X_t representing the region of the state space S corresponding to the renormalized state x^n under the time rescaling T_n. The precise statement of this correspondence will the subject of the next section (see in particular Theorem 5.1).

First of all we will define the *generalized basin of attraction* of each state and we will prove their main properties. Finally, by using these properties, we will be able to define the particular cycles C_{x^n} $\forall x^n \in S^{(n)}$.

We recall, once again, that the superscript (n) denote quantities related to the chain $X^{(n)}$. We will denote by x^n, y^n elements of the space $S^{(n)}$ and by $m_{x^n}^{(n-1)}$ the set of equivalent states in $M^{(n-1)} \subseteq S^{(n-1)}$ (with respect to the equivalence relation $\sim^{(n-1)}$) corresponding to the equivalence class of x^n.

Definition 4.1. For any $n \geqslant 1$ and for any $x^n \in S^{(n)}$ we define the *basin of attraction* of x^n as the subset of $S^{(n-1)}$ given by

$$B_{x^n}^{(n-1)} := \{z^{n-1} \in S^{(n-1)}: \exists y^{n-1} \in m_{x^n}^{(n-1)}: V^{(n-1)}(z^{n-1}, y^{n-1}) = 0\}$$

Definition 4.2. We say that a state $x \in S$ *is connected by a steep path* to $x^n \in S^{(n)}$ if and only if there exists a sequence of states of increasing stability $y^i \in S^{(i)}$, $i = 0,..., n$, such that $y^0 = x$, $y^n = x^n$, and

$$y^i \in B_{y^{i+1}}^{(i)} \qquad \forall i = 1,..., n-1 \tag{4.10}$$

Definition 4.3. Given $x^n \in S^{(n)}$, we define the *generalized basin of attraction* (GBA) of x^n at level n as the subset of S given by

$$\bar{\mathscr{B}}_{x^n}^n := \{x \in S \text{ connected by a steep path to } x^n\}$$

We can also define a GBA for a set of equivalent states (i.e., a "plateau") instead of a single state.

In the following we will often use the short notation $p^{(n)}$ to denote a plateau $m_{x^{n+1}}^{(n)}$ for some x^{n+1}. We recall that $p^{(n)}$ is made up of stable points of $S^{(n)}$: $p^{(n)} = (x^{n+1})_{\sim n}$ with $(x^{n+1} \in M^{(n)})$.

We define

$$\bar{\mathscr{B}}_{p^{(n)}}^n = \bigcup_{y^n \in p^{(n)}} \bar{\mathscr{B}}_{y^n}^n$$

Remarks. It is immediate to verify that

$$\bar{\mathscr{B}}_{x^n}^n = \bigcup_{x_l^{n-1} \in B_{x^n}^{(n-1)}} \bar{\mathscr{B}}_{x_l^{n-1}}^{n-1}$$

It is also immediate to verify that for each $x \in \mathscr{E}_{x^n}^n$ we have $x \in \bar{\mathscr{B}}_{x^n}^n$ and that for any $y \in \mathscr{E}_{y^n}^n$ with $y^n \neq x^n$ we have $y \notin \bar{\mathscr{B}}_{x^n}^n$.

On the other hand, we notice that for any $n \geqslant 1$ the GBAs $\bar{\mathscr{B}}_{x^n}^n$ define a covering of the space S for $x^n \in S^{(n)}$. Different GBAs, say $\bar{\mathscr{B}}_{x_1^n}^n$, $\bar{\mathscr{B}}_{x_2^n}^n$, with x_1^n, $x_2^n \in S^{(n)}$, can partially overlap since there may exist "saddle" states decaying to different equilibrium states.

We used a superscript n in the definition of the GBA to emphasize that it depends on the index n. More precisely, if a state belongs to two different state spaces $x \in S^{(n)}$ and $x \in S^{(n+1)}$, then we have $\mathscr{B}_x^n \subseteq \mathscr{B}_x^{n+1}$.

Definition 4.4. We will define the *strict generalized basin of attraction* (SBGA) of x^n, the set given by

$$\mathscr{B}^n_{x^n} := \left(\bigcup_{y^n \in S^{(n)},\, y^n \neq x^n} \bar{\mathscr{B}}^n_{y^n} \right)^c$$

For any plateau $p^{(n)} \in S^{(n)}$ we analogously define

$$\mathscr{B}^n_{p^{(n)}} := \left(\bigcup_{y^n \in S^{(n)} \setminus p^{(n)}} \bar{\mathscr{B}}^n_{y^n} \right)^c$$

Definition 4.5. For any $x \in \bar{\mathscr{B}}^n_{x^n}$, we say that the process X_t, starting from x, *falls to the bottom* x^n if the following event takes place: there exists a sequence $y^0, y^1,\ldots$ satisfying Eq. (4.10) with $y^0 = x$, $y^n = x^n$ such that

$$X_{\tau_{\bar{S}^{(i)}}} \in \mathscr{E}^i_{y^i} \qquad \forall i \leqslant n$$

One can show that with large probability, starting from any state x, the process first reaches the bottom of its SGBA. More precisely:

Proposition 4.5. Let $x^n \in S^{(n)}$; for each $x \in \mathscr{B}^n_{x^n}$ we have

$$P_x(\text{the process falls to the bottom } x^n) \geqslant 1 - e^{-K\beta} \tag{4.11}$$

which implies

$$P_x(X_{\tau_{\bar{S}^{(n)}}} \in \mathscr{E}^n_{x^n}) \geqslant 1 - e^{-K\beta} \tag{4.12}$$

and

$$P_x(\tau_{\bar{S}^{(n)}} \leqslant T_{n-1}) \geqslant 1 - e^{-K'\beta} \tag{4.13}$$

for some $K, K' > 0$ independent of β.

Moreover, for any $x \in S$ and for any $n \geqslant 1$, let $x_1^n, x_2^n,\ldots, x_i^n$ be the set of all the states in $S^{(n)}$ such that $x \in \bar{\mathscr{B}}^n_{x_j^n}$. Then

$$P_x(\text{the process falls to the bottom } x_j^n \text{ for some } j = 1,\ldots, i) \geqslant 1 - e^{-K\beta} \tag{4.14}$$

Proof. The proof immediately follows from the definition of the SGBA. Suppose that the process starting from x does not fall into x^n. For each trajectory of the process X_t, we can define a sequence of states of increasing stability as $x^0 = x$, and for each $i \geqslant 1$ we take $x^i \in S^{(i)}$ such that

$X_{\tau_{\bar{S}(i)}} \in \mathscr{E}^i_{x^i}$. If the process does not fall to x^n, there must be a state x^i of this sequence such that $i \leqslant n$ and $x^{i-1} \notin B^{(i-1)}_{x^i}$. This means that the sequence $X^{(i-1)}_s$, $s = 0, 1, 2,\dots$, that we can construct on the trajectory of the process (see Definition 2.1) does not follow a path of zero cost to reach a stable state. By using Lemmas 2.1 and 2.2 one can easily complete the proof. Equations (4.13), (4.14) follow from Proposition 2.1. ∎

Remark. If we define the *fall to the bottom* $p^{(n)}$ as the event that there exists a sequence $y^0, y^1,\dots$ satisfying Eq. (4.10) with $y^0 = x$, $y^n \in p^{(n)}$ such that

$$X_{\tau_{\bar{S}(i)}} \in \mathscr{E}^i_{y^i} \qquad \forall i \leqslant n$$

then, by using the definition of the SGBA of a plateau, it is immediate to prove that with probability of order one the process falls to the bottom $p^{(n)}$, i.e., the statement of Proposition 4.5 holds even in the case of a plateau.

We can prove something more: first of all, the mean exit time from a SGBA is exponentially large, namely:

Proposition 4.6. For any $n \geqslant 1$, for any $x^n_0 \in S^{(n)}$, and for any $x_0 \in \mathscr{E}^n_{x^n_0}$

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln E_{x_0} \tau_{(\mathscr{B}^n_{x^n_0})^c} = V_1 + \cdots + V_n + V^{(n)}_{x^n_0}$$

where V_i, $i = 1,\dots, n$, are defined by (2.14) and (2.11) and

$$V^{(n)}_{x^n_0} := \lim_{\beta \to \infty} \frac{1}{\beta} \ln E_{x^n_0} \tau^{(n)}_{(x^n_0)^c}$$

[see (2.11)].

The same holds for the SGBA of plateaus by replacing the single state x^n_0 with a plateau $p^{(n)}$ where

$$\mathscr{E}^n_{p^{(n)}} = \bigcup_{x^n \in p^{(n)}} \mathscr{E}^n_{x^n}$$

Proof. We define the set

$$D := \bigcup_{y^n \in S^{(n)},\, y^n \neq x^n_0} \mathscr{E}^n_{y^n}$$

By iteratively applying Theorem 2.1 we have that for any $x_0 \in \mathscr{E}^n_{x^n_0}$

$$E_{x_0} \tau_D \asymp \exp[(V_1 + \cdots + V_n + V^{(n)}_{x^n_0})\beta]$$

To get the theorem we have only to prove that

$$E_{x_0}\tau_D \asymp E_{x_0}\tau_{(\mathscr{B}^n_{x_0^n})^c}$$

By definition

$$(\mathscr{B}^n_{x_0^n})^c \supseteq D$$

Suppose now that

$$E_{x_0}\tau_D \succ E_{x_0}\tau_{(\mathscr{B}^n_{x_0^n})^c} \tag{4.15}$$

We will prove that this implies

$$P_{x_0}(\tau_D < e^{-a\beta}E_{x_0}\tau_D) \geqslant e^{-v\beta} \tag{4.16}$$

with $a>0$ independent of β and $v\to 0$ as $\beta\to\infty$, against (2.32).

Indeed if the inequality (4.15) holds, then there exists $d>0$ such that

$$E_{x_0}\tau_D e^{-(d/2)\beta} \geqslant E_{x_0}\tau_{(\mathscr{B}^n_{x_0^n})^c}e^{(d/2)\beta}$$

By Proposition 2.1 applied to $\mathscr{B}^n_{x_0^n}$ we have that for any $d>0$ there exists $k(d)>0$ such that

$$P_{x_0}(\tau_{(\mathscr{B}^n_{x_0^n})^c} \leqslant E_{x_0}\tau_{(\mathscr{B}^n_{x_0^n})^c}e^{+d\beta}) \geqslant 1 - e^{-k(d)\beta} \tag{4.17}$$

To prove (4.16) with $a=d/2$, since $E_{x_0}\tau_D \succcurlyeq T_n$, it is thus sufficient to show that for any

$$y \in D^c \cap (\mathscr{B}^n_{x_0^n})^c$$

there exists $y^n \neq x^n$ such that $y \in \bar{\mathscr{B}}^n_{y^n}$ and thus there exist $y^1, y^2,..., y^n$ such that, if d is sufficiently small (independent of β), we have

$$P_y(\tau_{\mathscr{E}^n_{y^n}} < T_n e^{-(d/2)\beta}) \geqslant e^{-\gamma\beta} \tag{4.18}$$

with $\gamma\to 0$ as $\beta\to\infty$.

Indeed we have

$$\begin{aligned} &P_y(\tau_{\mathscr{E}^n_{y^n}} < T_n e^{-(d/2)\beta}) \\ &\quad\geqslant P_y\left(X_{\tau_{\bar{S}^{(1)}}} \in \mathscr{E}^1_{y^1} \cap \tau_{\bar{S}^{(1)}} < \frac{T_n}{n}e^{-(d/2)\beta}\right) \end{aligned}$$

$$\times \min_{y_1 \in \mathscr{E}^1_{y^1}} P_{y_1}\left(X_{\tau_{\bar{S}(2)}} \in \mathscr{E}^2_{y^2} \cap \tau_{\bar{S}(2)} < \frac{T_n}{n} e^{-(d/2)\beta} \right) \cdots$$

$$\times \min_{y_{n-1} \in \mathscr{E}^{n-1}_{y^{n-1}}} P_{y_{n-1}}\left(X_{\tau_{\bar{S}(n)}} \in \mathscr{E}^n_{y^n} \cap \tau_{\bar{S}(n)} < \frac{T_n}{n} e^{-(d/2)\beta} \right) \geqslant e^{-\gamma\beta}$$

with $\gamma \to 0$ as $\beta \to \infty$. We have used the fact that the probabilities

$$P_{y_i}\left(\tau_{\bar{S}(i+1)} > \frac{T_n}{n} e^{-(a/2)\beta} \right)$$

are superexponentially small.

