Journal of Statistical Physics, Vol. 86, Nos. 3/4, 1997

The Consensus Times of the Majority Vote Process on a Torus

Dayue Chen[1]

Received June 2, 1995; final June 27, 1996

We study the majority vote process on a two-dimensional torus in which every voter adopts the minority of opinion with small probability δ. We identify the exponent that the mean of consensus time is asymptotically $(1/\delta)$ with that exponent as δ goes to 0. The proof is by a formula for mean exit time and by the metastable theory of Markov chains developed in the study of the stochastic Ising model.

KEY WORDS: Majority vote process; consensus time; attractors; mean exit time.

1. INTRODUCTION

The *majority vote process* is a spin system usually defined on an infinite lattice (ref. 10, p. 33). We now consider it on a two-dimensional torus $\mathbf{T} = \{1, 2, 3,..., N\} \times \{1, 2, 3,..., N\}$. A voter can have either of two opinions 0 or 1, and updates its opinion at exponential times with parameter 1. At an exponential time, the voter adopts the opinion of the majority of its four neighbors and itself with probability $1-\delta$ and the opinion of the minority with probability δ.

Points of $\mathbf{T}$ are denoted by x, y, and occasionally by two coordinates, e.g., (1,2). Let $\eta(x)$ be the opinion of the voter at x. Then $\eta(x)$ takes 0 or 1 only. The collection $\eta = \{\eta(x);\ x \in \mathbf{T}\}$ is called a configuration. $S = \{0, 1\}^{\mathbf{T}}$ is the set of all configurations. Let $x+e_1, x+e_2, x+e_3$ and $x+e_4$ be the adjacent sites of x. If $\sum_{i=1}^{4} |\eta(x) - \eta(x+e_i)| \leq 2$, the voter at

[1] Department of Probability and Statistics, Peking University, Beijing, 100871, China; e-mail: dayue@sxx0.math.pku.edu.cn.

822/86/3-4-21

x is in agreement with the majority of its neighbors and itself. η^x is the configuration that differs from η only at site x, i.e.,

$$\eta^x(y) = \begin{cases} \eta(y) & \text{if } \ y \neq x \\ 1-\eta(x) & \text{if } \ y = x \end{cases}$$

Infinitesimal rates are given as follows:

$$q(\eta, \xi) = \begin{cases} 1-\delta & \text{if } \ \xi = \eta^x, \ \ \sum_{i=1}^4 |\eta(x) - \eta(x+e_i)| \geqslant 3 \\ \delta & \text{if } \ \xi = \eta^x, \ \ \sum_{i=1}^4 |\eta(x) - \eta(x+e_i)| \leqslant 2 \\ 0 & \text{if } \ \xi \neq \eta^x, \ \ \forall x \in \mathbf{T} \end{cases} \tag{1.1}$$

The corresponding Markov chain $\{\xi_t\}$ on S is called the *majority vote process* on a torus. Here $0 \leqslant \delta \leqslant 1$. To emphasize the dependence of δ we sometimes put δ as the superscript, e.g., $q^\delta(\eta, \xi)$.

The majority vote process $\{\xi_t\}$ is recurrent, since S is finite. Let $\mathbf{0}$ ($\mathbf{1}$, respectively) be the configuration that all voters hold opinion 0 (1, respectively) and

$$\begin{aligned} \sigma(\mathbf{0}) &= \inf\{t \geqslant 0; \xi_t = \mathbf{0}\}, \qquad & \sigma(\mathbf{1}) &= \inf\{t \geqslant 0; \xi_t = \mathbf{1}\} \\ T_1 &= \sigma(\mathbf{0}) \wedge \sigma(\mathbf{1}), & T_2 &= \sigma(\mathbf{0}) \vee \sigma(\mathbf{1}) \end{aligned} \tag{1.2}$$

Here $a \wedge b$ means the minimum of a and b, $a \vee b$ the maximum of a and b. The time T_1 is the time when all voters reach an agreement in opinion for the first time, and is therefore called a consensus time. T_2 is the time that all voters adopt the same opinion 1 after they all adopted opinion 0 simultaneously, or vice versa. If δ is small, it would take a very long time for a voters to take the opinion of the minority. We have the following asymptotic estimates on ET_1 and ET_2.

Theorem 1. The state space S is partitioned into three disjoint sets S_0, S_1 and S_2. If the initial state $\xi \in S_k$, $k = 0$, 1, or 2, then $\lim_{\delta \to 0} \delta^k E_\xi T_1$ exists and is a rational number.

The sets S_0, S_1, and S_2 will be determined in Section 3 after more preparations. It follows from Theorem 1 that

$$\lim_{\delta \to 0} P_\xi \left(\left| \frac{\log T_1}{-\log \delta} \right| \leqslant \varepsilon \right) = 1 \qquad \forall \xi \in S_0, \quad \forall \varepsilon > 0$$

However, as we shall see later, for initial state $\xi \in S_1$, $(\log T_1)/(-\log \delta)$ converges in distribution to $\alpha_\xi \delta_0 + (1-\alpha_\xi)\, \delta_1$, where δ_0 and δ_1 are the probability measures concentrating on 0 and 1 respectively, and $\alpha_\xi \in [0, 1)$

is a number depending on ξ. By the end of this paper we will determine the ξ for which $\alpha_\xi = 0$, i.e.,

$$\lim_{\delta \to 0} P_\xi \left(\left| \frac{\log T_1}{-\log \delta} - 1 \right| \leqslant \varepsilon \right) = 1$$

By the same argument we will also give the range of $\xi \in S_2$ such that

$$\lim_{\delta \to 0} P_\xi \left(\left| \frac{\log T_1}{-\log \delta} - 2 \right| \leqslant \varepsilon \right) = 1 \tag{1.3}$$

Theorem 2. For any initial state, $\delta^N E T_2$ converges to a rational number as $\delta \to 0$; furthermore, T_2/ET_2 converges in law to a random variable having the exponential distribution with mean 1.

