5

A Martingale Approach to Metastability

We introduce in Definition 5.1 below the notion of scaling limit of a sequence of metastable Markov chains. To motivate this definition, we examine in the first section of this chapter the asymptotic behavior of a sequence of birth and death chains. The statement and the proof of the scaling limit of the birth and death chain are postponed to Section 5. In Section 2 we introduce the *metastable* Markov chains and we sketch in Section 3 and 4 the proof of the scaling limit of these chains.

1 A Birth and Death Chain

We examine in this section the asymptotic behavior of a sequence of birth and death chains which is reversible with respect to a Gibbs measure associated to a logarithmic energy. This sentence will be clarified as we introduce the stationary measure of the chain in (1.4). A

The energy. Fix $a < b$ in $\mathbb{R}$ and consider a nonnegative, continuously differentiable function $H : [a, b] \to \mathbb{R}_+$. Assume that H vanishes only at a finite number of points denoted by $a_1 < a_2 < \cdots < a_m$, $m \geq 2$:

$$H(x) = 0 \quad \text{if and only if} \quad x \in A \;:=\; \{a_1, \dots, a_m\}\;.$$

We do not exclude the possibility that H vanishes at the boundary points a, b. δ_0

Fix $\delta > 0$ such that $a_{i+1} - a_i > 4\delta$ for $1 \leq i < m$. Assume that there exist α_i $\alpha_i > 1$, and $0 < \delta_0 < \delta$ such that w_i

$$H(x) \;=\; |x - a_i|^{\alpha_i} \quad \text{for } |x - a_i| \leq 2\delta_0\;,\;\; 1 \leq i \leq m. \tag{1.1}$$

By the continuity of H, there exists $\delta_1 > 0$ such that

$$H(x) \;>\; \delta_1 \quad \text{for } x \notin \bigcup_{i=1}^{m} (a_i - \delta_0, a_i + \delta_0)\;. \tag{1.2}$$

In particular, for any increasing sequence J_N such that $1 \ll J_N \ll N$,

$$w_i \;:=\; \lim_{N\to\infty} \frac{1}{N^{\alpha_i}} \sum_{k=1}^{J_N} \frac{1}{H(a_i + k/N)} \;=\; \lim_{N\to\infty} \frac{1}{N^{\alpha_i}} \sum_{k=1}^{J_N} \frac{1}{H(a_i - k/N)} \;<\; \infty\,. \tag{1.3}$$

Of course, if a_i is one of the boundary points, we only assume the existence of the limit which makes sense. Here and below for two positive, non-decreasing sequences J_N, M_N, $J_N \ll M_N$ means that

$$\lim_{N\to\infty} \frac{J_N}{M_N} \;=\; 0\,.$$

$\ll$
α
b_i

The configuration space. Let $\alpha = \max\{\alpha_i : 1 \le i \le m\}$, and assume that there are at least two exponents α_i equal to α:

$$L \;:=\; \big|\{i : \alpha_i = \alpha\}\big| \;\ge\; 2\,,$$

where $|A|$ indicates the cardinality of a finite set A. Denote by $b_1 < b_2 < \cdots < b_L$ the elements of $\{a_1, \dots, a_m\}$ whose associated exponents are α.

E_N
ζ_i^N
H_N

Let $E_N = \{k/N : [aN] \le k \le [bN]\}$ be the configuration space, let ζ_i^N be the approximation of the critical point a_i, $\zeta_i^N = [a_i N]/N$, and let $H_N : E_N \to \mathbb{R}_+$ be the approximation of the function H, defined by

$$H_N(\zeta_i^N) = N^{-\alpha_i}\,, \quad H_N(\zeta_i^N \pm k/N) \;=\; H(a_i \pm k/N)\,, \quad 1 \le k \le [2\delta_0 N]\,,$$

and $H_N(\eta) = H(\eta)$ otherwise.

Z_N

The stationary state. Define a probability measure π_N on E_N by

$$\pi_N(\eta) \;=\; \frac{1}{Z_N}\,\frac{1}{H_N(\eta)}\,, \quad \eta \in E_N\,, \tag{1.4}$$

where Z_N is the normalizing constant. The measure π_N corresponds to the Gibbs measure associated to the energy $\log H_N$ at temperature 1.

By assumption (1.1) and by definition (1.3),

$$\lim_{N\to\infty} \frac{Z_N}{N^\alpha} \;=\; \sum_{i=1}^{L} \Big\{1 + \sigma_i\, w_i\Big\}\,, \tag{1.5}$$

σ_i
ξ_i^N
ℓ_N
$\mathcal{E}_N^i$

where $\sigma_i = 1$ if $b_i \in \{a, b\}$ and $\sigma_i = 2$ otherwise.

Let $\{\ell_N : N \ge 1\}$ be an increasing sequence of positive integers such that $1 \ll \ell_N \ll N$. For each $1 \le i \le L$, let $\xi_i^N = [b_i N]/N$ and define the subsets $\mathcal{E}_N^i$ of the configuration space E_N by

$$\mathcal{E}_N^i \;:=\; \Big\{\, \xi_i^N - \frac{\ell_N}{N}\,, \,\dots\,,\, \xi_i^N + \frac{\ell_N}{N} \,\Big\}\,.$$

Here again, if $b_1 = a$, $\mathcal{E}_N^1$ is the set $\{\xi_1^N, \dots, \xi_1^N + \ell_N/N\}$, with an analogous convention if $b_L = b$. Since $N^{-1}\ell_N \to 0$, for N large enough, $\mathcal{E}_N^i \cap \mathcal{E}_N^j = \varnothing$ for all $i \neq j$. Let

$$\mathcal{E}_N = \bigcup_{i=1}^{L} \mathcal{E}_N^i \;, \quad \Delta_N \;:=\; E_N \setminus \mathcal{E}_N \;.$$

Since $\ell_N \to \infty$, by (1.3) and (1.5), for all $1 \leq i \leq L$,

$$\lim_{N\to\infty} \pi_N(\Delta_N) \;=\; 0 \quad \text{and} \quad \lim_{N\to\infty} \pi_N(\mathcal{E}_N^i) \;=\; \frac{m(b_i)}{\sum_{j=1}^{L} m(b_j)} \;, \tag{1.6}$$

where $m(b_i) = 1 + \sigma_i\, w_i$. The stationary measure π_N is therefore concentrated on the sets $\mathcal{E}_N^i$. $\mathcal{E}_N$

The dynamics. Fix a positive function $\Phi : [a,b] \to \mathbb{R}_+$ bounded above and below by a strictly positive constant: Δ_N $m(b_i)$

$$0 \;<\; \delta_2 \;\leq\; \Phi(\eta) \;\leq\; \delta_2^{-1} \;.$$

This assumption is not necessary but we do not seek optimal conditions. Consider a birth and death chain $\{\eta^N(t) : t \geq 0\}$ on E_N with jump rates given by

$$\begin{aligned} R_N(\eta, \eta + N^{-1}) \;&=\; \Phi(\eta) \;, \\ R_N(\eta + N^{-1}, \eta) \;&=\; \frac{\pi_N(\eta)\, \Phi(\eta)}{\pi_N(\eta + N^{-1})} \end{aligned}$$

for $[aN]/N \leq \eta < ([bN]-1)/N$, and $R_N(\eta,\xi) = 0$ if $|\xi - \eta| \neq 1/N$. Clearly, $\eta^N(t)$ is Markov chain, reversible with respect to the probability measure π_N. It is called a birth and death chain because it evolves on $\mathbb{Z}$ and it may jump only to the nearest neighbors. It follows from the assumptions (1.1), (1.2) that the jump rates $R_N(\eta,\xi)$ to nearest neighbors sites are bounded below by a positive constant and bounded above by a finite constant. Φ

Asymptotic dynamics. Define the subsets $\mathcal{F}_N(\zeta_i^N)$, $1 \leq i \leq m$, of the configuration space by R_N

$$\mathcal{F}_N(\zeta_i^N) \;:=\; \Big\{\, \zeta_i^N - \frac{\ell_N}{N} \,, \,\dots\,, \,\zeta_i^N + \frac{\ell_N}{N} \,\Big\} \;,$$

with the same convention as the one adopted for the sets $\mathcal{E}_N^i$ if one of the configurations ζ_i^N is an endpoint of the set E_N. Let

$$\mathcal{F}_N = \bigcup_{i=1}^{m} \mathcal{F}_N(\zeta_i^N) \;, \quad \Delta_N^2 \;:=\; E_N \setminus \mathcal{F}_N \;.$$

In view of the definition of the jump rates $R_N(\eta,\xi)$, inside the sets $\mathcal{F}_N(\zeta_i^N)$ the chain has a drift towards the configuration ζ_i^N which increases as the chain approaches the configuration ζ_i^N. $F_N(\zeta_i^N)$

$\alpha(a_i)$

Let $F_N(\zeta_i^N)$ be the interval $\{\, \zeta_i^N - [\delta_0 N]/N\,, \,\dots\,,\, \zeta_i^N + [\delta_0 N]/N\}$, $\mathcal{F}_N(\zeta_i^N) \subset F_N(\zeta_i^N)$, and denote by $\alpha(a_i)$ the exponent associated to a_i by (1.1). Hence, $H(x) = |x - a_i|^{\alpha(a_i)}$ in a small neighborhood of a_i, $1 \le i \le m$. Inside the set $F_N(\zeta_i^N)$ the chain has a drift towards the point ζ_i^N. For this reason we call the set $F_N(\zeta_i^N)$ a *well* or a *trap*. We show in (6.2) below that it takes a time of order $N^{\alpha(a_i)+1}$ for the chain to reach the boundary of the set $F_N(\zeta_i^N)$ when it starts from a configuration in $\mathcal{F}_N(\zeta_i^N)$. This property allows us to call the exponent $\alpha(a_i)$ the *depth* of the well $F_N(\zeta_i^N)$. The configuration ζ_i^N is called the *bottom* or the *center* of the well.