Exactly the same proof holds by replacing the state x^n_0 with a plateau $p^{(n)}$ by using the set

$$D := S \Big\backslash \Big(\bigcup_{y^n \notin p^{(n)}} \mathscr{E}_{y^n} \Big) \quad \blacksquare$$

In order to relate cycles and GBAs we give now two technical lemmas. Their proofs are in the Appendix. The first lemma contains a result on graphs and renormalization. The same result in the reversible and non-degenerate case is given in ref. 13. Here te presence of equivalent states makes the statement more complicated and the proof much longer.

Lemma 4.1. For any $k \geqslant 1$ and for any set $D^k \subset S^{(k)}$ we consider an arbitrary set $D^{k-1} \subset S^{(k-1)}$ such that

$$D^{k-1} \subseteq \bigcup_{x \in D^k} m_x^{(k-1)}, \qquad m_x^{(k-1)} \cap D^{k-1} \neq \varnothing \qquad \forall x \in D^k \tag{4.19}$$

Let

$$Q^k := S^{(k)} \backslash D^k$$

and

$$Q^{k-1} := S^{(k-1)} \backslash D^{k-1}$$

As in Section 3, we define for any $k \geqslant 0$

$$W^{(k)}_{Q^k} := \min_{g \in G^{(k)}(Q^k)} \sum_{i \to j \in g} \Delta^{(k)}(i, j)$$

and let $g^{(k)}_{Q^k}{}^*$ be a graph (not necessarily unique) in $G^{(k)}(Q^k)$ minimizing $W^{(k)}_{Q^k}$.

Then given a graph $g_{Q^k}^{(k)*}$ we can construct a graph $g_{Q^{k-1}}^{(k-1)*}$ minimizing $W_{Q^{k-1}}^{(k-1)}$ and we have

$$W_{Q^{k-1}}^{(k-1)} = W_{Q^k}^{(k)} + V_k\,|Q^k| \tag{4.20}$$

Conversely, given two sets D^k and D^{k-1} satisfying (4.19), given a graph $g_{Q^{k-1}}^{(k-1)*}$ minimizing $W_{Q^{k-1}}^{(k-1)}$, we can construct a graph $g_{Q^k}^{(k)*}$ minimizing $W_{Q^k}^{(k)}$ and satisfying (4.20).

The SGBAs share with the cycles the property that with large probability the bottom x_0 is visited before the exit from $\mathscr{B}^n_{x_0^n}$. More precisely:

Lemma 4.2. For any $n \geq 1$, for any $x_0^n \in S^{(n)}$, and for any $x_0 \in \mathscr{E}^n_{x_0^n}$ any $(\mathscr{B}^n_{x_0^n} \backslash x_0)$-graph minimizing

$$W_{\mathscr{B}^n_{x_0^n} \backslash x_0}$$

[see Eq. (3.8)] has no arrows exiting from $\mathscr{B}^n_{x_0^n}$.

The same holds for each stable plateau, i.e., $p^{(n)} \subset S^{(n)}$ with $V^{(n)}(x^n, y^n) > 0$ for each $x^n \in p^{(n)}$, $y^n \notin p^{(n)}$.

Remark. We note here that, due to Proposition 3.1, the statement of this lemma is equivalent to the following one:

For any $n \geq 1$, for any $x_0^n \in S^{(n)}$, and for any $x_0 \in \mathscr{E}^n_{x_0^n}$ and $x \in \mathscr{B}^n_{x_0^n}$ we have

$$P_x(X_{\tau(\mathscr{B}^n_{x_0^n} \backslash x_0)} \neq x_0) < e^{-K\beta}$$

for some $K > 0$.

We stress that this result is stronger than Proposition 4.5. In fact here *every* state $x_0 \in \mathscr{E}^n_{x_0^n}$ is visited before the exit from $\mathscr{B}^n_{x_0^n}$.

The following result is an easy consequence of Lemma 4.2:

Proposition 4.7. Any cycle $C' \subset \mathscr{B}^N_{x_0^N}$ such that $\mathscr{E}^N_{x_0^N} \not\subset C'$ satisfies

$$R_{C'} \subset \mathscr{B}^N_{x_0^N} \tag{4.21}$$

The same holds for stable plateaus, by replacing x_0^n with $p^{(n)}$.

Proof. Suppose (4.21) is false and let $g^*_{C'}$ be a graph minimizing $W_{C'}$ and containing arrows ending outside $\mathscr{B}^N_{x_0^N}$. Moreover, let $x_0' \in \mathscr{E}^N_{x_0^N} \cap C'^c$. If we consider now a graph

$$g^*_{\mathscr{B}^N_{x_0^N} \backslash x_0'}$$

minimizing

$$W_{\mathscr{B}^{N}_{x_0 N}\backslash x'_0}$$

we can construct a graph g' coinciding with

$$g^*_{\mathscr{B}^{N}_{x_0 N}\backslash x'_0}$$

for all arrows with starting points in $(\mathscr{B}^{N}_{x_0 N}\backslash x'_0)\backslash C'$ and coinciding with $g^*_{C'}$ for the arrows starting in C'. The new graph g', because of the argument of proof of Proposition 4.2(a), is indeed a $(\mathscr{B}^{N}_{x_0 N}\backslash x'_0)$-graph minimizing

$$W_{\mathscr{B}^{N}_{x_0 N}\backslash x'_0}$$

Since $g^*_{C'}$ has a unique arrow exiting from $\mathscr{B}^{N}_{x_0 N}$, g' itself has an arrow exiting from $\mathscr{B}^{N}_{x_0 N}$ against the statement of Lemma 4.2. ∎

We prove now the main result of this section:

Proposition 4.8. For any $x'' \in S^{(n)}$ there exists a cycle $C_{x''}$ for the chain X_t such that

$$C_{x''} \cap \bar{S}^{(n)} = \mathscr{E}^{n}_{x''}$$

and

$$V_{C_{x''}} = V_1 + \cdots + V_n + V^{(n)}_{x''}$$

$C_{x''}$ turns out to be the maximum cycle containing $\mathscr{E}^{n}_{x''}$ and contained in the SGBA of x'': $\mathscr{B}^{n}_{x''}$.

Proof. For notational convenience we set in this proof $\mathscr{B}^{n}_{x''} = B$.

Suppose that for any $x_0 \in \mathscr{E}^{n}_{x''}$ we are able to show that the maximal cycle C contained in B and containing x_0 is such that:

(i) $\mathscr{E}^{n}_{x''} \subseteq C$.

(ii) $V_C = V_B(x_0) := \lim_{\beta \to \infty}(1/\beta) \ln E_{x_0}\tau_{B^c}$.

Then, by Proposition 4.6, we get the result. Point (i) means that the cycle C is the same for each $x_0 \in \mathscr{E}^{n}_{x''}$. To prove (i) and (ii) we fix $x_0 \in \mathscr{E}^{n}_{x''}$ and we consider the maximal rank k of the maximal cycle C contained in B: $C \subset B$, $x_0 \in C$, $C \in \mathscr{C}^k$, where the maximality of the rank k means that the cycle of rank $k+1$ containing C is not contained in B.

To prove (i) we note that, as a consequence of Proposition 4.7, the SGBA B is measurable with respect to the family $\mathscr{C}^k$ of cycles of rank k, i.e., $B = \bigcup_{i \in b} C^k_i$, where b is a suitable set of indices and $C = C^k_j$ for some

$j \in b$. Indeed, suppose *ab absurdo* that there is a $C' \in \mathscr{C}^k$, $C' \neq C$ with $C' \cap B \neq \emptyset$ and $C' \not\subset B$. Then there must be a cycle $C'' \subset B$ with $R_{C''} \not\subset B$ against Proposition 4.7.

Thus not only is B measurable with respect to $\mathscr{C}^k$, but also every $C_i^k \neq C$, $i \in b$, has its successor in B: $R_{C_i^k} \subset B$. This means that if (i) is not satisfied, if $C \not\supset \mathscr{E}^n_{x^n}$, then again by Proposition 4.7, $R_C \subset B$, so that the descendants of C are in B and thus the cycle C^{k+1} of rank $k+1$ containing C is contained in B against the hypothesis of maximality of C. This prove that (i) holds and that

$$R_C \cap B^c \neq \emptyset \tag{4.22}$$

By (i) we know that $V_C \leqslant V_B(x_0)$. If now $V_C(x_0) = V_B(x_0) - 2a$ with $a > 0$, we would have, using Proposition 2.1,

$$P_{x_0}(\tau_{B^c} < e^{\beta V_C(x_0) + a}) < e^{-K\beta}$$

with $K = K(a) > 0$; but, on the other hand, by (4.22) we get

$$\begin{aligned} &P_{x_0}(\tau_{B^c} < e^{\beta V_C(x_0) + a}) \\ &\qquad \geqslant P_{x_0}(\tau_{C^c} < e^{\beta V_C(x_0) + a} \cap X_{\tau_{C^c}} \notin B) \geqslant e^{-\delta\beta} \end{aligned}$$

with $\delta \to 0$ as $\beta \to \infty$. This leads to a contradiction, proving (ii). ∎

Proposition 4.9. For any plateau $p^{(n)}$ which is stable for the chain $X_t^{(n)}$, i.e., $p^{(n)} = m^{(n)}_{x^{n+1}}$ for some $x^{n+1} \in S^{(n+1)}$, there exists a cycle $C_{p^{(n)}}$ for the chain X_t such that

$$C_{p^{(n)}} \cap \bar{S}^{(n)} = \mathscr{E}^n_{p^{(n)}} = \mathscr{E}^{n+1}_{x^{n+1}}$$

and

$$V_{C_{p^{(n)}}} = V_1 + \cdots + V_n + V^{(n)}_{p^{(n)}}$$

$C_{p^{(n)}}$ turns out to be the maximum cycle contained in $\bar{\mathscr{B}}^n_{p^{(n)}}$ containing $\mathscr{E}^n_{p^{(n)}}$.

The proof can be obtained exactly as in Proposition 4.8.

5. THE MAIN THEOREM

In this section we give the main tool for the control of the behavior of the chain X_t on exponentially long time intervals. This result completes the analysis developed in ref. 12.

For any pair of states $x^n, y^n \in S^{(n)}$ we can in fact describe the behavior of the chain X_t by knowing that the corresponding chain $X_t^{(n)}$ is doing the transition $x^n \to y^n$. More precisely, for any state $x^n \in S^{(n)}$ we can define a set Q_{x^n} contained in the GBA of x^n as follows.

There are two possible cases: (a) $m_{x^n}^{(n-1)} = x^{n-1}$ and (b) $m_{x^n}^{(n-1)} = p^{(n-1)}$.

Let $C_{x^{n-1}}$ ($C_{p^{(n-1)}}$) be the maximal cycle contained in $\mathscr{B}^{n-1}_{x^{n-1}}$ ($\mathscr{B}^{n-1}_{p^{(n-1)}}$) given by Proposition 4.8 (Proposition 4.9).

For simplicity we will consider in what follows only case (a). The analysis of case (b) is exactly the same since the same properties have been proved for generic single states and for stable plateaus in Propositions 4.5–4.9 and Lemma 4.2.

As shown in the proof of Proposition 4.8, the cycles $C_1, \ldots, C_l$ intersecting $\mathscr{B}^{n-1}_{x^{n-1}}$ of the same rank as $C_{x^{n-1}}$, say $C_1 \equiv C_{x^{n-1}}$, are strictly contained in $\mathscr{B}^{n-1}_{x^{n-1}}$ and provide a partition of $\mathscr{B}^{n-1}_{x^{n-1}}$. Moreover, by the same argument (again by Proposition 4.7), we have that for any $x \in \mathscr{B}^{n-1}_{x^{n-1}}$ any cycle containing x either is contained in $\mathscr{B}^{n-1}_{x^{n-1}}$ or it contains $C_{x^{n-1}}$.