Corollary 3. For any initial state and for any $\varepsilon > 0$,

$$\lim_{\delta \to 0} P \left(\left| \frac{\log T_2}{-\log \delta} - N \right| \leqslant \varepsilon \right) = 1$$

Remark. It is interesting to note that in dimension one

$$\lim_{\delta \to 0} \frac{\log ET_1}{-\log \delta} \leqslant 1; \qquad \lim_{\delta \to 0} \frac{\log ET_2}{-\log \delta} = 2 \tag{1.4}$$

Namely, the volume N does not come into play.

Similar results have been established for the stochastic Ising model.[4, 5, 11, 12, 14-16] The inverse temperature β of the stochastic Ising model plays the same role as δ here. The motivation was to understand the so-called metastable behavior observed in physical experiments. There are hundreds of papers on metastable behavior in both the mathematical and physical literature (for example, refs. 2, 9, and 13).

The results reported in this paper are inspired by works on the stochastic Ising model. We have chosen a very simple model. The method can be extended to dealing with higher dimensions, larger neighborhoods, or the case that infinitesimal rates are biased in 0 and 1. Similar assertions hold with different constants, but require substantial analysis. In ref. 6, consensus times of a voter model on the torus in Z^d are studied as $N \to \infty$.

It is worth comparing the majority vote process with the stochastic Ising model. The stochastic Ising model is reversible with respect to the Gibbs measure, while the majority vote process is not. The reversibility is crucial to the analysis of attractors of the stochastic Ising model.[4, 5] It is the Hamiltonian that gives all the information we need to compute ET_1

and ET_2. On the other hand, the majority vote process enjoys a simple feature. It is shown in refs. 3 and 14 that there is a hierarchic structure among the attractors. There are only four levels of attractors, and a characteristic quantity called $C(\xi, \eta)$ is either 0,1, or ∞. This makes the analysis manageable. In the study of the stochastic Ising model a configuration called the *critical droplet* plays a crucial role. With probability very close to 1 the evolution from all spins up to all spins down goes through a critical droplet. The Hamiltonian of the critical droplet is the max–min if all configurations are classified according to the number of $+1$ spins. The critical droplet is unique in some sense. So an important object in the study of the stochastic Ising model is to identify the critical droplet. There is no counterpart in the majority vote process. However, there are many level 2 attractors and they somewhat share the role of the critical droplet.

This paper is self-contained, although Theorems 1 and 2 were first derived by applying the results of ref. 3. We redefine *attractors* in Section 3 and use implicitly the idea of ref. 3. Our approach does not increase redundancy, because there are only four levels in the hierarchic structure of attractors. It is for completeness that we include the proof of the second part of Theorem 2. The original proof can be found in ref. 3. In Section 2 we first recall the *matrix tree theorem*. Applying this theorem, we convert the proof of Theorems 1 and 2 to finding a *good* map from a subset of S into S. The desired estimate on the exponents is obtained through a careful (and somewhat tedious) study in eight lemmas. Our approach is quite simple if the detailed proof of the lemmas is skipped.

2. PRELIMINARIES

Suppose that $S = \{\xi, \eta, ..., \zeta\}$ is finite and K is a proper subset of S.

Definition 4. $G(K)$ is the set of maps $g: K \to S$ with the property that g maps no nonempty subset of K into itself. We say that $g \in G(K)$ leads $\xi \in K$ to $\eta \in S \setminus K$ if there is a sequence $\{\zeta_1, ..., \zeta_n\}$ of distinct elements in K such that

$$g(\xi) = \zeta_1, \qquad g(\zeta_n) = \eta, \qquad \text{and} \qquad g(\zeta_j) = \zeta_{j+1}, \qquad 1 \leqslant j \leqslant n-1 \quad (2.1)$$

Let $G_{\xi\eta}(K) = \{g \in G(K); g \text{ leads } \xi \text{ to } \eta\}$.

Let $\{X_t\}$ be a continuous-time Markov chain on S, $q(\xi, \eta)$ the infinitesimal rate from ξ to η. Let $\pi(g) = \prod_{x \in K} q(\xi, g(\xi))$ for $g \in G(K)$. The *first exit time* of Markov chain $\{X_t\}$ is

$$\tau(K) = \inf\{t \geqslant 0; X_t \notin K\}$$

Lemma 5. Suppose that K is a proper subset of S, $\xi \in K$, and $\eta \notin K$. Suppose that $\sum_{\xi \in S, \xi \neq \zeta} q(\zeta, \xi) = -q(\zeta, \zeta)$ and $0 < -q(\zeta, \zeta) < \infty$ for any $\zeta \in K$. Then

$$P(X_{\tau(K)} = \eta \mid X_0 = \xi) = \frac{\sum_{g \in G_{\xi\eta}(K)} \pi(g)}{\sum_{g \in G(K)} \pi(g)}$$

$$E_\xi \tau(K) = \frac{\sum_{g \in G(K\setminus\{\xi\})} \pi(g) + \sum_{\zeta \in K, \zeta \neq \xi} \sum_{g \in G_{\xi\zeta}(K\setminus\{\zeta\})} \pi(g)}{\sum_{g \in G(K)} \pi(g)} \tag{2.2}$$

This lemma is called the *matrix tree theorem* in ref. 7. It was known 40 years ago.[1] For the proof, see ref. 8, Lemmas 3.3 and 3.4 of Chapter 6. For discrete-time Markov chains it is also derived by an elementary method of determinant in ref. 3.