A_N^j

A_N

θ_j

Let $\theta_1 < \theta_2 < \cdots < \theta_\kappa$ be the possible depths of the traps $F_N(\zeta_i^N)$, i.e., the values of the exponents α_i introduced in (1.1): $\{\theta_1, \dots, \theta_\kappa\} = \{\alpha_1, \dots, \alpha_m\}$. In particular, $\theta_\kappa = \alpha$. Let $A^j = \{a_1^j, \dots, a_{m_j}^j\}$, $1 \le j \le \kappa$, be the set of points in A whose exponents are equal to θ_j,

$$A^j \;=\; \{a_1^j, \dots, a_{m_j}^j\} \;=\; \{a \in A : \alpha(a) = \theta_j\}\;.$$

Note that $m_\kappa = L$ and that $m_1 + \cdots + m_\kappa = m$. Let A_N^j be the discrete approximation of the set A^j,

$$A_N^j \;=\; \{[a_1^j N]/N, \dots, [a_{m_j}^j N]/N\}\;,$$

and let $A_N = A_N^1 \cup \cdots \cup A_N^\kappa$.

The previous considerations suggest that the chain $\eta^N(t)$ has a scaling limit in each time scale $N^2 \ll N^{\theta_1+1} \ll \cdots \ll N^{\theta_\kappa+1}$. On the smallest one, N^2, outside of the sets $F_N(\zeta_i^N)$, the chain behaves as a diffusion with a drift pointing towards the set A_N. The drift increases as the diffusion gets closer to the set A_N. When the diffusion reaches one of the traps $\mathcal{F}_N(\zeta_i^N)$ it is absorbed there since it takes the chain a time of order at least N^{θ_1+1} to exit any well. This informal description depicts the scaling limit of the chain in the diffusive time scale. A more profound analysis would show the convergence of the chain to a diffusion on the interval $[a,b]$ which is absorbed at A and is reflected at the boundary if the endpoint of the interval does not belong to A.

We turn to the scaling limit of the chain in the next time scale, N^{θ_1+1}. If the chain starts from a configuration which does not belong to any of the wells $\mathcal{F}_N(\zeta_i^N)$, in view of the conclusions of the previous paragraph, in the time scale N^{θ_1+1} the chain immediately reaches one of the wells. We have therefore to analyze the scaling limit when the chain starts from one of these wells.

We show in (**??**) that if the chain starts from a configuration ζ inside a well $\mathcal{F}_N(\zeta_i^N)$, with a probability asymptotically close to 1, it visits the bottom of the well $\mathcal{F}_N(\zeta_i^N)$ before it reaches any other well. We may therefore assume that the initial state belongs to A_N.

By (6.2), if the chain starts from a configuration ζ in $A_N \setminus A_N^1$, it remains in the well $F_N(\zeta)$ for ever in the time scale N^{θ_1+1}. The configurations in the set $A_N \setminus A_N^1$ act therefore as absorbing points for the asymptotic dynamics in the time scale N^{θ_1+1}.

Assume that the chain starts from a configuration $\zeta_k^N \in A_N^1$. We have seen that in this case the process hits the boundary of the set $F_N(\zeta_k^N)$ after a time of order N^{θ_1+1}. When the chain reaches this boundary, it starts evolving as a diffusion which is trapped after a time of order N^2, much smaller than the time N^{θ_1+1}, either in the same well $\mathcal{F}_N(\zeta_k^N)$ or in a neighboring well $\mathcal{F}_N(\zeta_{k-1}^N)$, $\mathcal{F}_N(\zeta_{k+1}^N)$. For the purpose of this analysis we denote these neighbor-hing wells by $\mathcal{F}'$.

If the chain is trapped in the the original well $\mathcal{F}_N(\zeta_k^N)$, it waits there a new time of order N^{θ_1+1} until it hits again the boundary of the set $F_N(\zeta_k^N)$. On the other hand, if it enters a neighbor well $\mathcal{F}'$, deeper than the well $\mathcal{F}_N(\zeta_k^N)$, in the time scale N^{θ_1+1} the chain is trapped in this well. If the well $\mathcal{F}'$ is as deep as $\mathcal{F}_N(\zeta_k^N)$, the same analysis carried out for the well $\mathcal{F}_N(\zeta_k^N)$ applies to this new trap.

These considerations show that on the time scale N^{θ_1+1} the chain visits the wells whose centers are in A_N^1 and it is trapped in the deeper wells whose centers are in $A_N \setminus A_N^1$. π_N^k

Consider the chain *reflected* in the set $F_N(\zeta_k^N)$. This means that forbid jumps from $F_N(\zeta_k^N)$ to its complement $F_N(\zeta_k^N)^c$ and jumps from $F_N(\zeta_k^N)^c$ to $F_N(\zeta_k^N)$, obtaining a birth and death chain on $F_N(\zeta_k^N)$. By (**??**), the stationary measure of this chain, denoted by π_N^k, is the stationary measure of the original chain conditioned to the set $F_N(\zeta_k^N)$:

$$\pi_N^k(\xi) \;=\; \frac{\pi_N(\xi)}{\pi_N(F_N(\zeta_k^N))}\;, \quad \xi \in F_N(\zeta_k^N)\;. \tag{1.7}$$

By (**??**) the process is reversible with respect to this measure.

It is shown in Lemma 5.6 below that the relaxation time of the reflected chain is of order N^2. Therefore, in view of (**??**), in a time of order N^2 the state of the reflected chain is very close to the stationary state π_N^k. As in (**??**), we refer to this proximity saying that the chain *equilibrated* or *thermalized* in $F_N(\zeta_k^N)$. Since this time is much smaller than the time needed to reach the boundary of the set $F_N(\zeta_k^N)$, before the chain exits the well it has reached a state very close to the stationary state, and for every purpouse we may assume that it has started from the stationary state.

With respect to the stationary state π_N^k, the set of boundary points of the well has a measure which vanishes as $N \uparrow \infty$. For this reason we may call this set a *rare event*. It has been established that in the realm of Markov chains the hitting time of rare events have asymptotically exponential distributions. Thus, the chain reaches the boundary of a well at an exponential time of order N^{θ_1+1}. Starting from the boundary of a well, the amount of time it takes for the chain to be absorbed by a new well is of order N^2, a negligible amount compared to the time N^{θ_1+1} needed to reach the boundary. We expect, therefore, that on the time scale N^{θ_1+1} the process visits the wells in A_N^1 at a succession of exponential times, i.e., that the chain converges to an A-valued Markov chain, whose points in $A \setminus A^1$ are absorbing.

A similar analysis at the longer time scale N^{θ_2+1} leads to the conclusion that the chain should converge to an $(A \setminus A^1)$-valued chain, whose points in $A \setminus [A^1 \cup A^2]$ are absorbing. In this time scale N^{θ_2+1}, the shallow traps $F_N(\zeta_k^N)$, $\zeta_k^N \in A_N^1$, are not seen since the time spent on these wells are of the order N^{θ_1+1}.

This investigation can be pursued in each time scale N^{θ_j+1} with similar conclusions. For each j, on the time scale N^{θ_j+1} the chain converges to an $[A^j \cup \cdots \cup A^\kappa]$-valued chain, whose points in $[A^{j+1} \cup \cdots \cup A^\kappa]$ are absorbing. In the last time scale, $N^{\alpha+1}$, only the deepest traps are seen and the chain converges to an A^κ-valued chain.

2 Metastable Markov Chains

We introduce in this section the concept of scaling limit of metastable Markov chains. The example of the previous section will help to understand the notion of metastable Markov chain.

Let E_N be a sequence of countable sets and consider a sequence $\{\eta^N(t) : t \ge 0\}$ of E_N-valued continuous-time, irreducible, positive-recurrent Markov chains. Denote by π_N the unique stationary probability measure of the chain $\eta^N(t)$.

$\mathbb{P}_\eta$

$\mathbb{E}_\eta$

Let $\mathbb{P}_\eta = \mathbb{P}_\eta^N$, $\eta \in E_N$, be the probability measure on $D(\mathbb{R}_+, E_N)$ induced by the Markov chain $\{\eta^N(t) : t \ge 0\}$ starting from η. Expectation with respect to $\mathbb{P}_\eta$ is denoted by $\mathbb{E}_\eta$.

S_L

Consider a finite number of disjoint subsets $\mathcal{E}_N^1, \dots, \mathcal{E}_N^L$, $L \ge 2$, of E_N: $\mathcal{E}_N^x \cap \mathcal{E}_N^y = \varnothing$, $x \ne y$. The sets $\mathcal{E}_N^x$ have to be interpreted as the wells of the Markov chains $\eta^N(t)$. Let $S_L = \{1, \dots, L\}$, let $\mathcal{E}_N = \cup_{x \in S_L} \mathcal{E}_N^x$ and let $\Delta_N = E_N \setminus \mathcal{E}_N$ so that

$$E_N \;=\; \mathcal{E}_N^1 \cup \cdots \cup \mathcal{E}_N^L \cup \Delta_N \; . \tag{2.1}$$

The set Δ_N is introduced to separate the wells $\mathcal{E}_N^x$. It is negligible in the sense that the amount of time the process $\eta^N(t)$ spends in $Delta_N$ is much smaller than the time needed to observe a jump from one well to another.