Let us now consider the set $\bar{\mathscr{B}}^{n-1}_{x^{n-1}} \backslash \mathscr{B}^{n-1}_{x^{n-1}}$. Each state x contained in this set is such that the minimum cycle containing x with nonempty intersection with some $\mathscr{B}^{n-1}_{y^{n-1}}$ contains $C_{x^{n-1}}$ and $C_{y^{n-1}}$. Indeed each $x \in \bar{\mathscr{B}}^{n-1}_{x^{n-1}} \backslash \mathscr{B}^{n-1}_{x^{n-1}}$ is such that there exist at least two sequences $x^0 = x$, x^1, $x^2, \ldots, x^{n-1}$ and $y^0 = x$, $y^1, y^2, \ldots, y^{n-1}$ such that $x^i \in B^{(i)}_{x^{i+1}}$ and $y^i \in B^{(i)}_{y^{i+1}}$. Moreover, if $x \in B_{x^1}$, then the minimum cycle different from $\{x\}$ containing x contains x^1, and since $x \in B_{y^1}$, the minimum cycle different from $\{x\}$ containing x contains also y^1. If $x^1 \neq y^1$, the minimum cycle containing x^1 and y^1 strictly contains also C_{x^1} and C_{y^1}. Again, if a cycle strictly contains C_{x^1} and $x^1 \in B^{(1)}_{x^2}$, then it contains x^2, and the same for y^1 and y^2. By iteration we can conclude that the minimum cycle containing x with nonempty intersection with some $\mathscr{B}^{n-1}_{y^{n-1}}$ contains $C_{x^{n-1}}$.

In conclusion we have that both the sets $\mathscr{B}^{n-1}_{x^{n-1}}$ and $\bar{\mathscr{B}}^{n-1}_{x^{n-1}} \backslash \mathscr{B}^{n-1}_{x^{n-1}}$ are measurable with respect to the family of cycles of the same rank as $C_{x^{n-1}}$:

$$\bar{\mathscr{B}}^{n-1}_{x^{n-1}} = C_1 \cup C_2 \cup \cdots \cup C_l \cup \cdots \cup C_L \tag{5.1}$$

For each one of these cycles we have that $V_{C_i} \leqslant V_1 + \cdots + V_{n-2}$ if $i > 1$.

We can now define the *permanence set* Q_{x^n} associated to each $x^n \in S^{(n)}$:

Definition 5.1. Among this set of cycles $\{C_i\}_{i=1,\ldots,L}$ we will consider the maximal subset $\{C_i\}_{i \in \mathscr{I}}$, with $\mathscr{I} \subseteq \{1, \ldots, L\}$ such that for any $i \in \mathscr{I}$ there exists a sequence $j_1(i), j_2(i), \ldots, j_m(i) \in \mathscr{I}$ with $j_1(i) = 1$, $j_m(i) = i$, and $R_{C_{j_k(i)}} \cap C_{j_{k+1}(i)} \neq \varnothing$. Then we define the permanence set associated to x^n:

$$Q_{x^n} = \bigcup_{i \in \mathcal{I}} C_i \tag{5.2}$$

The set Q_{x^n} is the set of states visited by the chain X_t in the interval of time corresponding to a transition $x^n \to y^n$ of the chain $X_t^{(n)}$. More precisely, let

$$\sigma^{(n-1)} := \tau^{(n-1)}_{(m_{x^n}^{(n-1)})^c}$$

and for any $T > 0$ let

$$\begin{aligned}
&\Phi^{(n-1)}(x^n, y^n, T)\\
&\quad := \{\phi^{(n-1)} \in \Phi(S^{(n-1)}) \text{ such that}\\
&\qquad\quad \phi_0^{(n-1)} \in m_{x^n}^{(n-1)}, \phi_T^{(n-1)} \in m_{y^n}^{(n-1)}, \phi_i^{(n-1)} \notin M^{(n-1)}\ \forall i = 1,..., T-1\}
\end{aligned}$$

and let Z_t^{n-1} be defined by (2.23).

We have the following:

Theorem 5.1. For any $a > 0$ sufficiently small (see Lemma 2.4), we define

$$\begin{aligned}
A_1 &:= \{Z^{n-1}_{\sigma^{(n-1)}-1} \in [T_n e^{-a\beta}, T_n e^{a\beta}]\}\\
A_2 &:= \{Z^{n-1}_{\sigma^{(n-1)}-1} < \tau_{(Q_{x^n})^c} \leqslant Z^{n-1}_{\sigma^{(n-1)}}\}\\
A_3 &:= \{\exists T, \phi^{(n-1)} \in \Phi^{(n-1)}(x^n, y^n, T) \text{ minimizing } I^{(n-1)}_{[0,T]}\\
&\qquad\qquad \text{such that } X_t \in \mathcal{T}(\phi^{(n-1)}, T)\ \forall t > Z^{n-1}_{\sigma^{(n-1)}-1}\}\\
G &:= \{X_0^{(n)} = x^n, X_1^{(n)} = y^n\}
\end{aligned}$$

There exists a positive constant K, depending on a but independent of β, such that for any sufficiently large β and for any $x \in \mathscr{E}^n_{x^n}$ we have

$$P_x(A_1 \cap A_2 \cap A_3 \mid G) \geqslant 1 - e^{-K\beta}$$

Moreover, if we add the hypothesis that

$$\text{there exists} \quad z^n \neq x^n \quad \text{such that} \quad P^{(n)}(x^n, z^n) \asymp 1 \tag{5.3}$$

and if we define

$$A_4 := \{\forall y \in C_{m_{x^n}^{(n-1)}}\ \exists t < Z^{(n-1)}_{\sigma^{(n-1)}-1} \text{ such that } X_t = y\}$$

then we have, for any sufficiently large β and for any $x \in \mathscr{E}^n_{x^n}$

$$P_x(A_1 \cap A_2 \cap A_3 \cap A_4 \mid G) \geqslant 1 - e^{-K\beta}$$

Proof. We will prove that for each $i = 1,..., 4$ we have

$$P_x(A_i^c \mid G) \leqslant e^{-K\beta}$$

Let $|m^{(n-1)}_{x^n}| = m$; we have

$$\begin{aligned}
P_x(G) &= \sum_{t=0}^{t_n-1} \sum_{z^{n-1} \in m^{(n-1)}_{x^n}} P_x(\{\sigma^{(n-1)} > t\} \cap \{X^{(n-1)}_t = z^{n-1}\}) \\
&\quad \times \sum_T \sum_{\phi^{(n-1)} \in \Phi^{(n-1)}(x^n, y^n, T)} P_{z^{n-1}}(X^{(n-1)}_s = \phi^{(n-1)}_s \ \forall s \leqslant T) \\
&= \sum_{t'=0}^{(t_n/m)-1} \sum_{t''=0}^{m-1} \sum_{z^{n-1} \in m^{(n-1)}_{x^n}} P_x(\{\sigma^{(n-1)} > mt' + t''\} \cap \{X^{(n-1)}_{mt'+t''} = z^{n-1}\}) \\
&\quad \times \sum_T \sum_{\phi^{(n-1)} \in \Phi^{(n-1)}(x^n, y^n, T)} P_{z^{n-1}}(X^{(n-1)}_s = \phi^{(n-1)}_s \ \forall s \leqslant T)
\end{aligned} \tag{5.4}$$

Since $m^{(n-1)}_{x^n}$ is a class of stable equivalent states in $S^{(n-1)}$, we have by definition that for each pair of states in $m^{(n-1)}_{x^n}$ there is a path $\psi^{(n-1)}$ in $S^{(n-1)}$ connecting these states of length at most m with $\psi^{(n-1)}_i \neq \psi^{(n-1)}_{i+1}$ and $I^{(n-1)}(\psi^{(n-1)}) = 0$, to which we can apply Lemma 2.1(ii).

In this way we obtain the following estimate:

$$\begin{aligned}
&\sum_{t''=0}^{m-1} P_x(\{\sigma^{(n-1)} > mt' + t''\} \cap \{X^{(n-1)}_{mt'+t''} = z^{n-1}\}) \\
&\qquad = \sum_{t''=0}^{m-1} \sum_{u^{n-1} \in m^{(n-1)}_{x^n}} P_x(\{\sigma^{(n-1)} > mt'\} \cap \{X^{(n-1)}_{mt'} = u^{n-1}\}) \\
&\qquad\quad \times P_{u^{n-1}}(\{\sigma^{(n-1)} > t''\} \cap \{X^{(n-1)}_{t''} = z^{n-1}\}) \\
&\qquad \geqslant e^{-\alpha\beta} P_x(\{\sigma^{(n-1)} > mt'\})
\end{aligned} \tag{5.5}$$

By using (5.5) from (5.4), we obtain

$$\begin{aligned}
P_x(G) &\geqslant \sum_{t'=0}^{t_n/m-1} P_x(\sigma^{(n-1)} > mt')\, e^{-\alpha\beta} \\
&\quad \times \sum_{z^{n-1} \in m^{(n-1)}_{x^n}} \sum_T \sum_{\substack{\phi^{(n-1)} \in \\ \Phi^{(n-1)}(x^n, y^n, T)}} P_{z^{n-1}}(X^{(n-1)}_s = \phi^{(n-1)}_s \ \forall s \leqslant T)
\end{aligned} \tag{5.6}$$

Let us now consider the case $i=1$. We have

$$P_x(A_1^c \mid G) \leqslant P_x(A_0^c \mid G) + P_x(A_0 \cap A_1^c \mid G)$$

with

$$A_0 := \{t_n > \sigma^{(n-1)} > t_n e^{-(a/2)\beta}\}$$

We remark that $\{\sigma^{n-1} > t_n\} \cap G = \varnothing$.

We can estimate from above, by using the same expansion used to estimate $P_x(G)$, the following probabilities:

$$\begin{aligned} P_x(A_0^c \cap G) = & \sum_{t < t_n e^{-(a/2)\beta}} \sum_{z^{n-1} \in m_{x^n}^{(n-1)}} P_x(\{\sigma^{(n-1)} > t\} \cap \{X_t^{(n-1)} = z^{n-1}\}) \\ & \times \sum_T \sum_{\substack{\phi^{(n-1)} \in \\ \Phi^{(n-1)}(x^n, y^n, T)}} P_{z^{n-1}}(X_s^{(n-1)} = \phi_s^{(n-1)}\ \forall s \leqslant T) \\ \leqslant & \sum_{t < t_n e^{-(a/2)\beta}} P_x(\sigma^{(n-1)} > t) \\ & \times \sum_{z^{n-1} \in m_{x^n}^{(n-1)}} \sum_T \sum_{\substack{\phi^{(n-1)} \in \\ \Phi^{(n-1)}(x^n, y^n, T)}} P_{z^{n-1}}(X_s^{(n-1)} = \phi_s^{(n-1)}\ \forall s \leqslant T) \end{aligned}$$

so that

$$\begin{aligned} P(A_0^c \mid G) & \leqslant \frac{\sum_{t < t_n e^{-(a/2)\beta}} P_x(\sigma^{(n-1)} > t)}{\sum_{t'=0}^{t_n/m - 1} P_x(\sigma^{(n-1)} > mt')\, e^{-\alpha\beta}} \\ & \leqslant \frac{\sum_{t < t_n e^{-(a/2)\beta}} P_x(\sigma^{(n-1)} > t)}{\sum_{t'=0}^{t_n/m e^{-(a/4)\beta}} P_x(\sigma^{(n-1)} > mt')\, e^{-\alpha\beta}} \\ & \leqslant \frac{e^{-(a/4)\beta} t_n}{e^{-\alpha'\beta} t_n} \end{aligned}$$

where $\alpha' \to 0$ as $\beta \to \infty$ and we used Proposition 2.1 to get the last inequality.

Let us consider the probability $P_x(A_0 \cap A_1^c \mid G)$. Since for any $t \in [t_n e^{-(a/2)\beta}, t_n]$ we have that

$$\begin{aligned} & P_x(\{Z^{n-1}_{\sigma^{(n-1)}-1} < T_n e^{-a\beta}\} \cup \{Z^{n-1}_{\sigma^{(n-1)}-1} > T_n e^{a\beta}\}) \\ & \quad \leqslant P_x(\{Z^{n-1}_{\sigma^{(n-1)}-1} < t \cdot T_{n-1} e^{-(a/2)\beta}\} \cup \{Z^{n-1}_{\sigma^{(n-1)}-1} > t \cdot T_{n-1} e^{(a/2)\beta}\}) \end{aligned}$$

Then, proceeding as before, we have

$$\begin{aligned} &P_x(A_0 \cap A_1^c \cap G) \\ &\quad \leqslant \sum_{t=t_n e^{-(a/2)\beta}}^{t_n - 1} \sum_{z^{n-1} \in m_{x^n}^{(n-1)}} P_x(\{\sigma^{(n-1)} > t\} \cap \{X_t^{(n-1)} = z^{n-1}\} \\ &\qquad \cap [\{Z_t^{n-1} < t \cdot T_{n-1} e^{-(a/2)\beta}\} \cup \{Z_t^{n-1} > t \cdot T_{n-1} e^{(a/2)\beta}\}]) \\ &\qquad \times \sum_T \sum_{\substack{\phi^{(n-1)} \in \\ \Phi^{(n-1)}(x^n, y^n, T)}} P_{z^{n-1}}(X_s^{(n-1)} = \phi_s^{(n-1)} \ \forall s \leqslant T) \end{aligned}$$

It is not difficult to extend the same argument of the proof of Lemma 2.4 to obtain, for any $t \in \mathbf{N}$, the inequality

$$P_x(t \cdot T_{n-1} e^{-(a/2)\beta} \leqslant Z_t^{n-1} \leqslant t \cdot T_{n-1} e^{(a/2)\beta}) \geqslant 1 - e^{-k\beta}$$

with $k = k(a)$, and thus we can prove as before that $P_x(A_0 \cap A_1^c \mid G)$ is exponentially small.