Applying this lemma to the majority vote process, we find that every $\pi(g)$ is 0 or of form $\delta^a(1-\delta)^b$ with very large integers a and b. Then $E_\xi \tau(K)$ is the ratio of two huge polynomials of δ by (2.2). To estimate $E\tau(K)$ asymptotically as $\delta \to 0$, we need to know the *difference of the minimum exponents* of δ in the two polynomials. Consider the family of infinitesimal rates $\{q^\delta(\xi, \eta); \delta \in [0, 1]\}$ indexed by δ. Introduce

$$C(\xi, \eta) = \begin{cases} \lim_{\delta \to 0} [-\log q^\delta(\xi, \eta)]/(-\log \delta) & \text{if } q^0(\xi, \eta) = 0 \\ 0 & \text{if } q^0(\xi, \eta) > 0 \end{cases} \tag{2.3}$$

with the convention that $\log 0 = -\infty$. In addition, let $C(\xi, \xi) = 0$ for any $\xi \in S$. For subset $K \subset S$, define

$$W(K) = \min_{g \in G(K)} \sum_{\zeta \in K} C(\zeta, g(\zeta)) \tag{2.4}$$

$$W_{\xi\eta}(K) = \min_{g \in G_{\xi\eta}(K)} \sum_{\zeta \in K} C(\zeta, g(\zeta)) \tag{2.5}$$

They are the minimum exponents of the related polynomials of δ. It follows from Lemma 5 that, assuming that $\xi \in K$ and $\eta \notin K$,

$$\lim_{\delta \to 0} \frac{-\log P_\xi^\delta(X_{\tau(K)} = \eta)}{-\log \delta} = -W(K) + W_{\xi\eta}(K) \tag{2.6}$$

$$\lim_{\delta \to 0} \frac{\log E_\xi^\delta \tau(K)}{-\log \delta} = W(K) - W(K\setminus\{\xi\}) \wedge \min_{\zeta \in K} W_{\xi\zeta}(K\setminus\{\zeta\}) \tag{2.7}$$

3. ANALYSIS ON ATTRACTORS

According to (2.3), for the majority vote process,

$$C(\xi, \eta) = \begin{cases} \infty & \text{if } \sum_{x \in \mathbf{T}} |\xi(x) - \eta(x)| \geqslant 2 \\ 0 & \text{if } \eta = \xi^x, \sum_{i=1}^{4} |\xi(\mathbf{x} + e_i) - \xi(x)| \geqslant 3 \\ 1 & \text{if } \eta = \xi^x, \sum_{i=1}^{4} |\xi(x + e_i) - \xi(x)| \leqslant 2 \end{cases} \tag{3.1}$$

Definition 6. Configuration ζ is called an attractor if

$$\sum_{i=1}^{4} |\zeta(x + e_i) - \zeta(x)| \leqslant 2 \qquad \text{for all} \quad x \in \mathbf{T}$$

We say attractor ζ is of *level 1* if $\max_{x \in \mathbf{T}} \sum_{i=1}^{4} |\zeta(x + e_i) - \zeta(x)| = 2$.
Let $\mathbf{A}$ be the set of all level 1 attractors.
We say attractor ζ is of *level* 2 if $\max_{x \in \mathbf{T}} \sum_{i=1}^{4} |\zeta(x + e_i) - \zeta(x)| = 1$.

Remark. The definition of *level* is altered from that of refs. 3 and 14. Notice that $\mathbf{0}$ and $\mathbf{1}$ are the only attractors that satisfy

$$\max_x \sum_{i=1}^{4} |\zeta(x + e_i) - \zeta(x)| = 0$$

It is often helpful to visualize ξ by identifying ξ with the subset $\{x \in \mathbf{T};\ \xi(x) = 1\}$. For example, if ξ is a level 2 attractor, $\{x \in \mathbf{T};\ \xi(x) = 1\}$ consists

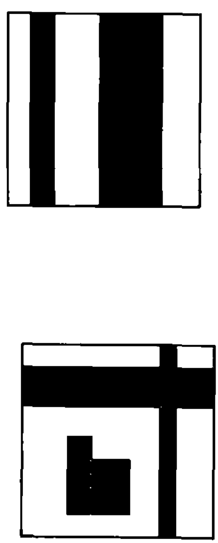

Fig. 1. Some attractors of level 1 (upper) and level 2 (lower). $\{x \in \mathbf{T};\ \eta(x) = 1\}$ is the shaded area.

of several columns or rows. Some typical subsets of **T** corresponding to attractors are shown in Fig. 1.

Definition 7. For attractor ζ, define $B^1(\zeta)=\{\eta\in S \mid \text{there is a sequence } \eta_0, \eta_1,..., \eta_n \text{ such that } \eta_0=\eta,\ \eta_n=\zeta, \text{ and } C(\eta_i, \eta_{i+1})=0 \text{ for } i=0, 1,..., n-1\}$.

Lemma 8.

(i) If ζ and η are attractors and if $\zeta\neq\eta$, then $\eta\notin B^1(\zeta)$.

(ii) If ζ is an attractor, then there is a map $g\in G(B^1(\zeta)\backslash\{\zeta\})$ such that g leads every $\xi\in B^1(\zeta)\backslash\{\zeta\}$ to ζ in the sense of (2.1) and

$$\sum_{\xi\in B^1(\zeta)\backslash\{\zeta\}} C(\xi, g(\xi))=0=W(B^1(\zeta)\backslash\{\zeta\})$$

(iii) $W(B^1(\zeta))=1$ if ζ is a level 1 attractor.

Proof. (i) If $\eta\in B^1(\zeta)$, then there exists a sequence leading η to ζ. In particular, there exists $\eta_1=\eta^x$ for some $x\in\mathbf{T}$ such that $C(\eta, \eta^x)=0$. On the other hand, η is an attractor if and only if $C(\eta, \eta^x)=1$ for all $x\in\mathbf{T}$. This contradiction shows that $\eta\notin B^1(\zeta)$.