$\mathcal{E}_N$

Denote by $\eta^{\mathcal{E}}(t)$ the trace of the process $\eta^N(t)$ on the set $\mathcal{E}_N$, and by $(\mathfrak{R}_{\mathcal{E}} \eta^N)(t)$ the process which records the last site visited by $\eta^N(t)$ in the set $\mathcal{E}_N$, as defined in (1.5).

Denote by $\Psi = \Psi_N : \mathcal{E}_N \mapsto S_L \cup \{N\}$, the projection given by

$$\Psi(\eta) \;=\; \sum_{x=1}^{L} x\, \mathbf{1}\{\eta \in \mathcal{E}_N^x\} \;+\; N \mathbf{1}\{\eta \in \Delta_N\} \; .$$

Let $\{X_N(t) : t \ge 0\}$ (resp. $\{X_N^L(t) : t \ge 0\}$, $\{\hat{X}_N(t) : t \ge 0\}$) be the stochastic process on $S_L \cup \{N\}$ (resp. S_L) defined by $X_N(t) = \Psi(\eta^N(t))$ (resp.

$X_N^L(t) = \Psi(\eta^{\mathcal{E}}(t))$, $\hat{X}_N(t) = \Psi((\mathfrak{R}_{\mathcal{E}}\eta^N)(t)))$. Besides trivial cases, $X_N(t)$ is not Markovian.

Note that $X_N^L(t)$ is the trace of $X_N(t)$ on the set S_L and that $\hat{X}_N(t)$ is the process which records the last site in S_L visited by $X_N(t)$.

Definition 5.1. *Let ν_N be a sequence of probability measures on $\mathcal{E}_N$ such that $\nu_N \circ \Psi^{-1}$ converges to a probability measure ν on S_L. The sequence of Markov chains $\{\eta^N(t) : t \geq 0\}$ is said to have a scaling limit starting from the initial state ν_N if there exist*

(a) *An increasing sequence θ_N, $\theta_N \gg 1$,*
(b) *A partition* (2.1) *of the configuration space E_N, and*
(c) *A S_L-valued Markov chain $\mathbb{X}(t)$ whose distribution we denote by $\mathbf{P}_x$, $x \in S_L$,*

such that the measure $\mathbb{P}_{\nu_N} \circ \mathbb{X}_N^{-1}$, $\mathbb{X}_N(t) = X_N(t\theta_N) = \Psi(\eta^N(t\theta_N))$, converges in the soft topology to $\mathbf{P}_\nu = \sum_{x\in S_L} \nu(x)\mathbf{P}_x$.

Note that the probability measure ν_N is defined on $\mathcal{E}_N$ and not on E_N. The sequence θ_N is called the time scale, the sets $\mathcal{E}_N^x$ the metastable sets and the Markov chain $X(t)$ the asymptotic dynamics of the scaling limit. An example of measure ν_N which will appear in the next chapters is the Dirac measure concentrated on a configuration which belongs to the same well for all N: $\nu_N = \delta_{\eta^N}$, $\eta^N \in \mathcal{E}_N^x$ for all $N \geq 1$.

3 The Martingale Approach

The proof of the scaling limit of a sequence of chains is based on the martingale characterization of Markov chains and is divided in four steps.

Step 1: We prove the convergence in the Skorohod topology of the trace of $\mathbb{X}_N(t)$ on S_L to a Markov chain $\mathbb{X}(t)$.

Step 2: We show that the time spent by the chain $\eta^N(t)$ on the set Δ_N in the time scale θ_N is negligible.

Step 3: Let $(\mathfrak{R}_L\mathbb{X}_N)(t)$ be the process which records the last site visited by $\mathbb{X}_N(t)$ in the set $\{1, \dots, L\}$, as defined in (1.5). By Theorem 5.3 and by the results proved in Step 1 and 2, the process $(\mathfrak{R}_L\mathbb{X}_N)(t)$ also converges in the Skorohod topology to $\mathbb{X}(t)$.

Step 4: We show that the assumptions of Proposition 4.12 are fulfilled. The convergence in the soft topology of $\mathbb{X}_N(t)$ to $\mathbb{X}(t)$ follows then from Step 3 and from the assertion of Proposition 4.12.

We conclude this section with a sketch of the proof of Step 1. Fix a sequence of probability measures ν_N on $\mathcal{E}_N$. Denote by $\mathbb{X}_N^L(t)$ the trace of the process $\mathbb{X}_N(t)$ on the set S_L. Clearly, $\mathbb{X}_N^L(t)$ $\eta^{\mathcal{E}}(t)$

$$\mathbb{X}_N^L(t) \;=\; \Psi(\eta^{\mathcal{E}}(t\theta_N)) \;. \tag{3.1}$$

As usual, the proof of the convergence of $\mathbb{X}_N^L(t)$ is divided in two steps. We first show that the sequence of measures $\mathbb{P}_{\nu_N} \circ (\mathbb{X}_N^L)^{-1}$ is a tight sequence in $D(\mathbb{R}_+, S_L)$. Actually, in all examples of this book, the arguments used to prove tightness of the sequence $\mathbb{X}_N^L(t)$ are identical to the ones needed to characterize the limit points of the sequence $\mathbb{X}_N^L(t)$.

Once tightness has been established, it remains to prove the uniqueness of the limit points. This part of the proof requires the sequence of chains to fulfill two conditions: the chains need to be *locally ergodic* and the averaged jump rates of the trace process $\eta^{\mathcal{E}}$ need to converge.

Averaged jump rates. Denote by $R^{\mathcal{E}_N}$ the jump rates of the trace process $\eta^{\mathcal{E}}(t)$. The averaged jump rates $r_N(x,y)$, x, $y \in S_L$, of $\eta^{\mathcal{E}}(t)$ are defined by

$$r_N(x,y) \;=\; \frac{1}{\pi_N(\mathcal{E}_N^x)} \sum_{\eta \in \mathcal{E}_N^x} \pi_N(\eta) R^{\mathcal{E}_N}(\eta, \mathcal{E}_N^y)\;, \tag{3.2}$$

where $R^{\mathcal{E}_N}(\eta, \mathcal{E}_N^y) = \sum_{\xi \in \mathcal{E}_N^y} R^{\mathcal{E}_N}(\eta, \xi)$. We will assume that the sequences of averaged jump rates multiplied by θ_N converge:

$$\lim_{N\to\infty} \theta_N\, r_N(x,y) \;=\; r(x,y) \quad \text{for all } x,\, y \in S_L\;. \tag{3.3}$$

$R^{\mathcal{E}_N}$

Local ergodicity. Consider a sequence of functions $G = G_N : \mathcal{E}_N \to \mathbb{R}$, the functions $G_y(\eta) = R^{\mathcal{E}_N}(\eta, \mathcal{E}_N^y)$, for example. We have seen in the previous section that the birth and death chain equilibrates in each well $\mathcal{E}_N^x$ before reaching another well $\mathcal{E}_N^z$. Hence, by the ergodic theorem, we expect the time average of $G(\eta^{\mathcal{E}}(s))$ in the portion of time where the chain visits the well $\mathcal{E}_N^x$ to be very close to the corresponding time average of the mean value, with respect to the stationary measure π_N, of the function $G(\eta)$ in the set $\mathcal{E}_N^x$. This is one of the fundamental ideas of the approach presented in this book to derive the scaling limit of metastable Markov chains. This idea has some similitudes with the "one and two blocks" estimates of the theory of scaling limit of interacting particles systems which permits to replace a local function by a function of the empirical measure using the local ergodicity of the dynamics. We refer to Kipnis and Landim [1999] for an exposition of this theory and of the one and two blocks estimates.

To formulate a rigorous version of the ideas presented in the previous paragraph, for a function $G : \mathcal{E}_N \to \mathbb{R}$, let $\widehat{G} : \mathcal{E}_N \to \mathbb{R}$ be the averaged function given by

$$\widehat{G}(\eta) \;:=\; \frac{1}{\pi_N(\mathcal{E}_N^x)} \sum_{\xi \in \mathcal{E}_N^x} \pi_N(\xi)\, G(\xi) \quad \text{for } \eta \in \mathcal{E}_N^x\;.$$

If we denote by $\mathcal{P}$ the sigma-algebra of subsets of $\mathcal{E}_N$ generated by the partition $\{\mathcal{E}_N^x : 1 \le x \le L\}$, the function $\widehat{G}$ corresponds to the conditional expectation of G given $\mathcal{P}$:

$$\widehat{G} \;=\; E_{\pi_N}[G \,|\, \mathcal{P}]\;.$$

The function $\widehat{G}$ is constant on each well $\mathcal{E}_N^x$. It can therefore be expressed as a function of the projection Ψ, $\widehat{G}(\eta) = g(\Psi(\eta))$ for some function $g : S_L \to \mathbb{R}$.