Let us now go to the case $i = 2$. We first observe that, by Lemma 2.3, since

$$\bar{\mathscr{B}}^{n-1}_{m_{x^n}^{n-1}} \cap \bar{S}^{n-1} = \mathscr{E}^{n-1}_{m_{x^n}^{n-1}}$$

we have that $Z^{n-1}_{\sigma^{(n-1)}} \geqslant \tau_{(Q_{x^n})^c}$; thus, by using the same expansion used in the case $i = 1$, we have only to prove that $P_x(\tau_{(Q_{x^n})^c} < Z^{n-1}_{\sigma^{(n-1)}-1})$ is exponentially small in β.

We have

$$\begin{aligned} &P_x(\tau_{(Q_{x^n})^c} < Z^{n-1}_{\sigma^{(n-1)}-1}) \\ &\quad \leqslant P_x\left(X_{\tau_{(Q_{x^n})^c}} \notin \left(\bigcup_{i \in \mathscr{I}} R_{C_i}\right)\right) \\ &\qquad + \sum_{z \in (Q_{x^n}^c \cap (\bigcup_{i \in \mathscr{I}} R_{C_i}))} P_z(X_{\tau_{\bar{S}^{(n-1)}}} \in \mathscr{E}^{n-1}_{m_{x^n}^{(n-1)}}) \end{aligned} \tag{5.7}$$

To estimate the first term in the r.h.s. of (5.7) we first note that, by the definition of $C_{x^{n-1}}$ and Q_{x^n}, there exists at least an index $i_0 \in \mathscr{I}$ such that $R_{C_{i_0}} \notin Q_{x^n}$ and a sequence $j_1 = 1, \ldots, j_k = i_0$ such that

$$R_{C_{j_i}} \cap C_{j_{i+1}} \neq \varnothing \tag{5.8}$$

On the other hand, since

$$C_i \subset \mathscr{B}^{n-1}_{m^{(n-1)}_{x^n}} \qquad \text{for any} \quad i \in \mathscr{I}$$

for any $i \neq \{j_1,..., j_k\}$ there is a sequence satisfying again (5.8) starting with i and ending in $j_1,..., j_k$. By iterating this argument we can conclude that we can choose a sequence of graphs $g^*_{C_i}$, $i \in \mathscr{I}$, minimizing, respectively, W_{C_i} such that

$$\bigcup_{i \in \mathscr{I}} g^*_{C_i} =: g_{Q_{x^n}} \tag{5.9}$$

is a Q_{x^n}-graph. As was noted in the proof of Proposition 4.2, we have, for any $g \in G(Q_{x^n})$

$$W(g) = \sum_{i \in \mathscr{I}} W(g \upharpoonright_{C_i})$$

and since $g \upharpoonright_{C_i}$ are C_i-graphs, we have

$$W(g) \geqslant \sum_{i \in \mathscr{I}} W(g^*_{C_i}) = \sum_{i \in \mathscr{I}} W_{C_i}$$

This implies that each Q_{x^n}-graph minimizing $W_{Q_{x^n}}$ has the form (5.9) and thus

$$R_{Q_{x^n}} \subset \bigcup_{i \in \mathscr{I}} R_{C_i} \tag{5.10}$$

By means of the FW results (see Lemma 3.2 and Proposition 3.2) we can then estimate the first term in the r.h.s. of (5.7) with an exponentially small quantity.

The second term in the r.h.s. of (5.7) can be estimated by Proposition 4.5 if we remark that, by construction, each point in $(Q_{x^n})^c \cap (\bigcup_{i \in \mathscr{I}} R_{C_i})$ belongs to

$$(\bar{\mathscr{B}}^{n-1}_{m^{(n-1)}_{x^n}})^c$$

Let us now consider the case $i = 3$. As in the case $i = 1$, we define an auxiliary event

$$A'_3 := \{\bar{\tau}^{(n-1)} \leqslant e^{\eta\beta}\}$$

where η is arbitrarily small and

$$\bar{\tau}^{(n-1)} := \min\{t > \sigma^{(n-1)} - 1; X^{(n-1)}_t \in M^{(n-1)}\}$$

We have

$$P_x(A_3^c \cap G) \leqslant P_x(A_3'^c \cap G) + P_x(A_3^c \cap A_3' \cap G) \tag{5.11}$$

By using the fact that the event A_3' depends only on the process $X^{(n-1)}$ after the time $\sigma^{(n-1)} - 1$ and by using again the expansion used to estimate $P_x(G)$, we immediately obtain a superexponentially small estimate of the first term in the r.h.s. of (5.11) by means of Lemma 2.2(ii) applied to the chain $X_t^{(n-1)}$. The second term can be estimated from above as follows:

$$\begin{aligned} P_x(A_3^c \cap A_3' \cap G) \leqslant \sum_{t=0}^{t_n-1} \sum_{z^{n-1} \in m_{x^n}^{(n-1)}} & P_x(\{\sigma^{(n-1)} > t\} \cap \{X_t^{(n-1)} = z^{n-1}\}) \\ \times \sum_{T \leqslant e^{\eta\beta}} \sum_{\substack{\phi^{(n-1)} \in \\ \bar{\Phi}^{(n-1)}(x^n, y^n, T)}} & P_{z^{n-1}}(X_s^{(n-1)} = \phi_s^{(n-1)} \ \forall s \leqslant T) \end{aligned} \tag{5.12}$$

where

$$\begin{aligned} &\bar{\Phi}^{(n-1)}(x^n, y^n, T) \\ &\quad = \{\phi \in \Phi^{(n-1)}(x^n, y^n, T) \text{ such that} \\ &\qquad\qquad I_{[0,T]}^{(n-1)}(\phi) > \min_{\psi^{(n-1)} \in \Phi^{(n-1)}(x^n, y^n, T)} I_{[0,T]}^{(n-1)}(\psi)\} \end{aligned}$$

and so, by applying Lemma 2.1(iii) we obtain an exponentially small estimate for $P_x(A_3 \mid G)$.

In the case in which there exists z^n such that $P^{(n)}(x^n, z^n) \asymp 1$, we can also prove that $P_x(A_4^c \mid G)$ is exponentially small. In fact, in this case, by Proposition 4.8 we have that

$$V_{C_{m_{x^n}^{(n-1)}}} = V_1 + \cdots + V_n$$

and, with probability exponentially near to one,

$$\tau_{C_{m_{x^n}^{(n-1)}}} \asymp Z_{\sigma^{(n-1)}-1}^{n-1}$$

The proof thus follows immediately by using Proposition 4.3(iv), the Chebyshev inequality, and the expansion used to estimate $P_x(G)$. ∎

Remark. We note that, if $P^{(n)}(x^n, y^n)$ is exponentially small, i.e., if the transition $x^n \to y^n$ is against the flow, with Theorem 5.1 we provide estimates from below of probabilities conditioned to an event G of exponentially small probability. This is the main difficulty of this theorem and for this reason we used explicit expansions in the proof.

The case $P^{(n)}(x^n, y^n) \asymp 1$ is easier. Indeed the process is "falling" to y^n, and we are conditioning to an event of probability of order one. The results proved in Theorem 5.1 can be stated in this case in the following form:

Theorem 5.2. Let $x^n, y^n \in S^{(n)}$ be such that $P^{(n)}(x^n, y^n) \asymp 1$. If we define the events

$$D_1 := \{T_n e^{-\alpha\beta} \leqslant \tau_{(Q_{x^n})^c} \leqslant T_n e^{+\alpha\beta}\}$$

$$D_2 := \{\forall y \in C_{x^n}\ \exists t < \tau_{(Q_{x^n})^c} \text{ such that } X_t = y\}$$

$$D_3 := \{X_{\tau_{(Q_{x^n})^c}} \in \bar{\mathscr{B}}^n_{y^n} \text{ and starting from } X_{\tau_{(Q_{x^n})^c}}, \text{ the process falls to } y^n\}$$

$$G := \{X_0^{(n)} = x^n, X_1^{(n)} = y^n\}$$

then for any $x \in \mathscr{E}^n_{x^n}$ we have

$$P_x(D_1 \cap D_2 \cap D_3 \mid G) \geqslant 1 - e^{-K\beta}$$

for some $K > 0$ independent of β.

Proof. By hypothesis, $P_x(G) = P^{(n)}(x^n, y^n) \asymp 1$.

By Proposition 2.1, since the states in $\mathscr{E}^n_{x^n}$ are the most stable states contained in Q_{x^n}, we have immediately

$$P(D_1^c \mid G) < e^{-K\beta} \tag{5.13}$$

for some $K > 0$.

Since

$$\tau_{(C_{m_{x^n}^{(n-1)}})^c} \leqslant \tau_{(Q_{x^n})^c}$$

by Proposition 4.3(iii) we have that

$$P_x(D_2^c \mid G) \leqslant e^{-K\beta} \tag{5.14}$$

for some $K > 0$.

To get $P(D_3^c \mid G) < e^{-K\beta}$ we notice that the probability that the process, starting from $X_{\tau_{(Q_{x^n})^c}}$, visits $\mathscr{E}^n_{z^n}$ before $\mathscr{E}^n_{y^n}$, $z_n \in S^{(n)}$, $z^n \neq y^n$, when conditioned to the event G, is zero. By using Proposition 4.5 we conclude the proof of the theorem. ▮

Remark. The main difference between the statements of Theorems 5.1 and 5.2 is that in Theorem 5.1 we were considering times of the form Z_t^{n-1}, in order to be able to iterate the theorem itself to obtain a complete description of the transition $x^n \to y^n$ in terms of the original chain X_t.

Here, in Theorem 5.2 we are considering times of the form $\tau_{(Q_{x^n})^c}$, since, as we will show in the following theorem, we will consider the iteration on the stability of the state to which the process is falling. Let $x^n \in S^{(n)}$ and $x \in \bar{\mathscr{B}}^n_{x^n}$.

Theorem 5.3. Let $B^{(n-1)}_{x^n} \equiv x^{n-1}_1, \dots, x^{n-1}_l$. For each of these states x^{n-1}_i there is at least a path $\bar{\phi}^{(n-1)}(x^{n-1}_i) = \{\bar{\phi}^{(n-1)}_k\}_{k=0,\dots,T}$ in $S^{(n-1)}$, with

$$\bar{\phi}^{(n-1)}_0 = x^{n-1}_i, \qquad \bar{\phi}^{(n-1)}_T \in m^{(n-1)}_{x^n}, \qquad P^{(n-1)}(\bar{\phi}^{(n-1)}_k, \bar{\phi}^{(n-1)}_{k+1}) \asymp 1 \qquad (5.15)$$

We define the following event:

$$\begin{aligned} E := \{ &\text{there exist } x^{n-1}_i \in B^{n-1}_{x^n} \text{ and } \bar{\phi}^{(n-1)}(x^{n-1}_i) \\ &\text{satisfying Eq. (5.15) such that } x \in \bar{\mathscr{B}}^{n-1}_{x^{n-1}_i}, \\ &x \text{ falls to } x^{n-1}_i, \text{ and } X^{(n-1)}_s = \bar{\phi}^{(n-1)}_s \ \forall s = 0,\dots, T\} \end{aligned}$$

Then we have

$$P_x(E \mid x \text{ falls to } x^n) \geqslant 1 - e^{-K\beta}$$

Proof. The proof of this theorem is an immediate consequence of Proposition 4.5 and Lemmas 2.1 and 2.2.

6. THE TUBE OF EXIT

We come now to the problem of the determination of the tube of exit. We will use the renormalization procedure and the results of the previous section to define the tube of typical exiting paths in terms of a sequence of permanence sets (see Definition 5.1).

We use the notation established in the previous sections. A generic state in $S^{(n)}$ will be denoted by x^n, and a path in $S^{(n)}$ will be denoted by ϕ^n. Since in this section states of different spaces $S^{(n)}$ will appear at the same time, we do not simplify the notation by omitting indices. With boldface letters, e.g., $\mathbf{y}$, we will denote sequences of states not necessarily belonging to the same state space.