(ii) We first define $g(\xi)=\zeta$ if $C(\xi, \zeta)=0$. Then $D=\{\xi; C(\xi, \zeta)=0, \xi\neq\zeta\}$ is a subset of $B^1(\zeta)\backslash\{\zeta\}$. Next we extend D step by step. If $\xi'\notin D$ and if there exists $\xi\in D$ such that $C(\xi', \xi)=0$, define $g(\xi')=\xi$. By definition of $B^1(\zeta)$ there is always such a pair (ξ and ξ') unless $D=B^1(\zeta)\backslash\{\zeta\}$ already. On the other hand, since $B^1(\zeta)$ is finite, the extension of D to $B^1(\zeta)\backslash\{\zeta\}$ can be completed in finite steps. Map g leads every $\xi\in B^1(\zeta)\backslash\{\zeta\}$ to ζ and $\sum_{\xi\in B^1(\zeta)\backslash\{\zeta\}} C(\xi, g(\xi))=0$. Now, $W(B^1(\zeta)\backslash\{\zeta\})\geqslant 0$ by (2.4) and is minimized by g.

(iii) If ζ is a level 1 attractor, by definition there is $x\in\mathbf{T}$ such that $\sum_{i=1}^4 |\zeta(x+e_i)-\zeta(x)|=2$. Notice that $\zeta^x\notin B^1(\zeta)$. Define $g'(\zeta)=\zeta^x$, and $g'(\xi)=g(\xi)$ defined in part (ii) if $\xi\in B^1(\zeta)\backslash\{\zeta\}$. Then $g'\in G(B^1(\zeta))$ and $\sum_{\xi\in B^1(\zeta)} C(\xi, g(\xi))=1$. On the other hand, for any $f\in G(B^1(\zeta))$, $\sum_{\xi\in B^1(\zeta)} C(\xi, f(\xi))\geqslant C(\zeta, f(\zeta))=1$. Hence $W(B^1(\zeta))\geqslant 1$ and is minimized by g'. ∎

Remarks. (1) This is a rather simple fact, but it is a basic ingredient of the subsequent arguments.

(2) The proof of (ii) and (iii) consists of two parts: to find a lower bound and to find a *good* map that reaches the lower bound. It is delicate to make sure that g maps no subset of K into itself. This is guaranteed, e.g., in part (ii) by the fact that g leads every ξ to ζ (see Fig. 2).

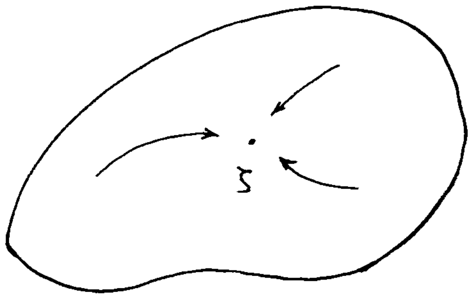

Fig. 2. Maps g constructed in the proof of Lemma 8, part (ii).

(3) It is clear from the proof that the statement is still valid if $B^1(\zeta)$ or $B^1(\zeta)\backslash\{\zeta\}$ is replaced by a *reasonable* subset, e.g., $B^1(\zeta)\backslash B^1(\eta)$.

(4) In most cases there are several x's such that $\sum_{i=1}^4 |\zeta(x+e_i)-\zeta(x)|=2$. This enables us to choose x with additional properties, e.g., $\zeta(x)=1$.

(5) In general, $B^1(\eta)\cap B^1(\zeta)\neq\emptyset$. For example, if N is even, configurations η' and η'' are in every $B^1(\cdot)$, where

$$\{x\in\mathbf{T};\eta'(x)=1\}=\{(m,n)\in\mathbf{T}; m+n=\text{even}\}$$
$$\{x\in\mathbf{T};\eta''(x)=1\}=\{(m,n)\in\mathbf{T}; m+n=\text{odd}\}$$

Definition 9. (i) For attractors η and ζ, we say $\eta \overset{(1)}{\Longrightarrow} \zeta$ if there is a sequence $\eta_0, \eta_1,..., \eta_n$ such that $\eta_0=\eta$, $\eta_n=\zeta$, and $\sum_{i=0}^{n-1} C(\eta_i, \eta_{i+1})=1$.

(ii) For attractors η and ζ, we say $\eta \overset{(1)}{\longrightarrow} \zeta$ if there is a sequence of attractors $\zeta_0, \zeta_1,..., \zeta_n$ such that $\zeta_0=\eta$, $\zeta_n=\zeta$, and $\zeta_i \overset{(1)}{\Longrightarrow} \zeta_{i+1}$ for $i=0, 1,..., n-1$.

(iii) For $\zeta=\mathbf{0}, \mathbf{1}$ or a level 2 attractor, define

$$B^2(\zeta)=\bigcup\{B^1(\eta)\mid\eta \text{ is an attractor and } \eta \overset{(1)}{\longrightarrow} \zeta\}$$

Lemma 10. (i) $B^2(\zeta)$ contains only one level 2 attractor, namely ζ itself.

(ii) If ζ is $\mathbf{0}, \mathbf{1}$, or a level 2 attractor, then there is a map $g\in G(B^2(\zeta)\backslash\{\zeta\})$ such that g leads every $\xi\in B^2(\zeta)\backslash\{\zeta\}$ to ζ and

$$\sum_{\xi\in B^2(\zeta)\backslash\{\zeta\}} C(\xi, g(\xi))=\sum_{\xi\in B^2(\zeta)\backslash\{\zeta\}} 1_{\{\xi\in\mathbf{A}\}}=W(B^2(\zeta)\backslash\{\zeta\})$$

where $\mathbf{A}$ is the set of all level 1 attractors. $\sum_{\xi\in B^2(\zeta)\backslash\{\zeta\}} 1_{\{\xi\in\mathbf{A}\}}$ is the number of level 1 attractors in $B^2(\zeta)\backslash\{\zeta\}$.