The sequence θ_N introduced in Definition 5.1 represents the time scale in which the chain jumps among the wells. The assumption that the trace process $\eta^{\mathcal{E}}(t)$ mixes on each well $\mathcal{E}_N^x$ before reaching another well signifies therefore that the trace process $\eta^{\mathcal{E}}(t)$ mixes locally on each well $\mathcal{E}_N^x$ in the time scale θ_N. If this occurs, we expect that for every sequence of uniformly bounded function $G_N : \mathcal{E}_N \to \mathbb{R}$ and every $t > 0$,

$$\lim_{N\to\infty} \mathbb{E}_{\nu_N}\Big[\Big| \int_0^t \Big\{ G_N(\eta^{\mathcal{E}}(s\theta_N)) - \widehat{G_N}(\eta^{\mathcal{E}}(s\theta_N)) \Big\}\, ds \Big|\Big] \;=\; 0\;. \tag{3.4}$$

In the characterization of the limit points of the sequence $\mathbb{X}_N^L$ the full strength of condition (3.4) is not needed. It is enough that (3.4) holds for the jump rates of the trace process. By the definition (3.2) of the averaged jumps rates $r_N(x,y)$,

$$\widehat{R^{\mathcal{E}_N}}(\eta, \mathcal{E}_N^y) \;=\; r_N(\Psi(\eta), y)\;.$$

The proof of the uniqueness of limits of the sequence $\mathbb{X}_N^L$ requires that for each $t > 0$

$$\lim_{N\to\infty} \mathbb{E}_{\nu_N}\Big[\Big| \int_0^t \theta_N \Big\{ R^{\mathcal{E}_N}(\eta^{\mathcal{E}}(s\theta_N), \mathcal{E}_N^y) - r_N(\mathbb{X}_N^L(s\theta_N), y) \Big\}\, ds \Big|\Big] \;=\; 0\;. \tag{3.5}$$

Theorem 5.2. *Let E_N be a sequence of countable spaces and let $\{\eta^N(t) : t \ge 0\}$ be a sequence of E_N-valued, irreducible and positive-recurrent Markov chains. Consider the partition* (2.1) *of the state spaces E_N and let $\eta^{\mathcal{E}}(t)$ be trace of $\eta^N(t)$ on $\mathcal{E}_N$. Define the projection $\mathbb{X}_N^L(t)$ by* (3.1)*. Assume that conditions* (3.5) *and* (3.3) *are in force. Then, the sequence of processes $\mathbb{X}_N^L$ has at most one limit point which is a S_L-valued Markov chain with jump rates given by the right hand side of* (3.3)*.*

Proof. Let $\nu_N^L = \nu_N \circ \Psi^{-1}$ be the projection of the measure ν_N on S_L.

Assume that the process $\mathbb{X}_N^L(t)$ starting from ν_N^L converges in the Skorohod topology to a process $\mathbb{X}(t)$. In view of Theorem 2.2, to characterize the distribution of $\mathbb{X}(t)$ as the distribution of the S_L-valued Markov chain whose jump rates $r(x,y)$ are given by the right hand side of (3.3), it is enough to show that $\mathbb{X}(t)$ solves the martingale problem associated to the generator $\mathcal{L}$ whose jump rates are $r(x,y)$: For every function $F : S_L \to \mathbb{R}$,

$$M^F(t) \;=\; F(\mathbb{X}(t)) \;-\; F(\mathbb{X}(0)) \;-\; \int_0^t (\mathcal{L}F)(\mathbb{X}(s))\, ds$$

is a martingale. We may write this martingale as

$$F(\mathbb{X}(t)) \; - \; F(\mathbb{X}(0)) \; - \; \int_0^t \sum_{y=1}^{L} r(\mathbb{X}(s), y) \, [F(y) - F(\mathbb{X}(s))] \, ds \; . \tag{3.6}$$

Denote by $\mathcal{L}_{\mathcal{E}}$ the generator of the trace process $\eta^{\mathcal{E}}(t)$. By (7.2), for every bounded function $G : \mathcal{E}_N \to \mathbb{R}$,

$$M_N^G(t) \; = \; G(\eta^{\mathcal{E}}(t\theta_N)) \; - \; G(\eta^{\mathcal{E}}(0)) \; - \; \int_0^t \theta_N \, (\mathcal{L}_{\mathcal{E}} G)(\eta^{\mathcal{E}}(s\theta_N)) \, ds$$

is a martingale. In particular, taking $G = F \circ \Psi$ for a function $F : S_L \to \mathbb{R}$, we have that

$$M_N(t) = F(\mathbb{X}_N^L(t)) \; - \; F(\mathbb{X}_N^L(0)) \; - \; \int_0^t \theta_N \, [\mathcal{L}_{\mathcal{E}}(F \circ \Psi)](\eta^{\mathcal{E}}(s\theta_N)) \, ds \tag{3.7}$$

is a martingale.

Recall that we denote by $R^{\mathcal{E}_N}$ the jump rates of the trace process $\eta^{\mathcal{E}}(t)$. With this notation, we have that

$$[\mathcal{L}_{\mathcal{E}}(F \circ \Psi)](\eta) \; = \; \sum_{\xi \in \mathcal{E}_N} R^{\mathcal{E}_N}(\eta, \xi) \, \{(F \circ \Psi)(\xi) - (F \circ \Psi)(\eta)\} \; .$$

We may write this sum as

$$\sum_{y=1}^{L} [F(y) - F(\Psi(\eta))] \sum_{\xi \in \mathcal{E}_N^y} R^{\mathcal{E}_N}(\eta, \xi) \; = \; \sum_{y=1}^{L} [F(y) - F(\Psi(\eta))] \, R^{\mathcal{E}_N}(\eta, \mathcal{E}_N^y) \; ,$$

and the integral part of the martingale $M_N(t)$ becomes

$$\theta_N \int_0^t \sum_{y=1}^{L} [F(y) - F(\mathbb{X}_N^L(s))] \, R^{\mathcal{E}_N}(\eta^{\mathcal{E}}(s\theta_N), \mathcal{E}_N^y) \, ds \; . \tag{3.8}$$

If we compare this expression with the integral part of the martingale $M^F(t)$ appearing in (3.6), we see that to complete the argument we need to replace

$$\theta_N \, R^{\mathcal{E}_N}(\eta^{\mathcal{E}}(s\theta_N), \mathcal{E}_N^y) \quad \text{by} \quad r(\Psi(\eta^{\mathcal{E}}(s\theta_N)), \mathcal{E}_N^y) \; .$$

The first step in the proof of this replacement consists in closing the equation in terms of the process $\mathbb{X}_N^L(t)$ using the local ergodicity. While the first two terms of the martingale $M_N(t)$ introduced in (3.7) are expressed in terms of the projection $\mathbb{X}_N^L(t)$, the integral part is a function of the trace process $\eta^{\mathcal{E}}(t)$ and not of its projection $\mathbb{X}_N^L(t)$.

By assumption (3.5),

$$\begin{aligned} M_N(t) \; &= \; F(\mathbb{X}_N^L(t)) \; - \; F(\mathbb{X}_N^L(0)) \\ &- \; \int_0^t \sum_{y=1}^{L} [F(y) - F(\mathbb{X}_N^L(s))] \, \theta_N \, r_N(\mathbb{X}_N^L(s), y) \, ds \; + \; R_N \; , \end{aligned}$$

where R_N is a remainder which vanishes in $L^1(\mathbb{P}_{\nu_N})$, $\lim_N \mathbb{E}_{\nu_N}[\,|R_N|\,] = 0$, and whose value may change from line to line. By assumption (3.3), we may replace $\theta_N\, r_N(\mathbb{X}^L_N(s), y)$ in the previous formula by $r(\mathbb{X}^L_N(s), y)$ and obtain that

$$\begin{aligned} M_N(t) \;&=\; F(\mathbb{X}^L_N(t)) \;-\; F(\mathbb{X}^L_N(0)) \\ &-\; \int_0^t \sum_{y=1}^{L} [F(y) - F(\mathbb{X}^L_N(s))]\, r(\mathbb{X}^L_N(s), y)\, ds \;+\; R_N\;. \end{aligned} \tag{3.9}$$

We have now all the elements to show that $M^F(t)$ given by (3.6) is a martingale. To prove this claim we need to show that for every $s < t$ and every bounded function $H : D(\mathbb{R}_+, S_L) \to \mathbb{R}$, continuous with respect to the Skorohod topology and measurable with respect to the σ-algebra spanned by $\{\mathbb{X}(r) : 0 \le r \le s\}$,

$$\mathbb{E}_\nu\big[\,\{M^F(t) - M^F(s)\}\, H\,\big] \;=\; 0\;. \tag{3.10}$$

Since $M_N(t)$ is martingale,

$$\mathbb{E}_{\nu_N}\big[\,\{M_N(t) - M_N(s)\}\, H\,\big] \;=\; 0\;.$$

Therefore, by (3.9),

$$\begin{aligned} \lim_{N\to\infty} \mathbb{E}_{\nu_N}\Big[\Big\{ & F(\mathbb{X}^L_N(t)) \;-\; F(\mathbb{X}^L_N(s)) \\ & -\; \int_s^t \sum_{y=1}^{L} [F(y) - F(\mathbb{X}^L_N(v))]\, r(\mathbb{X}^L_N(v), y)\, dv \Big\}\, H\Big] \;=\; 0\;. \end{aligned}$$

Claim (3.10) follows from this fact and from the convergence of $\mathbb{X}^L_N$ to $\mathbb{X}$ in the Skorohod topology. □

4 The Last Visit Approximates the Trace

Consider the set-up introduced in Section 2. Let ν_N be a sequence of probability measures on $\mathcal{E}_N$. The main result of this section asserts that the last visit process $\hat{X}_N(t)$ is close in the Skorohod topology to the trace process $X_N(t)$ if for all $t > 0$,

$$\lim_{N\to\infty} \mathbb{E}_{\nu_N}\Big[\int_0^t \mathbf{1}\{\eta^N(s) \in \Delta_N\}\, ds\Big] \;=\; 0\;. \tag{4.1}$$

For a trajectory $\mathbf{x} \in D(\mathbb{R}_+, S_L)$, denote by $\{T_n(\mathbf{x}) : n \ge 0\}$, $\{S_n(\mathbf{x}) : n \ge 0\}$ the sequence of holding times and the sequence of jump times of $\mathbf{x}$, respectively. Set $S_0(\mathbf{x}) = 0$ and, for $n \ge 1$, we define $S_n(\mathbf{x})$ as

$$S_n(\mathbf{x}) \;:=\; \inf\{t > S_{n-1}(\mathbf{x}) : \mathbf{x}(t) \ne \mathbf{x}(S_{n-1}(\mathbf{x}))\}\;, \tag{4.2}$$

with the convention that $S_n(\mathbf{x}) = \infty$ if $S_{n-1}(\mathbf{x}) = \infty$ and, as usual, $\inf \varnothing = +\infty$. Let

$$T_n(\mathbf{x}) \;=\; S_n(\mathbf{x}) \;-\; S_{n-1}(\mathbf{x})$$

if $S_{n-1}(\mathbf{x}) < \infty$, and $T_n(\mathbf{x}) = 0$ if $S_{n-1}(\mathbf{x}) = \infty$.