Let $Q \subset S$ be an arbitrary domain. As in Proposition 2.1, we can substitute our original chain X_t with the chain X^Q_t defined by (2.30). Let $N = N(Q)$ be defined by (2.31). We omit, from now on, the superscript Q, since only the chain X^Q_t will be considered in the rest of this section. The characterization of the typical exit of the chain $X^{(N)}_t$ from the domain

$Q \cap S^{(N)}$ is an easy problem since this domain does not contain stable states for the chain $X_t^{(N)}$ and so, with large probability, the chain $X_t^{(N)}$ exits from $Q \cap S^{(N)}$ in a time of order one by following a path falling to Q^c.

For any $x^N \in Q \cap S^{(N)}$ we define the set of typical exiting paths starting from x^N:

$$\Psi^{(N)}(x^N, Q^c) := \bigcup_{T=1}^{\infty} \{\phi_0^N, \phi_1^N, ..., \phi_T^N \in S^{(N)}: \phi_0^N = x^N, \phi_T^N \in Q^c, \phi_t^N \notin M^{(N)}$$

$$\forall t = 1, ..., T-1 \text{ and } I_{[0,T]}^{(N)}(\phi^N) = 0\} \tag{6.1}$$

By Lemma 2.2 we have the inequality

$$P_{x^N}(\exists \phi^N \in \Psi^{(N)}(x^N, Q^c): X_t^{(N)} = \phi_t^N \ \forall t \leqslant T(\phi^N)) > 1 - e^{-k\beta} \tag{6.2}$$

where $T(\phi^N)$ is the first hitting time to Q^c of the path ϕ^N.

This can be called "N-descent to Q^c" (in the reversible case the Nth renormalized energy function is decreasing on the paths in $\Psi^{(N)}(x^N, Q^c)$).

As noted in ref. 11, Section 2, (6.2) gives us a first knowledge about the exit of the chain X_t from the domain Q, providing, in particular, the results obtained by Freidlin and Wentzell about the mean exit time and the most probable exit point. A first preliminary, quite rough version of the exit tube starting in $\mathscr{E}_{x^N}^N$ is thus given by the following union of paths:

$$\mathscr{T} := \bigcup_{\phi^N \in \Psi^{(N)}(x^N, Q^c)} \mathscr{T}(\phi^N, T(\phi^N)) \tag{6.3}$$

where $\mathscr{T}(\phi^N, T(\phi^N))$ is defined in (2.19).

On the other hand, each descending path $\phi^N \in \Psi^{(N)}(x^N, Q^c)$ can be analyzed in terms of paths of the chain $X^{(N-1)}$. An N-descent to Q^c for the path ϕ^N will be a sequence of descents and ascents for the paths on the smaller scale $N-1$.

Now we want to use the results of the previous sections to give more details on the typical exiting paths, in order to describe these sequences of descents and ascents up to the first (better, zeroth) level of the renormalization procedure (corresponding to the original chain). In this way we want to narrow the tube $\mathscr{T}$ as much as possible keeping a good control in probability. This will be achieved by describing, for each path $\{\phi^N\}_{i=1}^T$ appearing in (6.3), the behavior of the original process X_t on the time intervals corresponding to each transition $\phi_i^N \to \phi_{i+1}^N$ at level N.

We can proceed as follows: by applying Theorem 5.1, we can associate to each transition $\phi_i^N \to \phi_{i+1}^N$ a permanence set $Q_{\phi_i^N}$ and a set of paths $\Psi^{(N-1)}$ in $\Phi(S^{(N-1)})$ defined by

$$\begin{aligned}
&\Psi^{(N-1)}(m_{\phi_i^N}^{(N-1)}, m_{\phi_{i+1}^N}^{(N-1)}) \\
&\quad := \bigcup_{T=1}^{\infty} \{\phi_0^{N-1}, \phi_1^{N-1},..., \phi_T^{N-1} \in S^{(N-1)}: \phi_0^{N-1} \in m_{\phi_i^N}^{(N-1)}, \\
&\qquad \phi_T^{N-1} \in m_{\phi_{i+1}^N}^{(N-1)}, \phi_t^{N-1} \notin M^{(N-1)}\ \forall t=1,..., T-1 \text{ and} \\
&\qquad I_{[0,T]}^{(N-1)}(\phi^{N-1}) = \Delta^{(N)}(\phi_i^N, \phi_{i+1}^N) + V_N\}
\end{aligned} \tag{6.4}$$

To each transition at level $N-1$ we can associate, again by Theorem 5.1, a permanence set $Q_{\phi_j^{N-1}}$ and the set of paths $\Psi^{(N-2)}(\phi_k^{N-1}, \phi_{k+1}^{N-1})$, where, for any $n<N$ and for any pair of states in $S^{(n)}$: ϕ_k^n, ϕ_{k+1}^n, we define

$$\begin{aligned}
&\Psi^{(n-1)}(m_{\phi_k^n}^{(n-1)}, m_{\phi_{k+1}^n}^{(n-1)}) \\
&\quad := \bigcup_{T=1}^{\infty} \{\phi_0^{n-1}, \phi_1^{n-1},..., \phi_T^{n-1} \in S^{(n-1)}: \phi_0^{n-1} \in m_{\phi_k^n}^{(n-1)}, \\
&\qquad \phi_T^{n-1} \in m_{\phi_{k+1}^n}^{(n-1)}, \phi_t^{n-1} \notin M^{(n-1)}\ \forall t=1,..., T-1 \text{ and} \\
&\qquad I_{[0,T]}^{(n-1)}(\phi^{n-1}) = \Delta^{(n)}(\phi_k^n, \phi_{k+1}^n) + V_n\}
\end{aligned} \tag{6.5}$$

By iterating this argument we have that the transition $\phi_i^N \to \phi_{i+1}^N$ is described by a family of sequences of permanence sets, one sequence for each choice of possible paths at each step of the iteration corresponding to lower and lower levels of renormalization:

$$Q_{y_0^{n_0}}, Q_{y_1^{n_1}},..., Q_{y_L^{n_L}}$$

For $n_i=0$, $Q_{y_i^{n_i}}$ reduces to the single point $Q_{y_i^{n_i}} \equiv y_i^0$ and $C_{y_i^0} = y_i^0$.

Let us analyze the sequences $\mathbf{y} = y_0^{n_0},..., y_L^{n_L}$ arising from the above iteration.

At each step of the iteration we insert between two states of order n a sequence of states belonging to $S^{(n-1)}$. For this reason we give the following definition.

Definition 6.1. We will denote by $\mathbf{y}^\mathbf{n}$ a sequence of states $(\mathbf{y}^\mathbf{n})_i^{n_i} \in S^{(n_i)}$, $i=1,..., T$, such that $n_i \geqslant n$, $\forall i=1,..., T$. Given $\mathbf{y}^\mathbf{n}$, a sequence

$\mathbf{y^{n-1}}$ is called a *refinement* of $\mathbf{y^n}$ and we write $\mathbf{y^{n-1}}\, ref\, \mathbf{y^n}$ if the following two conditions hold:

(i) $\mathbf{y^{n-1}}$ is a set of states in $\{S^{(n_i)}\}$, with $n_i \geqslant n-1$, containing the set $\mathbf{y^n}$.

(ii) It is obtained by inserting on the r.h.s. of each element $(\mathbf{y^n})_i^{n_i}$ with $n_i = n$ a sequence $\mathbf{y_i^{n-1}}$ of elements of $S^{(n-1)}$ such that there exist x^{n-1} and x^n with

$$x^{n-1} \in m_{x^n}^{(n-1)} \qquad \text{and} \qquad \mathscr{E}_{x^n}^{n} \subseteq \mathscr{E}_{(\mathbf{y^n})_{i+1}}^{n_i+1}$$

and the sequence

$$\{\mathbf{y_i^{n-1}}, x^{n-1}\} \in \Psi^{(n-1)}(m_{(\mathbf{y^n})_i^{n_i}}^{(n-1)}, m_{x^n}^{(n-1)})$$

Such sequences $\mathbf{y_i^{n-1}}$ that we add to $\mathbf{y^n}$ in order to get $\mathbf{y^{n-1}}$ will be called the *refining paths* of $\mathbf{y^{n-1}}$.

Each sequence $\mathbf{y}$ arising from the above iterative application of Theorem 5.1 will be a refinement of a refinement ... of a refinement of an $\mathbf{y^N} \in \Psi^{(N)}(x^N, Q^c)$.

Now for any given constant $a > 0$ and for each state $x^n \in S^{(n)}$ we define, as in Theorem 5.1 the following events (sets of paths):

$$A_1(x^n, a) := \{\phi \in \Phi(\mathbf{S})\colon \phi_0^{(n-1)} \in m_{x^n}^{(n-1)} \text{ and } Z_{\sigma^{(n-1)}-1}^{n-1} \in (T_n e^{-a\beta}, T_n e^{a\beta})\}$$

$$A_2(x^n) := \{\phi \in \Phi(\mathbf{S})\colon \phi_0^{(n-1)} \in m_{x^n}^{(n-1)} \text{ and } Z_{\sigma^{(n-1)}}^{n-1} \geqslant \tau_{(Q_{x^n})^c} > Z_{\sigma^{(n-1)}-1}^{n-1}\}$$

where

$$\sigma^{(n-1)} = \sigma^{(n-1)}(\phi) := \tau_{(m_{x^n}^{(n-1)})^c}^{(n-1)}$$

The first step of our iterative construction is given by the following expression for $\mathscr{T}_N(x^N, Q^c)$:

$$\mathscr{T}_N(x^N, Q^c) = \bigcup_{\mathbf{y^N} \in \Psi^{(N)}(x^N, Q^c)} \mathscr{T}(\mathbf{y^N}, T(\mathbf{y^N})) \tag{6.6}$$

Then we define

$$\mathscr{T}_{N-1}(x^N, Q^c, a) := \bigcup_{\substack{\mathbf{y^N} \in \\ \Psi^{(N)}(x^N, Q^c)}} \; \bigcup_{\mathbf{y^{N-1}}\, ref\, \mathbf{y^N}} \mathscr{T}_{\mathbf{y^N}, \mathbf{y^{N-1}}}$$

where

$$\mathscr{T}_{\mathbf{y}^N,\mathbf{y}^{N-1}} := \{\phi \in \mathscr{T}_N(x^N, Q^c)\text{: for each transition } \phi_i^{(N)} = (\mathbf{y}^N)_i \to \phi_{i+1}^{(N)} = (\mathbf{y}^N)_{i+1} \text{ the path } \phi \text{ on the corresponding interval of time } [Z_i^N, Z_{i+1}^N] \text{ belongs to } A_1((\mathbf{y}^N)_i, a) \cap A_2((\mathbf{y}^N)_i) \text{ and } \phi_t \in \mathscr{T}(\mathbf{y}_i^{N-1}, T(\mathbf{y}_i^{N-1}))\ \forall t > Z_i^N + Z_{\sigma^{(N-1)}-1}^{N-1}\}$$

where $\mathbf{y}_i^{N-1}$ are the refining paths of $\mathbf{y}^{N-1}$.

By iteration

$$\mathscr{T}_n(x^N, Q^c, a) := \bigcup_{\substack{\mathbf{y}^N \in \\ \psi^{(N)}(x^N, Q^c)}} \bigcup_{\mathbf{y}^{N-1}\ ref\ \mathbf{y}^N} \cdots \bigcup_{\mathbf{y}^n\ ref\ \mathbf{y}^{n+1}} \mathscr{T}_{\mathbf{y}^N,\mathbf{y}^{N-1},\dots,\mathbf{y}^n}$$

with

$$\mathscr{T}_{\mathbf{y}^N,\mathbf{y}^{N-1},\dots,\mathbf{y}^n} := \{\phi \in \mathscr{T}_{\mathbf{y}^N,\mathbf{y}^{N-1},\dots,\mathbf{y}^{n+1}}\text{: for each transition of a refining path of } \mathbf{y}^{n+1}\text{: } \phi_i^{(n+1)} = (\mathbf{y}_k^{n+1})_i \to \phi_{i+1}^{(n+1)} = (\mathbf{y}_k^{n+1})_{i+1} \text{ the path } \phi \text{ on the corresponding interval of time belongs to } A_1((\mathbf{y}^{n+1})_i, a) \cap A_2((\mathbf{y}^{n}{}_{+}{}^{1})_i) \text{ and } \phi_t \in \mathscr{T}(\mathbf{y}_k^n, T(\mathbf{y}_k^n))\ \forall t > Z_k^{n+1} + Z_{\sigma^{(N-1)}-1}^{n}\}$$

where $\mathbf{y}_k^n$ are the refining paths of $\mathbf{y}^n$.