(iii) $W(B^2(\zeta))=2+\sum_{\xi\in B^2(\zeta)} 1_{\{\xi\in\mathbf{A}\}}$ if ζ is a level 2 attractor.

Proof. (i) Suppose that ζ is a level 2 attractor and that $\{\eta_0, \eta_1,..., \eta_n\}$ is a sequence of distinct elements such that $\zeta = \eta_0$ and η_n is another attractor. Then $C(\eta_0, \eta_1) = \infty$ or $\eta_1 = \zeta^x$. In the latter case,

$$\sum_{i=1}^{4} |\zeta(x+e_i) - \zeta(x)| \leqslant 1 \Rightarrow \sum_{i=1}^{4} |\eta_1(x+e_i) - \eta_1(x)| \geqslant 3$$

So η_1 is not an attractor and $n \geqslant 2$. Either $C(\eta_1, \eta_2) = \infty$ or we can assume that $\eta_2 = \eta_1^y$, $x \neq y$. We have

$$\sum_{i=1}^{4} |\zeta(y+e_i) - \zeta(y)| \leqslant 1 \Rightarrow \sum_{i=1}^{4} |\eta_1(y+e_i) - \eta_1(y)| \leqslant 2$$

So $C(\eta_1, \eta_2) = 1$. Consequently $\sum_{i=0}^{n-1} C(\eta_i, \eta_{i+1}) \geqslant C(\eta_0, \eta_1) + C(\eta_1, \eta_2) \geqslant 2$. Hence there is no attractor η such that $\zeta \overset{(1)}{\Longrightarrow} \eta$. In other words, ζ is not contained in any $B^2(\eta)$ if $\eta \neq \zeta$.

(ii) The proof is very similar to that of Lemma 8. For any $f \in G(B^2(\zeta)\backslash\{\zeta\})$,

$$\sum_{\xi \in B^2(\zeta)\backslash\{\zeta\}} C(\xi, f(\xi)) \geqslant \sum_{\xi \in B^2(\zeta)\backslash\{\zeta\}} C(\xi, f(\xi)) 1_{\{\xi \in A\}} = \sum_{\xi \in B^2(\zeta)\backslash\{\zeta\}} 1_{\{\xi \in A\}}$$

The last equality holds because $C(\xi, \xi^x) = 1$ for any x if ξ is an attractor. Since f is arbitrary, $W(B^2(\zeta)\backslash\{\zeta\}) \geqslant$ number of level 1 attractors in $B^2(\zeta)\backslash\{\zeta\}$.

We now construct a map $g \in G(B^2(\zeta)\backslash\{\zeta\})$ with the desired properties. First take $D = B^1(\zeta)\backslash\{\zeta\}$ and $g|_{B^1(\zeta)\backslash\{\zeta\}}$ by part (ii) of Lemma 8. Then g leads every $\xi \in D$ to ζ and

$$\sum_{\xi \in D} C(\xi, g(\xi)) = \sum_{\xi \in D} 1_{\{\xi \in A\}} \tag{3.2}$$

Suppose now that $\eta \overset{(1)}{\Longrightarrow} \zeta$. By definition, there is a sequence $\{\eta_0, \eta_1,..., \eta_n\}$ such that $\eta_0 = \eta$, $\eta_n = \zeta$ and $\sum_{i=0}^{n-1} C(\eta_i, \eta_{i+1}) = 1$. Let $l = \min\{i; \eta_i \in B^1(\zeta)\}$. We extend D to $\{\eta_0, \eta_1,..., \eta_{l-1}\} \cup B^1(\zeta)\backslash\{\zeta\}$ by defining $g(\eta_i) = \eta_{i+1}$, $i = 0, 1, 2,..., l-1$. Then extend D to $B^1(\eta) \cup B^1(\zeta)\backslash\{\zeta\}$ as we did in the proof of part (ii) of Lemma 8. After extension, (3.2) holds with $D = B^1(\eta) \cup B^1(\zeta)\backslash\{\zeta\}$. Next, choose another attractor ξ such that $\xi \overset{(1)}{\Longrightarrow} \zeta$ or $\xi \overset{(1)}{\Longrightarrow} \eta$ and repeat the procedure again. After a finite number of extensions we get a map g defined on the entire $B^2(\zeta)\backslash\{\zeta\}$ with the desired properties.

(iii) It is easy to prove that $W(B^2(\zeta)) \geqslant 2 + \sum_{\xi \in B^2(\zeta)} 1_{\{\xi \in A\}}$. We now construct a map $g \in G(B^2(\zeta))$ that reaches the minimum. By the symmetry between the two coordinates we may assume that there exists k such that either

$$\zeta((l,k))=1$$
$$\zeta((l,k+1))=\zeta((l,k+2))=\zeta((l,k+3))=0 \qquad \text{for all} \quad 1\leqslant l\leqslant N$$

or

$$\zeta((l,k))=\zeta((l,k+3))=1$$
$$\zeta((l,k+1))=\zeta((l,k+2))=0 \qquad \text{for all} \quad 1\leqslant l\leqslant N$$