Theorem 5.3. *Let $\{\eta^N(t) : t \geq 0\}$, $N \geq 1$, be a sequence of Markov chains fulfilling the conditions of Section 2. Fix $T > 0$ and denote by $\mathbb{Q}_N$ the probability measure on $D([0,T], S_L)$ induced by the trace process $X_N^L(t)$ starting from $\nu_N \circ \Psi^{-1}$, $\mathbb{Q}_N = \mathbb{P}_{\nu_N} \circ (X_N^L)^{-1}$. Assume that the sequence of measures $\mathbb{Q}_N$ converges in the Skorohod topology to a probability measure $\mathbb{Q}$, and that $\mathbb{Q}$ is the measure induced by a S_L-valued Markov chain. Assume, furthermore, that condition (4.1) is in force. Then,*

$$\lim_{N\to\infty} \mathbb{E}_{\nu_N}\big[\, d_T(X_N^L, \hat{X}_N)\,\big] \;=\; 0\, .$$

In particular, the measure $\hat{\mathbb{Q}}_N = \mathbb{P}_{\nu_N} \circ \hat{X}_N^{-1}$ on the space $D([0,T], S_L)$ converges in the Skorohod topology to $\mathbb{Q}$.

Proof. Fix $\epsilon > 0$ and $T > 0$. Denote by N_T the number of jumps in the time interval $[0,T]$ of a trajectory $\mathbf{x} \in D([0,T], S_L)$.
Claim A. There exists $K = K(\epsilon, T) \geq 1$ such that

$$\limsup_{N\to\infty} \mathbb{Q}_N[N_T \geq K] \;\leq\; \epsilon\, . \tag{4.3}$$

By assumption, the sequence $\mathbb{Q}_N$ converges to $\mathbb{Q}$. Since the set $\{N_T \geq K\}$ is closed for the Skorohod topology,

$$\limsup_{N\to\infty} \mathbb{Q}_N[N_T \geq K] \;\leq\; \mathbb{Q}[N_T \geq K]\, .$$

The measure $\mathbb{Q}$ corresponds to a Markov chain with jump rates $r(x,y)$. Under the measure $\mathbb{Q}$, conditionally on the jump chain $\{Y_n : n \geq 0\}$, $\{T_j : j \geq 0\}$ is a sequence of independent times of rates $\lambda(Y_j)$. Let $\lambda = \max_{x\in S_L} \lambda(x)$ and let $\{\tilde{T}_j : j \geq 0\}$ be an i.i.d. sequence of exponential random times of parameter λ. We have that

$$\mathbb{Q}[N_T \geq K] \;=\; \mathbb{Q}[T_1 + \cdots + T_K \;\leq\; T] \;\leq\; \mathbb{Q}[\tilde{T}_1 + \cdots + \tilde{T}_K \;\leq\; T]\, .$$

The last expression vanishes as $K \uparrow \infty$. This proves (4.3).
Claim B. There exists $\delta > 0$ such that

$$\limsup_{N\to\infty} \mathbb{Q}_N\Big[\bigcup_{j\geq 1} \big\{S_j \in [T-\delta, T]\big\}\Big] \;\leq\; \epsilon\, . \tag{4.4}$$

By (4.3), it is enough to show that

$$\limsup_{N\to\infty} \mathbb{Q}_N\Big[\bigcup_{j\ge 1}\{S_j\in[T-\delta,T]\}\,,\,\{N_T\le K\}\Big] \;\le\; \epsilon$$

for all $K\ge 1$. Fix $K\ge 1$. On the set $\{N_T\le K\}$ we may restrict the union to indices $j\le K$. To prove (4.4), it is therefore enough to show that

$$\limsup_{N\to\infty}\sum_{j=1}^{K}\mathbb{Q}_N\Big[S_j\in[T-\delta,T]\Big] \;\le\; \epsilon\;.$$

For each $j\ge 0$, the set $\{\mathbf{x}\in D([0,T],S_L): T-\delta\le S_j\le T\}$ is closed for the Skorohod topology. Therefore,

$$\limsup_{N\to\infty}\sum_{j=1}^{K}\mathbb{Q}_N\Big[S_j\in[T-\delta,T]\Big] \;\le\; \sum_{j=1}^{K}\mathbb{Q}\Big[S_j\in[T-\delta,T]\Big]\;.$$

Since the measure $\mathbb{Q}$ corresponds to the distribution of a Markov chain, for each $j\ge 1$, $\mathbb{Q}\Big[S_j\in[T-\delta,T]\Big]$ vanishes as $\delta\downarrow 0$. This proves (4.4).

Denote by S_n, $\hat S_n$, $n\ge 0$, the sequence of jump times of the processes X^L_N, $\hat X_N$, respectively, as defined in (4.2). Define the random variable

$$\mathfrak{n} \;:=\; \sup\{j\ge 0: \hat S_j<T\}\;.$$

Suppose that $S_{\mathfrak{n}+1}<T$. In this case,

$$T \;-\; S_{\mathfrak{n}+1}\le \int_0^T \mathbf{1}\{\eta^N(s)\in\varDelta_N\}\,ds\;. \tag{4.5}$$

Indeed, since,

$$\hat S_{\mathfrak{n}}-S_{\mathfrak{n}} \;=\; \int_0^{\hat S_{\mathfrak{n}}} \mathbf{1}\{\eta^N(s)\in\varDelta_N\}\,ds\;, \tag{4.6}$$

decomposing the integral appearing on the right hand side of (4.5), we see that (4.5) holds if and only if

$$(T-\hat S_{\mathfrak{n}}) \;-\; \int_{\hat S_{\mathfrak{n}}}^T \mathbf{1}\{\eta^N(s)\in\varDelta_N\}\,ds \;\le\; S_{\mathfrak{n}+1}-S_{\mathfrak{n}}\;.$$

The left hand side of this inequality is equal to

$$\int_{\hat S_{\mathfrak{n}}}^T \mathbf{1}\{\eta^N(s)\in\mathcal{E}_N\}\,ds \;\le\; \int_{\hat S_{\mathfrak{n}}}^{\hat S_{\mathfrak{n}+1}} \mathbf{1}\{\eta^N(s)\in\mathcal{E}_N\}\,ds \;=\; S_{\mathfrak{n}+1}-S_{\mathfrak{n}}\;.$$

This proves (4.5).

Claim C.

$$\limsup_{N\to\infty}\mathbb{P}_{\nu_N}\big[S_{\mathfrak{n}+1}\le T\big] \;=\; 0\;. \tag{4.7}$$

Denote by Ω_δ the event $\cup_{j\ge 1}\{S_j \in [T-\delta, T]\}$. Fix $\epsilon > 0$ and choose $\delta > 0$ for which (4.4) holds. On the event Ω_δ^c, $\{S_{\mathfrak{n}+1} \le T\} \subset \{S_{\mathfrak{n}+1} \le T-\delta\} = \{T - S_{\mathfrak{n}+1} \ge \delta\}$. Hence, by (4.4), (4.5), and assumption (4.1), the left hand side of (4.7) is bounded above by

$$\begin{aligned}\epsilon \;+\; &\limsup_{N\to\infty} \mathbb{P}_{\nu_N}\big[T - S_{\mathfrak{n}+1} \ge \delta\big] \\ &\le\; \epsilon \;+\; \limsup_{N\to\infty} \frac{1}{\delta}\,\mathbb{E}_{\nu_N}\Big[\int_0^T \mathbf{1}\{\eta^N(s)\in\Delta_N\}\,ds\Big] \;\le\; \epsilon\,.\end{aligned}$$

This proves (4.7).

Claim D. Let $N_T(X^L_N)$ be the number of jumps of the process X^L_N in the time interval $[0,T]$. For every $\epsilon > 0$, there exists $\delta > 0$ such that

$$\limsup_{N\to\infty} \mathbb{P}_{\nu_N}\Big[\bigcup_{j=1}^{N_T(X^L_N)} |S_j - S_{j-1}| \le \delta\Big] \;\le\; \epsilon\,.$$

By (4.3), it is enough to prove this claim with the additional constraint that $N_T(X^L_N) \le K$ for some $K = K(\epsilon, T)$ large enough. Since the set $\{|S_j - S_{j-1}| \le \delta\} \cap \{S_j \le T\}$ is closed for the Skorohod topology, and since $\mathbb{Q}_N$ converges to $\mathbb{Q}$,

$$\begin{aligned}\limsup_{N\to\infty} \mathbb{Q}_N &\Big[\{|S_j - S_{j-1}| \le \delta\} \cap \{S_j \le T\}\Big] \\ &\le\; \mathbb{Q}\Big[\{|S_j - S_{j-1}| \le \delta\} \cap \{S_j \le T\}\Big]\,.\end{aligned}$$

Since $\mathbb{Q}$ is the probability measure induced by a S_L-valued Markov chain, with bounded holding rates, the right hand side vanishes as $\delta \downarrow 0$. This concludes the proof of Claim D.