Finally we obtain

$$\begin{aligned}\mathscr{T}_0(x^N, Q^c, a) &:= \bigcup_{\substack{\mathbf{y}^N \in \\ \psi^{(N)}(x^N, Q^c)}} \bigcup_{\mathbf{y}^{N-1}\ ref\ \mathbf{y}^N} \cdots \bigcup_{\mathbf{y}^0\ ref\ \mathbf{y}^1} \mathscr{T}_{\mathbf{y}^N,\mathbf{y}^{N-1},\dots,\mathbf{y}^0} \\ &= \bigcup_{\substack{\mathbf{y}^N \in \\ \psi^{(N)}(x^N, Q^c)}} \bigcup_{\mathbf{y}^{N-1}\ ref\ \mathbf{y}^N} \cdots \bigcup_{\mathbf{y}^0\ ref\ \mathbf{y}^1} \{\phi \text{ visits the ordered sequence of sets } Q_{\mathbf{y}^0} \text{ and in each set } Q_{y_i^{n_i}} \text{ spends a time in } [T_{n_i}e^{-a\beta}, T_{n_i}e^{a\beta}]\}\end{aligned} \tag{6.7}$$

where we define $Q_{x^0} = x^0$ and $T_0 = 1$ if $x^0 \in S$.

A first result on the typical exiting tube can thus be stated as follows:

Theorem 6.1. For each positive a there exists a positive constant K such that for each sufficiently large β and for each $x \in \mathscr{E}^N_{x^N}$ we have

$$P_x(X_t \in \mathscr{T}_0(x^N, Q^c, a)) \geqslant 1 - e^{-K\beta}$$

Proof. The proof follows immediately by Theorem 5.1 if we note that for some positive K', K''

$$\begin{aligned} &P_x(X_t \in \mathscr{T}_n(x^n, Q^c, a)) \\ &\quad = \sum_{\substack{\mathbf{y}^N \in \\ \psi^{(N)}(x^N, Q^c)}} \sum_{\mathbf{y}^{N-1}\, ref\, \mathbf{y}^N} \cdots \sum_{\mathbf{y}^{n+1}\, ref\, \mathbf{y}^{n+2}} \\ &\qquad \times P_x(X_t \in \mathscr{T}_n(x^n, Q^c, a) \mid X_t \in \mathscr{T}_{\mathbf{y}^N \ldots \mathbf{y}^{n+1}})\, P_x(X_t \in \mathscr{T}_{\mathbf{y}^N \ldots \mathbf{y}^{n+1}}) \\ &\quad \geqslant (1 - e^{-K'\beta})\, P_x(X_t \in \mathscr{T}_{n+1}(x^n, Q^c, a)) \end{aligned}$$

and $P_x(X_t \in \mathscr{T}_N(x^n, Q^c, a)) \geqslant 1 - e^{-K''\beta}$ by (6.2). ∎

Let us now make some remarks on this theorem.

Given the sequence $\mathbf{y} = \mathbf{y}^0$, consider each element $y^{n_j}_j$ belonging to a portion of the sequence where n_j is monotonically decreasing, i.e., such that

$$\mathscr{E}^{n_j}_{y^{n_j}_j} \subset \mathscr{E}^{n_{j-1}}_{y^{n_{j-1}}_{j-1}} \tag{6.8}$$

By construction, (6.8) implies that

$$n_j < n_{j-1} \text{ and } Q_{y^{n_j}_j} \subset Q_{y^{n_{j-1}}_{j-1}}$$

This means that the information given by the event $A_1 \cap A_2$

$$\tau_{Q^{n_j}_{y_j}} \asymp T_{n_j} \tag{6.9}$$

can be considered as a "negligible correction" to the statement

$$\tau_{Q^{n_{j-1}}_{y_{j-1}}} \asymp T_{n_{j-1}}$$

Then from any sequence of the form

$$\mathbf{y} \in \bigcup_{\substack{\mathbf{y}^N \in \\ \psi^{(N)}(x^N, Q^c)}} \bigcup_{\mathbf{y}^{N-1}\, ref\, \mathbf{y}^N} \cdots \bigcup_{\mathbf{y}^0\, ref\, \mathbf{y}^1} \mathbf{y}^0 \tag{6.10}$$

we can extract a new sequence $\mathbf{y}'$ in which the terms satisfying (6.8) have been canceled.

Moreover, by construction, each element $y_j^{n_j}$ appearing in $\mathbf{y}$ given by (6.10) either is in $S^{(n_j)}\backslash M^{(n_j)}$ [see the definition of the sets $\Psi^{(n)}(x^{n+1}, y^{n+1})$], and so by applying Theorem 5.1 without losing in probability we can also include the event A_4 in the definition of our tube, or $y_j^{n_j}$ is the starting point of the path starting at $y_{j-1}^{n_j-1}$, i.e., (6.8) holds.

Following this remark let us extract from the sequence $\mathbf{y}$ sequences $\{\mathbf{y}''\}_i$ in which (6.8) holds, or, more generally, the maximal segment of $\mathbf{y}$ in which there is a term of the segment such that its permanent set contains all the permanents sets of the segment:

Definition 6.2. Let $\{\mathbf{y}''\}_i$ be a maximal segment of $\mathbf{y}$:

$$\{\mathbf{y}''\}_i := \{\mathbf{y}_{l_i}^{n_{l_i}}, \mathbf{y}_{l_i+1}^{n_{l_i+1}}, ..., \mathbf{y}_{m_i}^{n_{m_i}}\}$$

such that for any $j \in [l_i, m_i]$ we have

$$Q_{y_j^{n_j}} \subseteq Q_{y_{k_i}^{n_{k_i}}} \qquad \text{for some} \quad k_i \in [l_i, m_i]$$

Let $\mathbf{y}'$ be the sequence obtained from $\mathbf{y}$ by replacing each segment $\{\mathbf{y}''\}_i$ with the unique term $y_{k_i}^{n_{k_i}}$.

This means that by considering the sequence $\mathbf{y}'$ instead of $\mathbf{y}$, we have canceled all the "negligible corrections" like (6.9).

Definition 6.3. Given a sequence $\mathbf{y}$ appearing in (6.7) and the corresponding sequence $\mathbf{y}'$ obtained according to Definition 6.2, we call $Q_{\mathbf{y}'}$ the *standard descent* outside Q emerging from x^N.

With Theorem 6.1 we have described in full detail our exit; however, with the refinement procedure we have obtained at the same time two different kinds of information. First of all, we have obtained what we call *primary information* (or *level-one approximation of the tube of exit*): the family of ordered sequences of different permanence sets visited by the process (namely the set of standard descents $Q_{\mathbf{y}'}$) and the typical times spent by our process inside them. The *secondary information* (or *level-two approximation of the tube of exit*) concerns further details about the history of the process inside each cycle that it visits when it stays inside some Q_i in $Q_{\mathbf{y}'}$. It will describe the first descent to the bottom of the cycles and the first excursion outside them.

Let us discuss in some more detail the primary and the secondary information on the tube contained respectively in the parts $\mathbf{y}'$ and $\mathbf{y}''$ of the sequences $\mathbf{y}$ appearing in Theorem 6.1.

We want to characterize the standard descents $Q_{\mathbf{y}'}$. We denote by ω_i the subsequences of sets in the sequence $Q_{\mathbf{y}'}$ coinciding with single states and by using a simpler enumeration of the permanence sets we write

$$Q_{\mathbf{y}'} \equiv Q_1, \omega_1, Q_2, ..., Q_k, \omega_k$$

Proposition 6.1. $Q_1, \omega_1, Q_2, ..., Q_k, \omega_k$ is such that:

(a) $Q_i \cap Q_{i+1} = \varnothing\ \forall i$.

(b) ω_i are downhill paths [i.e., $\omega = x_1, x_2, ..., x_m$ satisfies $\Delta(x_j, x_{j+1}) = 0$, $\forall i = 1, 2, ..., m$] going from Q_i to Q_{i+1} in the following sense: if $\omega_i \equiv x_{i,1}, x_{i,2}, ..., x_{i,m}$, then $x_{i,1}$ is one of the optimal exit points (belonging to the set Y_x of Proposition 3.2) of Q_i and $x_{i,m} \in Q_{i+1}$.

(c) ω_k goes from Q_k to Q^c.

Proof. We know that $\mathbf{y}^{\mathbf{N}} \in \Psi^N(x^N, Q^c)$, i.e., $\Delta^{(N)}((\mathbf{y}^{\mathbf{N}})_i, (\mathbf{y}^{\mathbf{N}})_{i+1}) = 0$. We will iteratively prove that if $\Delta^{(n)}((\mathbf{y}^{\mathbf{n}})_i, (\mathbf{y}^{\mathbf{n}})_{i+1}) = 0$, then the refining $\mathbf{y}_i^{\mathbf{n}-1}$, i.e., the subsequence of elements in $S^{(n-1)}$ belonging to $\mathbf{y}$ between $(\mathbf{y}^{\mathbf{n}})_i$ and $(\mathbf{y}^{\mathbf{n}})_{i+1}$, can be divided into three parts:

(i) A first segment, say $y_{i,1}^{n-1}, ..., y_{i,m}^{n-1}$, in $Q_{(\mathbf{y}^{\mathbf{n}})_i}$.

(ii) A second segment, say $y_{i,m+1}^{n-1}, ..., y_{i,l}^{n-1}$, downhill in $\mathscr{B}_{(\mathbf{y}^{\mathbf{n}})_i}$.

(iii) The last segment $y_{i,l+1}^{(n-1)}, ..., y_{i,T}^{n-1}$ in $Q_{(\mathbf{y}^{\mathbf{n}})_{i+1}}$.

We notice that each of these parts can be empty. The first and the last segments consist, respectively, in the ascent outside $Q_{(\mathbf{y}^{\mathbf{n}})_i}$ and the descent to the bottom of $Q_{(\mathbf{y}^{\mathbf{n}})_{i+1}}$; they will be studied later when we analyze the secondary information. The only remaining part in our level-one analysis, i.e., in the sequence $\mathbf{y}'$, is part (ii), and so the proposition follows once we prove the iteration step.

We know, by the definition of $\mathbf{y}$, that $\mathbf{y}_i^{\mathbf{n}-1}$ is a path going from $(\mathbf{y}^{\mathbf{n}})_i$ to $(\mathbf{y}^{\mathbf{n}})_{i+1}$ and minimizing the functional $I^{(n-1)}$. More precisely, we know that $I^{(n-1)}(\mathbf{y}_i^{\mathbf{n}-1}) = \Delta^{(n)}((\mathbf{y}^{\mathbf{n}})_i, (\mathbf{y}^{\mathbf{n}})_{i+1}) + V_n = V_n$ and, on the other hand, each path in $S^{(n-1)}$ exiting from $m_{(\mathbf{y}^{\mathbf{n}})_i}^{(n-1)}$ gives to the functional $I^{(n-1)}$ a value larger than or equal to V_n and so the part of the refining path outside $Q_{(\mathbf{y}^{\mathbf{n}})_i}$ [the second segment (ii)] must give a zero contribution to the functional $I^{(n-1)}$. ▮

Definition 6.4. Given a standard descent $Q_1, \omega_1, Q_2, ..., Q_k, \omega_k$ emerging from $x \in \mathscr{E}_{x^N}^N$, we say that our process follows *regularly* $Q_1, \omega_1, Q_2, ..., Q_k, \omega_k$ if:

1. It stays inside Q_1 during a suitable random interval of time, exponentially long in β.

2. The length of this random interval of time as well as the way our process spends its time in Q_1 is specified as follows: during its permanence in Q_1 our process visits one or more cycles $C_{1,k}$ belonging to Q_1. When it enters inside one of the cycles $C_{1,k}$ it has the typical behavior described in Propositions 3.1, 3.2, 4.3, and 4.4; in particular it visits all the points before

exiting from a point y minimizing the quantity $W_Q(x, y)$ in (3.8); it stays in $C_{1,k}$ a random time logarithmically equivalent to its expectation given in (3.10).

3. Then the process gets out of Q_1 following ω_1.
4. Subsequently it enters Q_2 and continues as before.
5. Finally, following ω_k, it gets out of Q.