In the first case, let $\zeta_0=\zeta$ and $\zeta_l=\zeta_{l-1}^{x_l}$, where $x_l=(l,k+1)$ for $l=1, 2,..., N$. Then ζ_N is a level 2 attractor; hence $\zeta_N\notin B^2(\zeta)$ by part (i). ζ_l is a level 1 attractor for $l=2, 3,..., N-2$; ζ_1 and ζ_{N-1} are not attractors. $C(\zeta_{N-1},\zeta_N)=0$ and $C(\zeta_l,\zeta_{l+1})=1$ for $l=0, 1, 2, 3,..., N-2$. We have

$$\sum_{l=0}^{N-1}(C(\zeta_l,\zeta_{l+1})-1_{\{\zeta_l\in A\}})=(N-1)-(N-3)=2$$

In the second case, let $x_{2l-1}=(l,k+1)$, $x_{2l}=(l,k+2)$ for $l=1, 2,..., N$, and $\zeta_0=\zeta, \zeta_n=\zeta_{n-1}^{x_n}$ for $n=1, 2,..., 2N$. Then $\zeta_{2N}\notin B^2(\zeta)$; ζ_{2l} is a level 1 attractor if $l=1, 2, 3,..., N-2$, and ζ_{2N-2} and ζ_{2l-1} are not attractors for $l=1, 2, 3,..., N$; $C(\zeta_n,\zeta_{n+1})=1$ if and only if $n=0, 1, 2, 4, 6,..., 2N-4$. We have

$$\sum_{n=0}^{2N-1}(C(\zeta_n,\zeta_{n+1})-1_{\{\zeta_n\in A\}})=N-(N-2)=2$$

In either case, define $g(\zeta_n)=\zeta_{n+1}$, $n=0, 1, 2,...$, and extend the domain of definition to the whole $B^2(\zeta)$, as we did in the proof of Lemma 8, so that $\sum_{\xi\in B^2(\zeta)}(C(\xi, g(\xi))-1_{\{\xi\in A\}})=2$. The desired conclusion now follows. ∎

4. LEMMAS FOR THEOREM 1

Let $K_1=S\backslash\{\mathbf{0},\mathbf{1}\}$,

$$S_2=\bigcup\{B^2(\eta)\mid \eta \text{ is a level 2 attractor}\}$$
$$S_1=\bigcup\{B^1(\eta)\mid \eta \text{ is a level 1 attractor}\}\backslash S_2$$
$$S_0=S\backslash(S_1\cup S_2)$$

Notice that S_0, S_1 and S_2 are mutually disjoint. This section is devoted to the proof of the following statement.

Lemma 11. We have

$$W(K_1)-W(K_1\backslash\{\xi\})\wedge\min_{\zeta\in K_1}W_{\xi\zeta}(K_1\backslash\{\zeta\})=\begin{cases}0 & \text{if} \quad \xi\in S_0\\ 1 & \text{if} \quad \xi\in S_1\\ 2 & \text{if} \quad \xi\in S_2\end{cases}$$

Lemma 12. $W(K_1)$ = number of level 1 attractors in $K_1 + 2\times$ number of level 2 attractors in K_1.

Proof. As observed in the proof of Lemma 10, if ζ is a level 2 attractor, ζ^x is not an attractor and $C(\zeta, \zeta^x) + C(\zeta^x, (\zeta^x)^y) = 2$ for any $x, y \in \mathbf{T}$, $x \neq y$. Suppose that ζ and η are level 2 attractors, $g \in G(K_1)$, $g(\zeta) = \zeta^x$, $g(\eta) = \eta^y$ for some $x, y \in \mathbf{T}$; then $g(\zeta) \neq g(\eta)$. Otherwise $\{\zeta, \zeta^x, \eta\}$ is a sequence leading ζ to η, and $C(\zeta^x, \eta) = C(\eta^y, \eta) = 0$. This implies that $\zeta \overset{(1)}{\Longrightarrow} \eta$, contradicting the fact that ζ is a level 2 attractor. Therefore for any $g \in G(K_1)$,

$$\begin{aligned}\sum_{\zeta \in K_1} C(\zeta, g(\zeta)) \geqslant & \sum_{\substack{\zeta \text{ is a level} \\ 1 \text{ attractor}}} C(\zeta, g(\zeta)) \\ & + \sum_{\substack{\zeta \text{ is a level} \\ 2 \text{ attractor}}} [C(\zeta, g(\zeta)) + C(g(\zeta), g(g(\zeta)))] \\ & = \text{number of level 1 attractors} \\ & \quad + 2 \times \text{number of level 2 attactors}\end{aligned}$$

and $W(K_1)$ shares the same lower bound.

For every level 2 attractor we constructed in the proof of part (iii) of Lemma 10 a sequence $\{\zeta_i; i = 0, 1, 2, 3,...\}$ leading the level 2 attractor $\zeta = \zeta_0$ to another level 2 attractor or $\mathbf{1}$. Let D be the union of these sequences. Define $g(\zeta_i) = \zeta_{i+1}$ for $\zeta_i \in D$. Then g maps no subset of D into itself because $\{x \in \mathbf{T};\ \zeta_i(x) = 1\}$ is increasing in i. Furthermore

$$\begin{aligned}\sum_{\xi \in D} C(\xi, g(\xi)) - 1_{\{\xi \in \mathbf{A}\}}) &= \sum_{\substack{\zeta \text{ is a level} \\ 2 \text{ attractor}}} \sum_{l} (C(\zeta_l, \zeta_{l+1}) - 1_{\{\zeta_l \in \mathbf{A}\}}) \\ &= \sum_{\substack{\zeta \text{ is a level} \\ 2 \text{ attractor}}} 2 = 2 \times \text{number of level 2 attactors} \qquad (4.1)\end{aligned}$$

Extend D to K_1 and keep (4.1), as we did in the proof of Lemma 8. Then $g \in G(K_1)$ and $\sum_{\zeta \in K_1} C(\zeta, g(\zeta))$ reaches the lower bound. This proves the lemma. ∎

Lemma 13. We have

$$W(K_1 \backslash \{\zeta\}) = \begin{cases} W(K_1) - 2 & \text{if } \zeta \text{ is a level 2 attractor} \\ W(K_1) - 1 & \text{if } \zeta \text{ is a level 1 attractor or if } \zeta^x \text{ is} \\ & \quad \text{a level 2 attractor for some } x \in \mathbf{T} \\ W(K_1) & \text{otherwise} \end{cases}$$

Furthermore, $W_{\xi\zeta}(K_1\backslash\{\zeta\})=W(K_1\backslash\{\zeta\})$ if ζ is a level 2 attractor and $\xi\in B^2(\zeta)$, or if ζ is a level 1 attractor and $\xi\in B^1(\zeta)$.