Claim E. For every $\delta > 0$,

$$\limsup_{N\to\infty} \mathbb{P}_{\nu_N}\big[|S_{\mathfrak{n}} - \hat{S}_{\mathfrak{n}}| \ge \delta\big] \;=\; 0\,.$$

This assertion follows from (4.6), the bound $\hat{S}_{\mathfrak{n}} \le T$ and assumption (4.1).

We are now in a position to prove the theorem. In view of Claims C, D and E, we may restrict our analysis to the set

$$\Omega \;=\; \{S_{\mathfrak{n}} + \delta < T < S_{\mathfrak{n}+1}\} \,\cap\, \{|S_{\mathfrak{n}} - \hat{S}_{\mathfrak{n}}| \le \delta/2\} \,\cap \bigcap_{j=1}^{N_T(X^L_N)} \{|S_j - S_{j-1}| \ge \delta\}$$

for some $\delta > 0$. On the set Ω, define the function $\lambda \in \Lambda_T$ by $\lambda(\hat{S}_j) = S_j$, $0 \le j \le \mathfrak{n}$, $\lambda(T) = T$, and complete λ on $[0,T]$ by linear interpolation.

Since $X^L_N(s) = \hat{X}_N(t)$ for $s \in [S_j, S_{j+1})$, $t \in [\hat{S}_j, \hat{S}_{j+1})$, $X^L_N(t) = \hat{X}_N(\lambda(t))$ for all $0 \le t \le T$. Therefore, by definition of the distance d_T,

$$d_T(X_N^L, \hat{X}_N) \;=\; \|\lambda\|^o \;.$$

By definition of λ, for $0 \le s < t \le T$,

$$\min_{0\le j<\mathfrak{n}} \frac{S'_{j+1} - S'_j}{\hat{S}'_{j+1} - \hat{S}'_j} \;\le\; \frac{\lambda(t) - \lambda(s)}{t - s} \;\le\; \max_{0\le j<\mathfrak{n}} \frac{S'_{j+1} - S'_j}{\hat{S}'_{j+1} - \hat{S}'_j} \;,$$

where $S'_j = S_j \wedge T$, $\hat{S}'_j = \hat{S}_j \wedge T$, $0 \le j \le \mathfrak{n}+1$. For $0 \le j < \mathfrak{n} - 1$,

$$\frac{\hat{S}_{j+1} - \hat{S}_j}{S_{j+1} - S_j} \;=\; 1 \;+\; \frac{1}{S_{j+1} - S_j} \int_{\hat{S}_j}^{\hat{S}_{j+1}} \mathbf{1}\{\eta^N(s) \in \Delta_N\}\, ds \;,$$

while by (4.6),

$$\frac{T - S_{\mathfrak{n}}}{T - \hat{S}_{\mathfrak{n}}} \;=\; 1 \;+\; \frac{\hat{S}_{\mathfrak{n}} - S_{\mathfrak{n}}}{T - \hat{S}_{\mathfrak{n}}} \;=\; 1 \;+\; \frac{1}{T - \hat{S}_{\mathfrak{n}}} \int_0^{\hat{S}_{\mathfrak{n}}} \mathbf{1}\{\eta^N(s) \in \Delta_N\}\, ds \;.$$

Therefore, since $\log(1+x) \le x$, and since on the set Ω, $|S_{j+1} - S_j| \ge \delta$, $T - \hat{S}_{\mathfrak{n}} \ge (T - S_{\mathfrak{n}}) - (\hat{S}_{\mathfrak{n}} - S_{\mathfrak{n}}) \ge \delta/2$,

$$\|\lambda\|^o \;\le\; \frac{2}{\delta} \int_0^T \mathbf{1}\{\eta^N(s) \in \Delta_N\}\, ds \;.$$

Hence, on the set Ω,

$$d_T(X_N^L, \hat{X}_N) \;\le\; \frac{2}{\delta} \int_0^T \mathbf{1}\{\eta^N(s) \in \Delta_N\}\, ds \;,$$

which concludes the proof of the theorem in view of assumption (4.1). □

5 Scaling Limit of Birth and Death Chains

In this section we present a rigorous formulation of the phenomena described in the previous section. A complete investigation of the asymptotic behavior of the chain requires an analysis of the convergence of the chain in the diffusive scaling to a diffusion with absorption points. We skip this analysis to concentrate on the behavior in the longer time scales.

We analyze the behavior of the chain in the time scales N^{θ_1+1} and $N^{\alpha+1}$, starting with the first. As explained in the previous section, the chain spends a negligible amount of time outside the union of the wells and it equilibrates inside each well before leaving the well. It is therefore natural to examine the asymptotic behavior of $X_N(t) = \Psi_N(\eta^N(t))$ instead of the one of $\eta^N(t)$, where $\Psi = \Psi_N : E_N \to \{0, 1, \dots, m\}$ is the function defined by

$$\Psi(\eta) \;=\; \sum_{k=1}^{m} x\, \mathbf{1}\{\eta \in \mathcal{E}_N^k\} \;.$$

Note that $\Psi(\eta) = 0$ if η belongs to $\Delta_N = E_N \setminus \mathcal{E}_N$, and that $X_N(t)$ is not a Markov chain.

A typical trajectory of the rescaled process $\mathbb{X}_N(t) = X_N(tN^{\theta_1+1})$ is presented in Figure **??**. The intervals of time $[S_j, T_j]$ correspond to the transitions from one well to the other, when the chain traverses the set Δ_N. The length of these intervals are of order $N^{1-\theta_1}$ and vanishes in the limit as $N \uparrow \infty$.

Such trajectories do not converge in the Skorohod topology. Therefore, to prove the convergence of the process $\mathbb{X}_N(t)$ to a finite state Markov chain, it is necessary to introduce a topology weaker than the Skorohod one. This is the content of Chapter xxx, where we define the soft topology.

Let $\mathbb{X}(t)$ be the Markov chain on $\{1, \dots, m\}$ characterized by the rates

$$
\begin{aligned}
r(i, i+1) \;&=\; \frac{1}{m(b_i)} \, \frac{1}{\int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\}\, du} \;, \\
r(i+1, i) \;&=\; \frac{1}{m(b_{i+1})} \, \frac{1}{\int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\}\, du} \;, \quad 1 \le i < L \;.
\end{aligned}
$$

Proposition 5.4. *for each $1 \le i \le m$,*

(**M1**) *For any sequence $\eta_N \in \mathcal{E}_N^i$, $N \ge 1$, the law of the stochastic process $\{\mathbb{X}(t) : t \ge 0\}$ under $\mathbf{P}_{x_N}$ converges to $\mathfrak{P}_i$ as $N \uparrow \infty$;*

(**M2**) *For every $T > 0$,*

$$
\lim_{N\to\infty} \sup_{x \in \mathcal{E}_N^i} \mathbf{E}_x \Big[\int_0^T \mathbf{1}\big\{ X^N(sN^{\alpha+1}) \in \Delta_N \big\}\, ds \Big] \;=\; 0 \;.
$$

Moreover,

$$
\lim_{N\to\infty} \inf_{x,y \in \mathcal{E}_N^i} \mathbf{P}_x \big[H_y < H_{\breve{\mathcal{E}}_N^i} \big] \;=\; 1 \;, \tag{5.1}
$$

where

$$
\breve{\mathcal{E}}_N^i \;:=\; \bigcup_{j \ne i} \mathcal{E}_N^j \;.
$$

Proof. a

We show that the hypotheses of Theorem **??** are in force. Let $\xi_N^i = b_i$ for $1 \le i \le L$. The asymptotic dynamics has no absorbing point and condition **(H2)** of [**?**, Theorem 2.7] follows from (1.6).