Our main result about the level-one approximation is contained in the following theorem, which immediately follows from Theorem 6.1 and Proposition 6.1:

Theorem 6.2. With probability tending to one as β tends to infinity, if at $t=0$ our process starts from $x \in \mathscr{E}^N_{x^N}$, the first exit from Q follows regularly one of the possible descents emerging from x.

Sometimes the primary information can be completely simplified. This happens, for instance, if Q itself is a single cycle.

In this case the primary information clearly reduces to a triviality: we have that the descent consists only in the set Q itself and the typical point of first exit in ∂Q. In other words, our primary information in this case is almost empty since it only gives us the typical random time spent in Q and the typical point of first exit in ∂Q. As far as the problem of the determination of the tube of exit is concerned, in the present case only what we have called the level-two approximation starts to be interesting.

Indeed the natural question that arises in this case concerns the typical first excursion outside Q, i.e., the path

$$x_{\theta_{x^N}}, x_{\theta_{x^N}+1}, \ldots, x_{\tau_{Q^c}}$$

where

$$\theta_{x^N} = \sup\{t < \tau_{Q^c}: X_t = x_N\}$$

The level-two approximation, in particular, will characterize this first excursion outside Q.

Let us consider the simplest case in which Q is a cycle whose bottom is made up of a single point x^N. We have $\mathbf{y}^N = x^N, z^N$ for some $z^N \in Q^c$ with $\Delta^{(N)}(x^N, z^N) = 0$ and

$$\mathbf{y}^{N-1} = x^N, x_1^{N-1}, \ldots, x_T^{N-1}, z^N$$

$$\mathbf{y}^{N-2} = x^N, x_1^{N-1}, x_{1,1}^{N-2}, \ldots, x_{1,T_1}^{N-2}, x_2^{N-1}, x_{2,1}^{N-2}, \ldots, x_{2,T_2}^{N-2}, \ldots, x_T^{N-1}, x_{T,1}^{N-2}, \ldots, z^N$$

The whole sequence $\mathbf{y}$ arising from the refinement procedure, except for the final state z^N, satisfies the condition defining the sequences $\{\mathbf{y}''\}_i$, since in this case

$$Q_{x^N} \supset Q_{y_i^{n_i}}, \qquad \forall y_i^{n_i} \in \mathbf{y}$$

On the other hand, if we want to get information about the first excursion, i.e., about the behavior of the chain in the time interval $[\theta_{x^N}, \tau_{Q^c}]$ (which is exponentially smaller than the time spent by the process in Q), we are exactly interested in the "negligible corrections" contained in $\{\mathbf{y}''\}_i$. Thus we can ignore the first term x^N in the sequence $\mathbf{y}$, describing exactly the long time interval spent in Q, and we can consider, as before, the remaining sequence $\mathbf{y}\backslash x^N$. Again we can extract from this a sequence $(\mathbf{y}\backslash x^N)'$ (see Definition 6.2) and we can apply the analysis developed up to now to this beheaded sequence. We obtain in this way a sequence of permanent sets $Q_{\{\mathbf{y}\backslash x^N\}'}$ describing the ascent from x^N to Q^c.

In the reversible case this ascent can be characterized by a sequence of disjoint sets Q's connected by uphill sequences ω_i. In the general nonreversible case this result is no longer true and more and more complicated paths ω_i (partially uphill and partially downhill) will appear in a standard ascent. On the other hand, we notice that in the case of standard descents the simple feature of the corresponding ω_i (only downhill paths) valid in the reversible case is preserved in the general case, as follows from Proposition 6.1.

The procedure of beheading applied to the discussion of the case when Q is a cycle with a single bottom x^N can also be applied to describe the segments $\{\mathbf{y}\}_i''$ in the case of a general Q, i.e., to obtain the secondary information.

Let $\{\mathbf{y}''\} = \{y_j^{n_j}\}_{j=l}^{m}$ and let k be such that

$$Q_{y_k^{n_k}} \supseteq Q_{y_j^{n_j}} \qquad \text{for each} \quad j = l, \ldots, m$$

Since in each standard descent each element of $\mathbf{y}$ appears only once, in each segment $\{\mathbf{y}''\}$ there is only one k satisfying the previous inclusion. We can thus define

$$\mathbf{y}_d = \{y_j^{n_j}\}_{j=l,\ldots,k-1} \qquad \text{and} \qquad \mathbf{y}_a = \{y_j^{n_j}\}_{j=k+1,\ldots,m}$$

so that $\{\mathbf{y}''\} = \mathbf{y}_d, y_k^{n_k}, \mathbf{y}_a$, where $\mathbf{y}_d$ and $\mathbf{y}_a$ correspond respectively to the descent to the bottom $y_k^{n_k}$ and to the ascent from $y_k^{n_k}$ to Q_{y}^c. For each of these segments we can extract again the subsequences $(\mathbf{y}_d)'$ and $(\mathbf{y}_a)'$.

In other words, the level-two approximation on the tube tells us how our process reaches for the first time the bottom of the visited cycles $C_{i,k}$ constituting the permanent set $Q_{y_i^{n_i}}$ and how it performs its first excursion from this bottom to $Q^c_{y_i^{n_i}}$. Between these relatively rapid "transients" the system will typically spend a much (exponentially in β) longer time performing random oscillations in $C_{i,k}$ visiting many times all its points before exiting. Of course for the degenerate permanence sets given by single points the secondary information loses any sense and we have to stop our analysis. In general we can continue our description of the tube of first exit trajectories by specifying the level-three approximation, namely by describing the first and last transients of the history of our process when it enters some smaller cycle. This means that each segment $\{\mathbf{y}''\}$ contained in $(\mathbf{y}_d)'$ and $(\mathbf{y}_a)'$ can again be analyzed in the same way.

Up to now we have considered a starting point $x \in \mathscr{E}^N_{x^N}$. We now discuss the case of a general starting point.

For any set Q, let $N = N(Q)$ be defined as usual [see (2.31)]. Consider the covering of S given by the union of the generalized basins of attraction of all the points $x^{(N)} \in S^{(N)}$:

$$S = \bigcup_{x^{(N)} \in S^{(N)}} \bar{\mathscr{B}}(x^{(N)})$$

Given any $x \in Q$, we look at an $x^{(N)}$ such that $x \in \bar{\mathscr{B}}(x^{(N)})$. If some of those $x^{(N)}$ do not belong to Q, there are possible typical first excursions outside Q starting from x which are just descents to $x^{(N)} \notin Q$.

Suppose now that there exists at least a point $x^{(N)} \in S^{(N)} \cup Q$ such that $x \in \bar{\mathscr{B}}(x^{(N)})$. We will show that in this case we can distinguish a first interval of time in which the process falls to x^N, and a second, much longer interval of time in which the process exits from Q following the description obtained above, since we can now consider x^N as starting point. This means that in order to complete our description of the typical exiting tube we have to study this first transient descent to the bottom of a basin of attraction.

More precisely, let $x^n \in S^{(n)}$ and $x \in \mathscr{B}^n_{x^n}$. We now apply iteratively Theorems 5.2 and 5.3 to obtain the tube of typical paths starting at x up to the first fall to x^n.

The sequence of states $\mathbf{y}$ in this case has the following property:

$$(\mathbf{y}^{\mathbf{N}-\mathbf{1}})_0 \in B^{(N-1)}_{x^N}$$

and there exists $x^{N-1} \in m^{(N-1)}_{x^N}$ such that

the path $\{\mathbf{y}^{\mathbf{N}-\mathbf{1}}, x^{N-1}\} \in \{\phi^{(N-1)}\}_{i=1}^{T(\phi^{(N-1)})} \in \Phi(S^{(N-1)})$:

$$I^{(N-1)}(\phi_i^{(N-1)}, \phi_{i+1}^{(N-1)}) = 0\} \qquad (6.11)$$

Moreover, we can define the sequence of states arising from this iteration by defining a new refining as follows:

Definition 6.5. For each $n < N-1$, $\mathbf{y}^n$ *is a fall-refining* of $\mathbf{y}^{n+1}$ and we write $\mathbf{y}^n$ *f-ref* $\mathbf{y}^{n+1}$ if $\mathbf{y}^n$ is obtained by inserting on the l.h.s. of each element $(\mathbf{y}^{n+1})_i$ a sequence of states in $S^{(n)}$: $\mathbf{y}_i^n$ such that there exist x_i^{n+1}, x_i^n such that

$$x_i^n \in m_{x_i^{n+1}}^{(n)}, \qquad \mathscr{E}_{x_i^{n+1}}^{n+1} \subseteq \mathscr{E}_{(\mathbf{y}^{n+1})_i^{n_i}}^{n_i}, \qquad (\mathbf{y}_i^n)_0 \in B_{x^{n+1}}^{(n)}$$

and the path

$$\{\mathbf{y}_i^n, x_i^n\} \in \{\{\phi^{(n)}\}_{i=1}^{T(\phi^{(n)})} \in \Phi(S^{(n)}) \colon I^{(n)}(\phi_i^{(n)}, \phi_{i+1}^{(n)}) = 0\}$$

As before, we can define the tube

$$\begin{aligned}\mathscr{T}_0(x, x^N, a) :&= \bigcup_{\substack{\mathbf{y}^{N-1} \\ \text{satisfying (6.11)}}} \ \bigcup_{\mathbf{y}^{N-2} f\text{-ref}\, \mathbf{y}^{N-1}} \cdots \bigcup_{\mathbf{y}^0 f\text{-ref}\, \mathbf{y}^1} \mathscr{T}_{\mathbf{y}^N, \mathbf{y}^{N-1} \dots \mathbf{y}^0} \\ &= \mathscr{T}_0(x, x^N, a) \\ &:= \bigcup_{\substack{\mathbf{y}^{N-1} \\ \text{satisfying (6.11)}}} \ \bigcup_{\mathbf{y}^{N-2} f\text{-ref}\, \mathbf{y}^{N-1}} \cdots \bigcup_{\mathbf{y}^0 f\text{-ref}\, \mathbf{y}^1} \\ &\quad \{\phi \text{ visits the ordered sequence of sets } Q_{\mathbf{y}^0} \\ &\quad \text{and in each set } Q_{y_i^{n_i}} \text{ spends a time in } [T_{n_i} e^{-a\beta}, T_{n_i} e^{a\beta}]\}\end{aligned}$$

Analogously to Theorem 6.1 we can prove the following result.

Theorem 6.3. For any $a > 0$ there exists $K > 0$ such that for any β sufficiently large and for any $x \in \mathscr{B}_{x^N}^N$

$$P_x(X_t \in \mathscr{T}(x, x^N)) \geq 1 - e^{-K\beta}$$

If

$$x \in \bar{\mathscr{B}}_{x_i^N}^N, \qquad i = 1, \dots, j(x)$$

then

$$P_x\left(X_t \in \bigcup_{i=1}^{j(x)} \mathscr{T}(x, x_i^N)\right) \geq 1 - e^{-K\beta}$$

By combining Theorems 6.1 and 6.3 we complete our description of the tube of typical paths starting from each point x up to the first exit from Q.

Indeed, if $x \in \mathscr{B}^N_{x^N}$ for some $x^N \in Q$, then the typical tube is just made up of paths which are in $\mathscr{T}(x, x^N, a)$ up to the time $\tau_{\mathscr{B}^N_{x^N}}$ and from this moment they are in the tube $\mathscr{T}(x^N, Q^c, a)$. We call it $\mathscr{T}(x, x^N, a) \circ \mathscr{T}(x^N, Q^c, a)$. If

$$x \in \bar{\mathscr{B}}^N_{x_i^N}, \qquad i = 1,..., j(x)$$

then we have to consider the tube

$$\bigcup_{i \in [1,..., j(x)]:\, x_i^N \in Q} \mathscr{T}(x, x_i^N, a) \circ \mathscr{T}(x_i^N, Q^c, a) \qquad \bigcup_{l \in [1,..., j(x)]:\, x_l^N \in Q^c} \mathscr{T}(x, x_l^N, a)$$

Similarly to what we did before, we can extract from the final sequence $\mathbf{y}$ a sequence $\mathbf{y}'$ and distinguish between primary and secondary information.

APPENDIX

Proof of Lemma 4.1.