The proof is identical with that of Lemmas 10 and 12, so we skip the details.

Proof of Lemma 11. If $\xi\in S_2$, there is a level 2 attractor ζ such that $\xi\in B^2(\zeta)$. By Lemma 13,

$$W_{\xi\zeta}(K_1\backslash\{\zeta\})=W(K_1\backslash\{\zeta\})=W(K_1)-2$$

$$W_{\xi\eta}(K_1\backslash\{\eta\})\geqslant W(K_1\backslash\{\eta\})\geqslant W(K_1)-2 \qquad \text{for any} \quad \eta\in K_1$$

So

$$W(K_1)-W(K_1\backslash\{\xi\})\wedge\min_{\zeta\in K_1}W_{\xi\zeta}(K_1\backslash\{\zeta\})$$
$$=[W(K_1)-W(K_1\backslash\{\xi\})]\vee\max_{\zeta\in K_1}[W(K_1)-W_{\xi\zeta}(K_1\backslash\{\zeta\})]=2$$

If $\xi=S_1$, there is a level 1 attractor ζ such that $\xi\in B^1(\zeta)$. By Lemma 13, $W(K_1)-W_{\xi\zeta}(K_1\backslash\{\zeta\})=1$; and for any $\eta\in K_1$ other than a level 2 attractor

$$W_{\xi\eta}(K_1\backslash\{\eta\})\geqslant W(K_1\backslash\{\eta\})\geqslant W(K_1)-1 \tag{4.2}$$

If (4.2) also holds for all level 2 attractors, then

$$W(K_1)-W(K_1\backslash\{\xi\}\wedge\min_{\zeta\in K_1}W_{\xi\zeta}(K_1\backslash\{\zeta\})$$
$$=[W(K_1)-W(K_1\backslash\{\xi\})]\vee\max_{\zeta\in K_1}[W(K_1)-W_{\xi\zeta}(K_1\backslash\{\zeta\})]=1$$

Suppose now that ζ' is a level 2 attractor and $g\in G_{\xi\zeta'}(K_1\backslash\{\zeta'\})$. Let $\{\xi_0,\xi_1,...,\xi_n\}$ be the sequence leading $\xi=\xi_0$ to $\zeta'=\xi_n$. Namely, $\xi_{k+1}=g(\xi_k)$, $k=0,1,2,...,n-1$. Suppose that ξ_l is the first attractor in the sequence. If $\sum_{i=0}^{l-1}C(\xi_i,\xi_{i+1})=0$ and $\sum_{i=0}^{n-1}C(\xi_i,\xi_{i+1})=\sum_{i=l}^{n-1}1_{\{\xi_i\in\Lambda\}}$, it would follow that $\xi\in B^1(\xi_l)\subset B^2(\zeta')$, contradicting the fact that $\xi\notin S_2$. Thus

$$\sum_{i=0}^{n-1}C(\xi_i,\xi_{i+1})\geqslant 1+\sum_{i=0}^{n-1}1_{\{\xi_i\in\Lambda\}}$$

and

$$\sum_{\eta\in K_1\backslash\{\zeta'\}}C(\eta,g(\eta))=\sum_{i=0}^{n-1}C(\xi_i,\xi_{i+1})+\sum_{\eta\neq\xi_i}C(\eta,g(\eta))$$
$$\geqslant 1+W(K_1\backslash\{\zeta'\})=W(K_1)-1$$

Therefore $W_{\xi\zeta'}(K_1\backslash\{\zeta'\}) \geqslant W(K_1)-1$. This verifies (4.2) and completes the analysis of the case that $\xi \in S_1$.

We skip the similar proof of the case that $\xi \in S_0$. ▮

5. LEMMAS FOR THEOREM 2

Lemma 14. Let $K_2 = S\backslash\{\mathbf{1}\} = K_1 \cup \{\mathbf{0}\}$. Then

$$W(K_2) - W(K_2\backslash\{\mathbf{0}\}) \wedge \min_{\eta \in K_1} W_{0\eta}(K_2\backslash\{\eta\}) = N$$

The proof is very long, and is divided into three parts. First, consider a sequence $\{x_i, i=1, 2,..., 2N\}$ consisting of

$$(1, 1), (1, 2), (1, 3),..., (1, N), (2, 1), (2, 2), (2, 3),..., (2, N) \tag{5.1}$$

(with possibly different order). Then $\{\zeta_0 = \mathbf{0}, \zeta_i = \zeta_{i-1}^{x_i}, i=1, 2,..., 2N\}$ is a sequence leading $\mathbf{0}$ to the level 2 attractor ζ_{2N} (Fig. 3a). ζ_{2m} is a level 1 attractor (Fig. 3b) if

$$\{x \in \mathbf{T}; \zeta_{2m}(x) = 1\} = \bigcup_{j=1}^{k} \{(1, a_j), (2, a_j),..., (1, b_j), (2, b_j)\}$$

$$\sum_{j=1}^{k} (b_j - a_j + 1) = m; \qquad b_j - a_j \geqslant 1, \quad a_{j+1} - b_j \geqslant 3 \quad \forall j \tag{5.2}$$

Lemma 15. We have

$$\sum_{j=0}^{2N-1} (C(\zeta_j, \zeta_{j+1}) - 1_{\{\zeta_j \in \mathbf{A}\}}) \geqslant N$$

In general, if ζ_{2m} is of form (5.2), then

$$\sum_{j=0}^{2m-1} (C(\zeta_j, \zeta_{j+1}) - 1_{\{\zeta_j \in \mathbf{A}\}}) \geqslant m + k \tag{5.3}$$

Proof. If ζ_{2m} is the first level 1 attractor in the sequence, then $\sum_{i=0}^{2m-1} 1_{\{\zeta_i \in \mathbf{A}\}} = 0$ and $C(\zeta_{i-1}, \zeta_i) = 1$ if x_i is the first of $(1, s)$ and $(2, s)$ in the sequence or $x_i = (1, a_j), (2, a_j), (1, b_j)$, or $(2, b_j)$ for some j. So (5.3) is true.