To check conditions **(H0)**, **(H1)** we take advantage from the one-dimensional setting to get explicit expressions for capacities. For two disjoint subsets A, B of E_N, denote by $\mathrm{cap}_N(A, B)$ the capacity between A and B. When $A = \{a\}$ we represent $\mathrm{cap}_N(A, B)$ by $\mathrm{cap}_N(a, B)$ with the same convention for B. Let $x < y$ be points in E_N. Recall that $\mathrm{cap}_N(x, y) = D_N(f_{x,y})$ where $f_{x,y} : E_N \mapsto \mathbb{R}$ solves the equation $L_N f_{x,y}(z) = 0$ for $z \notin \{x, y\}$ with boundary conditions $f_{x,y}(x) = 1$ and $f_{x,y}(y) = 0$. An elementary computation gives that $f(z) = 1$ for $z \le x$, $f(z) = 0$ for $z \ge y$ and

$$f(z+1/N) - f(z) \;=\; \frac{\big\{\pi_N(z) R_N(z,z+1/N)\big\}^{-1}}{\sum_{z=x}^{y-1/N} \big\{\pi_N(z) R_N(z,z+1/N)\big\}^{-1}}$$

for $z \in E_N \cap [x,y)$. Hence,

$$\mathrm{cap}_N(x,y) \;=\; \frac{1}{\sum_{z=x}^{y-1/N} \big\{\pi_N(z) R_N(z,z+1/N)\big\}^{-1}} \; . \tag{5.2}$$

In last two formulae, there is a slight abuse of notation since E_N is not the set $\{z/N : z \in \mathbb{Z} \cap [aN, bN]\}$, but the meaning is clear. In particular, if $\{x^N : N \ge 1\}$, $\{y^N : N \ge 1\}$ are two sequences in E_N such that $x^N \to a'$ and $y^N \to b'$ for some $a \le a' < b' \le b$, by (5.2),

$$\lim_{N\to\infty} N^{1+\alpha} \mathrm{cap}_N(x^N, y^N) \;=\; \frac{1}{\sum_{i=1}^L m(b_i)} \Big\{ \int_{a'}^{b'} \{H(u)/\Phi(u)\}\, du \Big\}^{-1} \;>\; 0\;. \tag{5.3}$$

Denote by $r_N(\mathcal{E}_N^i, \mathcal{E}_N^j)$ the average jump rate from $\mathcal{E}_N^i$ to $\mathcal{E}_N^j$ of the trace of the process Z_t^N on $\mathcal{E}_N = \cup_{1\le m\le L} \mathcal{E}_N^m$, defined by (**??**) in a general context. Clearly, $r_N(\mathcal{E}_N^i, \mathcal{E}_N^j) = 0$ for $|i-j| > 1$. Fix an arbitrary $1 \le i < L$ and let $G_1 = \cup_{j\le i} \mathcal{E}_N^j$, and $G_2 = \cup_{j>i} \mathcal{E}_N^j$ so that

$$\begin{aligned}\pi_N(\mathcal{E}_N^i)\, r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i+1}) &= \pi_N(b_i + \frac{\ell_N}{N})\, R_N^{\mathcal{E}}(b_i + \frac{\ell_N}{N}, b_{i+1} - \frac{\ell_N}{N}) \\ &= \pi_N(G_1)\, r_N(G_1, G_2)\;,\end{aligned}$$

where $R_N^{\mathcal{E}}(x,y)$, $x \ne y \in \mathcal{E}_N$, represents the jumps rates of the trace of Z_t^N on $\mathcal{E}_N$. Therefore, by (**??**),

$$\begin{aligned} r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i+1}) = \frac{\pi_N(G_1)\, r_N(G_1,G_2)}{\pi_N(\mathcal{E}_N^i)} \;&=\; \frac{\mathrm{cap}_N(G_1,G_2)}{\pi_N(\mathcal{E}_N^i)} \\ = \frac{\mathrm{cap}_N(b_i + \ell_N/N, b_{i+1} - \ell_N/N)}{\pi_N(\mathcal{E}_N^i)}\;. \end{aligned}$$

Analogously, we obtain that

$$r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i-1}) \;=\; \frac{\mathrm{cap}_N(b_i - \ell_N/N, b_{i-1} + \ell_N/N)}{\pi_N(\mathcal{E}_N^i)}$$

for any $1 < i \le L$. Therefore, by (5.3) and (1.6),

$$\lim_{N\to\infty} N^{1+\alpha}\, r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i+1}) \;=\; \frac{1}{m(b_i)} \Big\{ \int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\}\, du \Big\}^{-1}$$

and

$$\lim_{N\to\infty} N^{1+\alpha}\, r_N(\mathcal{E}_N^i, \mathcal{E}_N^{i-1}) \;=\; \frac{1}{m(b_{i+1})} \Big\{ \int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\}\, du \Big\}^{-1}$$

for any $1 \le i < L$, which concludes the proof of assumption **(H0)**.

The same arguments show that

$$\lim_{N\to\infty} N^{1+\alpha}\, \mathrm{cap}_N(\mathcal{E}_N^i, \breve{\mathcal{E}}_N^i)$$
$$= \frac{1}{\sum_{1\le i\le L} m(b_i)} \left\{ \frac{1}{\int_{b_i}^{b_{i+1}} \{H(u)/\Phi(u)\}\, du} + \frac{1}{\int_{b_{i-1}}^{b_i} \{H(u)/\Phi(u)\}\, du} \right\} , \tag{5.4}$$

provided $b_i \neq a$, b, with similar identities if $b_i = a$ or if $b_i = b$.

It remains to check condition **(H1)**. For any $1 \le i \le L$ and N large enough, $H(x) = |x - b_i|^\alpha$ for all $x \in \mathcal{E}_N^i$. In consequence, by (1.5) and 5.2, there exists a positive constant C_0, independent of N, such that

$$\mathrm{cap}_N(x, b_i) \;\ge\; \frac{C_0}{\ell_N^{\alpha+1}} ,$$

for any $1 \le i \le L$ and $x \in \mathcal{E}_N^i$. Therefore, since $\ell_N \ll N$, by (5.4),

$$\lim_{N\to\infty} \sup_{x\in\mathcal{E}_N^i} \frac{\mathrm{cap}_N(\mathcal{E}_N^i, \breve{\mathcal{E}}_N^i)}{\mathrm{cap}_N(x, b_i)} \;=\; 0$$

for all $1 \le i \le L$, which concludes the proof of the proposition. □

6 Ergodic Properties of the Birth and Death Chain

We present in this section some ergodic properties of the birth and death chain introduced in Section 1. We start with an estimation of the hitting time of the boundary of a well. Fix $1 \le i \le m$ and recall that we represent the interval $\{\, \zeta_i^N - [\delta_0 N]/N\, , \,\dots\, , \,\zeta_i^N + [\delta_0 N]/N\}$ by $F_N(\zeta_i^N)$. Let

$$m(a_i) \;=\; 1 \;+\; 2\sum_{k\ge 1} \frac{1}{k^{\alpha_i}} ,$$

$$I_-^i \;=\; \int_{a_i-\delta_0}^{a_i} \frac{H(x)}{\Phi(x)}\, dx\ , \quad I_+^i \;=\; \int_{a_i}^{a_i+\delta_0} \frac{H(x)}{\Phi(x)}\, dx\ . \tag{6.1}$$

We assume in the statement below that a_i is a point in the interior of the interval $[a, b]$. The same arguments apply to the case in which a_i is an endpoint of the interval $[a, b]$ and provide a formula for the expectation of the hitting time of the same order as the one below.

Lemma 5.5. *Assume that a_i is a point in the interior of the interval $[a, b]$, $a < a_i < b$. Denote by H_N the hitting time of $F_N(\zeta_i^N)^c$, and let η_N be a sequence of configurations such that $\lim_N |\eta_N - a_i| = 0$. Then,*

$$\lim_{N\to\infty} \frac{1}{N^{\alpha_i+1}}\, \mathbf{E}_{\eta_N}[H_N] \;=\; m(a_i)\, \frac{I_-^i \times I_+^i}{I_-^i + I_+^i} \cdot \tag{6.2}$$

η_i^-

η_i^+

Proof. Let

$$F(\zeta) \;=\; \mathbf{E}_\zeta[H_N]\,, \quad \zeta \in E_N\,,$$

and let η_i^-, η_i^+, be the outer boundary of the set $F_N(\zeta_i^N)$,

$$\eta_i^- = \zeta_i^N - ([\delta_0 N]+1)/N\,, \quad \eta_i^+ = \zeta_i^N + ([\delta_0 N]+1)/N$$

so that $F(\eta_i^-) = F(\eta_i^+) = 0$. Let $\epsilon = \epsilon_N = (1/N)$,

$$R(\eta) = R_N(\eta, \eta+\epsilon)\,, \quad Q(\eta) = R_N(\eta, \eta-\epsilon)\,.$$

By the detailed balance condition, $\pi(\eta)R(\eta) = \pi(\eta+\epsilon)Q(\eta+\epsilon)$. ϵ

Define B as the ratio $R(\eta)$

$Q(\eta)$

$$B \;:=\; \sum_{\eta=\eta_i^-+\epsilon}^{\eta_i^+-\epsilon} \frac{\pi[\eta_i^-+\epsilon,\eta]}{R(\eta)\,\pi(\eta)} \Big/ \sum_{\eta=\eta_i^-}^{\eta_i^+-\epsilon} \frac{1}{R(\eta)\pi(\eta)}\,,$$

where $\pi[\eta_i^-+\epsilon,\eta] = \pi\{\eta_i^-+\epsilon,\dots,\eta\}$. We claim that

$$F(\eta_i^-+\epsilon) - F(\eta_i^-) \;=\; \frac{B}{R(\eta_i^-)\pi(\eta_i^-)}\,\cdot \tag{6.3}$$

$\pi[\eta_i^-+\epsilon,\eta]$

To derive (6.3), apply the strong Markov property to obtain that

$$R(\eta)\,\{F(\eta+\epsilon)-F(\eta)\} \;=\; -1 \;+\; Q(\eta)\,\{F(\eta)-F(\eta-\epsilon)\}$$

for $\eta_i^- < \eta < \eta_i^+$. Iterating this relation we get that

$$\begin{aligned} F(\eta+\epsilon)-F(\eta) \;&=\; \frac{Q(\eta)\cdots Q(\eta_i^-+\epsilon)}{R(\eta)\cdots R(\eta_i^-+\epsilon)}\big\{F(\eta_i^-+\epsilon)-F(\eta_i^-)\big\} \\ &- \Big\{\frac{1}{R(\eta)} \;+\; \frac{Q(\eta)}{R(\eta)R(\eta-\epsilon)} \;+\;\cdots\;+\; \frac{Q(\eta)\cdots Q(\eta_i^-+2\epsilon)}{R(\eta)\cdots R(\eta_i^-+\epsilon)}\Big\} \end{aligned}$$

for $\eta_i^- < \eta < \eta_i^+$. This equation and the reversibility relation $\pi(\eta)R(\eta) = \pi(\eta+\epsilon)Q(\eta+\epsilon)$ yields that

$$F(\eta+\epsilon)-F(\eta) \;=\; -\frac{\pi[\eta_i^-+\epsilon,\eta]}{R(\eta)\pi(\eta)} \;+\; \frac{R(\eta_i^-)\pi(\eta_i^-)}{R(\eta)\pi(\eta)}\big\{F(\eta_i^-+\epsilon)-F(\eta_i^-)\big\} \tag{6.4}$$

for $\eta_i^- < \eta < \eta_i^+$. Since

$$0 \;=\; \sum_{\eta=\eta_i^-}^{\eta_i^+-\epsilon} \{F(\eta+\epsilon)-F(\eta)\}\,,$$

(6.3) follows from (6.4).