The proof is organized as follows:

1. First we consider a graph $g^{(k)*}_{Q^k}$ and we construct a graph $\bar{g}$ on $S^{(k-1)}$ satisfying property 2' of Definition 3.1 of B-graphs.
2. As a second step we extract from $\bar{g}$ a Q^{k-1} graph on $S^{(k-1)}$: $g^{(k-1)}_{Q^{k-1}}$.
3. We show that

$$W^{(k-1)}_{Q^{k-1}} \leqslant W(g^{(k-1)}_{Q^{k-1}}) = W^{(k)}_{Q^k} + V_k\, |Q^k|$$

4. Given a graph $g^{(k-1)*}_{Q^{k-1}}$, we construct a Q^k-graph on $S^{(k)}$: $g^{(k)}_{Q^k}$.
5. We verify that

$$W^{(k)}_{Q^k} \leqslant W(g^{(k)}_{Q^k}) = W^{(k-1)}_{Q^{k-1}} - V_k\, |Q^k|$$

From 3 and 5 we immediately obtain (4.20) and thus we can conclude that the graph $g^{(k-1)}_{Q^{k-1}}$ constructed in 1 and 2 minimizes $W^{(k-1)}_{Q^{k-1}}$ and the graph $g^{(k)}_{Q^k}$ constructed in 4 minimizes $W^{(k)}_{Q^k}$.

1. By Eq. (4.19) we have that Q^{k-1} contains all the unstable states of the chain $X^{(k-1)}_t$, say $S^{(k-1)} \backslash M^{(k-1)}$, and each state in $m^{(k-1)}_x$, if $x \in Q^k$. Here Q^{k-1} possibly contains some states in $m^{(k-1)}_x$ with $x \in D^k$, but not all of them.

For any $i \in Q^k$ let $m^{(k-1)}_i$ be the corresponding equivalence class in $S^{(k-1)}$. Given a graph $g^{(k)*}_{Q^k}$, we can now construct a graph $\bar{g}$ of arrows starting in states in Q^{k-1} as follows:

(i) To each arrow in $g_{Q^k}^{(k)*}$: $i \to j$ we associate a sequence of arrows $x_1^{k-1} \to \cdots \to x_l^{k-1}$ between states in $S^{(k-1)}$ such that $x_1^{k-1} \in m_i^{(k-1)}$, $x_l^{k-1} \in m_j^{(k-1)}$, $x_j^{k-1} \in S^{(k-1)} \backslash M^{(k-1)}$, and $x_1^{k-1} \to \cdots \to x_l^{k-1}$ is a path minimizing $\Delta^{(k)}(i, j)$ [see (2.18) and (2.9)].

(ii) For any equivalent class $m_i^{(k-1)}$, $i \in Q^k$, let i_0^{k-1} be a state which is the starting point of an arrow constructed in the previous step (i). For any other state in $m_i^{(k-1)}$ we can construct a sequence of arrows between points in $m_i^{(k-1)}$ leading to i_0^{k-1}. Such a sequence exists by definition of $\sim^{(k-1)}$ and for each such arrow the associated function $\Delta^{(k-1)}$ is zero.

(iii) Since $D^{k-1} \subset \bigcup_{x \in D^k} m_x^{(k-1)}$, it may happen that there are stable states in Q^{k-1} (for the chain $X_t^{(k-1)}$) which were not touched by the construction (i) and (ii). They must be equivalent to some state in D^{k-1} and we can then draw a sequence of arrows from each of them to a state in D^{k-1} with a zero contribution to the quantity $W(\bar{g})$.

(iv) Let us now consider the states in Q^{k-1} which were not touched by the previous construction, i.e., which are not starting points of any arrow constructed in steps (i)–(iii). These are unstable states. In this set of states we consider the equivalence classes with respect to the relation $\sim^{(k-1)}$, c_i^{k-1}. For each such equivalent class, due to unstability, we can draw an arrow emerging from a state contained in it and ending outside it and corresponding to a zero value of the function $\Delta^{(k-1)}$. Moreover, for each of these equivalence classes we can find a sequence of such arrows of zero cost and leading to a state in D^{k-1} or to a state considered in (i), (ii), or (iii) (i.e., starting point of an arrow already drawn). Inside each equivalence class c_i^{k-1} we can draw arrows between equivalent states as in point (ii).

In this way we have constructed a set of arrows, say $\bar{g}$, such that at least one arrow emerges from each state in Q^{k-1} and condition 2′ of Definition 3.1 of a Q^{k-1}-graph for the chain $X_t^{(k-1)}$ is satisfied.

2. We will show now that $\bar{g}$ contains a Q^{k-1}-graph: $g_{Q^{k-1}}^{(k-1)}$. This is a standard proof (ref. 13, Theorem 1). In fact by the previous remark we have only to show that we can satisfy also condition 1 of Definition 3.1 of B-graphs only by removing arrows in $\bar{g}$. This can be done with the following prescription:

(a) Introduce in the set of states $S^{(k-1)} \backslash D^{k-1}$ in an arbitrary order $x_1, x_2, \ldots$; set $g_0 = \bar{g}$ and $i = 1$.

(b) Since g_{i-1} satisfies condition 2′, there is at least a sequence of arrows in g_{i-1} leading from x_i to some point in D^{k-1}; choose one of these sequences; let $x_i \to x_i'$ be its first arrow.

(c) Define g_i as the set of arrows constructed starting from g_{i-1} by erasing all the arrows exiting from x_i different from $x_i \to x_i'$.

(d) Verify that g_i satisfies condition 2′ and every $x_j, j \leqslant i$, is the initial point of exactly one arrow in g_i.

(e) Make $i \to i+1$ and go back to point (b).

The graph

$$g_{Q^{k-1}}^{(k-1)} \equiv g_i \qquad \text{for} \quad i = |Q^{k-1}|$$

is by definition a Q^{k-1}-graph included in $\bar{g}$, since it satisfies conditions 1 and 2′.

3. By using the definition of $\Delta^{(k)}(\cdot, \cdot)$ we have

$$\begin{aligned} W_{Q^{k-1}}^{(k-1)} \leqslant W(g_{Q^{k-1}}^{(k-1)}) &= \sum_{i \to j \in g_{Q^{k-1}}^{(k-1)}} \Delta^{(k-1)}(i,j) \leqslant \sum_{i \to j \in \bar{g}} \Delta^{(k-1)}(i,j) \\ &\leqslant \sum_{i \to j \in g_{Q^k}^{(k)*}} (\Delta^{(k)}(i,j) + V_k) = W_{Q^k}^{(k)} + V_k\,|Q^k| \end{aligned}$$

4. Let $g_{Q^{k-1}}^{(k-1)*}$ be a Q^{k-1}-graph minimizing $W_{Q^{k-1}}^{(k-1)}$. For each $i \in S^{(k)}$ let $m_i^{(k-1)}$ be its equivalent class. If $m_i^{(k-1)} \cap D^{k-1} = \varnothing$, then $i \in Q^k$ and in $g_{Q^{k-1}}^{(k-1)*}$ there is a unique arrow exiting from $m_i^{(k-1)}$. Indeed, if there were two arrows exiting from $m_i^{(k-1)}$ in $g_{Q^{k-1}}^{(k-1)*}$ we could construct a new Q^{k-1}-graph $g'^{(k-1)}$ by changing only arrows starting in $m_i^{(k-1)}$ with

$$W(g'^{(k-1)}) < W(g_{Q^{k-1}}^{(k-1)*}) = W_{Q^{k-1}}^{(k-1)}$$

contradicting the fact that $W_{Q^{k-1}}^{(k-1)}$ is the minimum.

This means that from the graph $g_{Q^{k-1}}^{(k-1)*}$ we can construct a Q^k-graph on $S^{(k)}$ by looking, for any $i \in Q^k$, at the sequence of arrows exiting from it and at the first state in $M^{(k-1)}$ hit by this sequence. The set of transitions starting in states in Q^k constructed in this way is a Q^k-graph $g_{Q^k}^{(k)}$, since $g_{Q^{k-1}}^{(k-1)*}$ was a $Q^{(k-1)}$-graph.

5. We have immediately

$$\begin{aligned} W_{Q^k}^{(k)} \leqslant W(g_{Q^k}^{(k)}) &= \sum_{i \to j \in g_{Q^k}^{(k)}} \Delta^{(k)}(i,j) \\ &\leqslant W(g_{Q^{k-1}}^{(k-1)*}) - V_k\,|Q_k| \quad \blacksquare \end{aligned}$$

Proof of Lemma 4.2. The proof is based on Lemma 4.1.

For any $x_0 \in \mathscr{E}^n_{x_0^n}$ let $x_0^1 \in S^{(1)}$, $x_0^2 \in S^{(2)}, \ldots, x_0^n \in S^{(n)}$ be the sequence of renormalized states such that $x_0 \in \mathscr{E}^i_{x_0^i}$ $\forall i = 1, \ldots, n$.

Let us first consider the chain at level $n-1$. The stable states of the chain $X_t^{(n-1)}$ contained in $\mathscr{B}^n_{x_0^n}$, say $y_1^{n-1}, y_2^{n-1}, \ldots, y_k^{n-1}$, are all contained in an equivalence class $m_{x_0^n}^{(n-1)}$, and one of them coincides with x_0^{n-1}, i.e., contains the point x_0. If we now consider the quantity

$$W^{(n-1)}_{B^{(n-1)}_{x_0^n} \backslash x_0^{n-1}}$$

we immediately obtain that it vanishes and each $(B^{(n-1)}_{x_0^n} \backslash x_0^{n-1})$-graph on $S^{(n-1)}$ minimizing this quantity has only arrows ending in x_0^{n-1}.

Now, by iteratively applying Lemma 4.1 with $k = n - l$, $l = 1, 2, \ldots, n-1$, and

$$D^{n-1} = (B^{(n-1)}_{x_0^n} \backslash x_0^{n-1})^c, \qquad D^{k-1} = \bigcup_{x \in D^k} m_x^{(k-1)} \backslash x_0^{k-1}$$

we conclude the proof of Lemma 4.2 in the case of a single state.

The proof in the case of a stable plateau $p^{(n)}$ is exactly the same. ∎

ACKNOWLEDGMENTS

We thank Gerard Ben Arous, Raphael Cerf, and Alain Touvé for stimulating discussions. We thank a referee for pointing out refs. 1 and 3. E.S. thanks the Physics Department of Università "La Sapienza" for its kind hospitality. This work was partially supported by grant CHRX-CT93-0411 of the Commission of European Communities.

REFERENCES

1. D. Chen, J. Feng, and M. Qian, The metastability of exponentally perturbed Markov chains, *Chinese Sci. A* **25**(6):590–595 (1995).
2. T. S. Chiang and Y. Chow, A limit theorem for a class of inhomogeneous Markov processes, *Ann. Prob.* **17**:1483–1502 (1989).
3. T. S. Chiang and Y. Chow, On the exit problem from a cycle of simulated annealing processes with application—A backward equation approach, Tech. Rept. Inst. Math. Academia Sinica (1995).
4. T. S. Chiang and Y. Chow, Asymptotic behavior of eigenvalues and random updating schemes, *Appl. Math. Optim.* **28**:259–275 (1993).
5. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer-Verlag, Berlin, 1984).
6. C. R. Hwang and S. J. Sheu, Singular perturbed Markov chains and exact behaviors of simulated annealing process, *J. Theor. Prob.* **5**:223–249 (1992).
7. R. Kotecky and E. Olivieri, Droplet dynamics for asymmetric Ising model, preprint (1992).

8. R. Kotecky and E. Olivieri, Shapes of growing droplets—A model of escape from a metastable phase, in preparation.
9. E. J. Neves and R. H. Schonmann, Behaviour of droplets for a class of Glauber dynamics at very low temperatures, *Commun. Math. Phys.* **137**:209 (1991).
10. E. J. Neves and R. H. Schonmann, Critical droplets and metastability for a Glauber dynamics at very low temperatures, *Prob. Theory Related Fields* **91**:331 (1992).
11. E. Olivieri and E. Scoppola, Markov chains with exponentially small transition probabilities: First exit problem from a general domain—I. The reversible case, preprint.
12. E. Scoppola, Renormalization group for Markov chains and application to metastability, *J. Stat. Phys.* **73**:83 (1993).
13. E. Scoppola, Renormalization and graph methods for Markov chains, in *Proceedings of the Conference "Advances in Dynamical Systems and Quantum Physics"—Capri 1993*, in press.
14. A. Touvé, Cycle decompositions and simulated annealing, preprint LMENS-94.
15. A. Touvé, Rough large deviation estimates for the optimal convergence speed exponent of generalized simulated annealing algorithms, *Ann. Inst. H. Poincaré* (1995).
16. O. Catoni, Rough large deviation estimates for simulated annealing. Application to exponential schedules, *Ann. Prob.* **20**:1109–1146 (1992).
17. R. Z. Has'minskii, *Stochastic Stability of Differential Equations* (Sijthoff and Noordhoff, 1980).