As $F(\eta_i^-)=0$, we obtain from (6.4) and from the explicit formula (6.3) for the difference $F(\eta_i^- +\epsilon) - F(\eta_i^-)$ that

$$F(\eta) \;=\; B \sum_{\xi=\eta_i^-}^{\eta-\epsilon} \frac{1}{R(\xi)\pi(\xi)} \;-\; \sum_{\eta=\eta_i^-+\epsilon}^{\eta-\epsilon} \frac{\pi[\eta_i^- +\epsilon, \xi]}{R(\xi)\,\pi(\xi)}$$

for $\eta_i^- < \eta < \eta_i^+$.

Fix a sequence of configurations η_N satisfying the assumption of the lemma. Let $A_N = \sum_{\eta_i^-\le\xi\le\eta_i^+-\epsilon}[R(\xi)\pi(\xi)]^{-1}$. By the explicit formula for B, and some simple algebra, we may rewrite $A_N\, F(\eta)$ as

$$\sum_{\xi=\eta}^{\eta_i^+-\epsilon} \frac{\pi[\eta_i^- +\epsilon, \xi]}{R(\xi)\pi(\xi)} \sum_{\xi=\eta_i^-}^{\eta-\epsilon} \frac{1}{R(\xi)\pi(\xi)} \;-\; \sum_{\xi=\eta_i^-+\epsilon}^{\eta} \frac{\pi[\eta_i^- +\epsilon, \xi]}{R(\xi)\pi(\xi)} \sum_{\xi=\eta}^{\eta_i^+-\epsilon} \frac{1}{R(\xi)\pi(\xi)}\;.$$

Denote the first term of this difference with $\eta=\eta_N$ by $A_N^{1,1}\times A_N^{1,2}$ and the second one by $A_N^{2,1}\times A_N^{2,2}$. By definition of π and R,

$$\lim_{N\to\infty} \frac{A_N}{NZ_N} \;=\; I_-^i \;+\; I_+^i\;, \quad \lim_{N\to\infty} \frac{A_N^{1,2}}{NZ_N} \;=\; I_-^i\;, \quad \lim_{N\to\infty} \frac{A_N^{2,2}}{NZ_N} \;=\; I_+^i\;,$$

where $I_\pm^i$ has been introduced in (6.1). On the other hand, the main contribution of $\pi[\eta_i^- +\epsilon,\xi]$ occurs for $\xi\ge\zeta_i^N - M_N/N$, where $1\ll M_N \ll N$. It follows from this observation that

$$\lim_{N\to\infty} \frac{A_N^{1,1}}{N^{\alpha_i+1}} \;=\; m(a_i)\, I_+^i\;, \quad \lim_{N\to\infty} \frac{A_N^{2,1}}{N^{\alpha_i+1}} \;=\; 0\;,$$

which concludes the proof of the lemma. □

We conclude this section with an estimation of the spectral gap of the birth and death chain reflected on a well. Fix $1\le i\le m$ and denote by $\xi^N(t)$ the chain reflected at $F(\zeta_i^N)$ and assume that a_i is an interior point of the interval $[a,b]$. As in the previous lemma, the arguments presented below can be adapted to the case in which a_i is an endpoint of the interval and provide a similar bound for the spectral gap.

Recall from (1.7) that we denoted by π^i the stationary measure of the reflected chain at $F(\zeta_i^N)$ and that π^i is the measure π conditioned to set $F(\zeta_i^N)$. Recall also that the spectral gap of the reflected chain, denoted by $\mathfrak{g}_N^i$, is given by

$$\mathfrak{g}_N^i \;=\; \inf_f \frac{\langle(-\mathcal{L}_i f), f\rangle_{\pi^i}}{\langle f, f\rangle_{\pi^i}}\;,$$

where the infimum is carried over all functions $f : F(\zeta_i^N)\to\mathbb{R}$ which have mean zero with respect to π^i, and where $\mathcal{L}_i$ represents the generator of the reflected process.

Lemma 5.6. *There exist constants* $0 < c_0 < C_0 < \infty$*, independent of* N*, such that*

$$\frac{c_0}{N^2} \;\le\; \mathfrak{g}_N^i \;\le\; \frac{C_0}{N^2}\;.$$

Proof. We start with the proof of the upper bound for the spectral gap. Consider the function $g : F(\zeta_i^N) \to \mathbb{R}$ such that

$$g(\eta) \;=\; \begin{cases} -1 \text{ for } \eta_i^- < \eta \le \eta_i^- + \delta_0/3, \\ \;\;0 \;\text{ for } |\eta - \zeta_i^N| \le \delta_0/3, \\ \;\;1 \;\text{ for } \eta_i^+ - \delta_0/3 \le \eta < \eta_i^+, \end{cases}$$

g is linear in the remaining two intervals. Let f be the mean-zero function defined by $f(\eta) = g(\eta) - E_{\pi^i}[g]$. It is not difficult to check that

$$E_{\pi^i}[g^2] \;\ge\; \frac{c_0}{N^{\alpha_i - 1}}\;, \quad \big|\, E_{\pi^i}[g] \,\big| \;\le\; \frac{C_0}{N^{\alpha_i - 1}}\;,$$

and that

$$\langle (-\mathcal{L}_i f), f \rangle_{\pi^i} \;=\; \frac{1}{2} \sum_{\eta = \eta_i^- + \epsilon}^{\eta_i^+ - 2\epsilon} \pi^i(\eta)\, R(\eta) \{ f(\eta + \epsilon) - f(\eta) \}^2 \;\le\; \frac{1}{N^2} \frac{C_0}{N^{\alpha_i - 1}}$$

for some constants $0 < c_0 < C_0 < \infty$ independent of N and whose value may change from line to line. This proves that $\mathfrak{g}_N^i \le C_0 N^{-2}$.

The proof of the lower bound for the spectral gap is simple. Denote by $\mathrm{Var}_{\pi^i}[f]$ the variance of a function f with respect to the stationary measure π^i. We need to show that there exists a finite constant C_0 such that for all functions $f : F(\zeta_i^N) \to \mathbb{R}$

$$\mathrm{Var}_{\pi^i}[f] \;\le\; C_0\, N^2\, \langle (-\mathcal{L}_i f), f \rangle_{\pi^i}\;. \tag{6.5}$$

By Schwarz inequality,

$$\begin{aligned} \mathrm{Var}_{\pi^i}[f] \;&\le\; \sum_{\eta = \eta_i^- + \epsilon}^{\eta_i^+ - \epsilon} \pi^i(\eta) \{ f(\eta) - f(\zeta_i^N) \}^2 \\ &\le\; N \sum_{\eta = \eta_i^- + \epsilon}^{\eta_i^+ - \epsilon} \pi^i(\eta) \sum_{\xi \in \gamma(\zeta_i^N, \eta)} \{ f(\xi + \epsilon) - f(\xi) \}^2\;, \end{aligned}$$

where $\gamma(\zeta_i^N, \eta)$ represents a path from ζ_i^N, to η, i.e., a sequence of nearest-neighbor configurations from ζ_i^N to η. We estimate the piece of the sum corresponding to the configurations $\eta \ge \zeta_i^N$. In this case, $\gamma(\zeta_i^N, \eta) = \{\zeta_i^N, \dots, \eta - \epsilon\}$. Interchanging the order of summation, the previous sum in the range $\zeta_i^N \le \eta \le \eta_i^+ - \epsilon$ becomes

$$N \sum_{\xi=\zeta_i^N}^{\eta_i^+ - 2\epsilon} \{f(\xi+\epsilon) - f(\xi)\}^2 \, \pi^i\{\xi+\epsilon, \dots, \eta_i^+ - \epsilon\} \; .$$

Since $R(\xi)$ is bounded below by a strictly positive constant, and since

$$\frac{1}{\pi^i(\xi)} \, \pi^i\{\xi+\epsilon, \dots, \eta_i^+ - \epsilon\} \;\le\; C_0 \, N \, (\xi - \zeta_i^N) \;\le\; C_0 \, N \; ,$$

the previous expression is bounded above by

$$C_0 N^2 \sum_{\xi=\zeta_i^N}^{\eta_i^+ - 2\epsilon} \pi^i(\xi) \, R(\xi) \, \{f(\xi+\epsilon) - f(\xi)\}^2 \;\le\; 2 \, C_0 \, N^2 \langle (-\mathcal{L}_i f), f \rangle_{\pi^i} \; ,$$

which concludes the proof of (6.5) and the one of the lemma. $\square$