We now apply induction. Suppose (5.3) is true for attractor ζ_{2m}, and $\zeta_{2m'}$ is the next level 1 attractor in the sequence and is of the form (5.2).

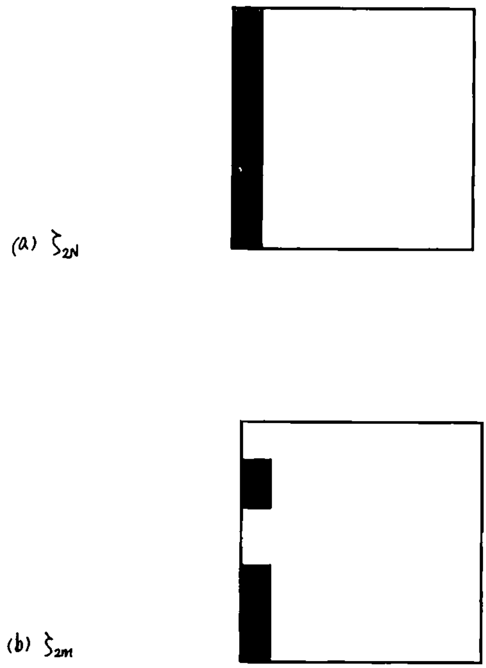

Fig. 3. (a) ζ_{2N} and (b) ζ_{2m} of form (5.2) $\{x \in \mathbf{T}; \zeta_i(x) = 1\}$ is the shaded area.

Any sequence $\{\zeta_{2m}, \zeta_{2m+1}, ..., \zeta_{2m'}\}$ leading ζ_{2m} to $\zeta_{2m'}$ is virtually one of the following three cases (Fig. 4):

(i)

$$\{x \in \mathbf{T}; \zeta_{2m}(x) = 1\} = \{(1, 1), (2, 1), (1, 2), (2, 2), ..., (1, m), (2, m)\}$$

$$\{x \in \mathbf{T}; \zeta_{2m'}(x) = 1\} = \{(1, 1), (2, 1), (1, 2), (2, 2), ..., (1, m'), (2, m')\}$$

Assume that $\zeta_i = \zeta_{i-1}^{x_i}$. Then $C(\zeta_{j-1}, \zeta_j) = 1$ if $x_j = (1, m'), (2, m')$, and first of $(1, s)$ and $(2, s)$ in the sequence for $m + 1 \leqslant s \leqslant m' - 1$. Note again that ζ_j is not an attractor for $2m < j < 2m'$, but ζ_{2m} is. Hence

$$\sum_{j=2m}^{2m'-1} C(\zeta_j, \zeta_{j+1}) \geqslant m' - m + 1, \qquad \sum_{j=2m}^{2m'-1} 1_{\{\zeta_j \in \Lambda\}} = 1$$

By the induction hypothesis, (5.3) holds with $k = 1$. We have

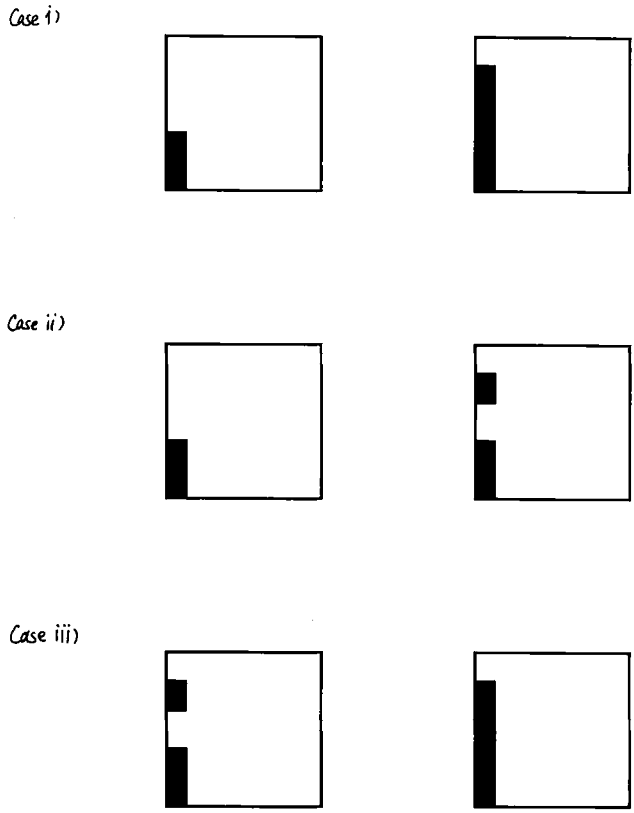

Fig. 4. Three basic cases of ζ_{2m} (left) and $\zeta_{2m'}$ (right).

$$\sum_{j=2m}^{2m'-1} \left(C(\zeta_j, \zeta_{j+1}) - 1_{\{\zeta_j \in \Lambda\}} \right)$$

$$= \sum_{j=0}^{2m-1} + \sum_{j=2m}^{2m'-1} \geqslant (m+1) + (m' - m + 1 - 1) = m' + 1$$

(ii) Assume that ζ_{2m} is the same as case (i) and (5.3) holds with $k = 1$, and that

$$\{x \in \mathbf{T}; \zeta_{2m'}(x) = 1\}$$
$$= \{(1, 1), (2, 1), (1, 2), (2, 2), \ldots, (1, m), (2, m)\}$$
$$\cup \{(1, a+1+m), (2, a+1+m), \ldots, (1, a+m'), (2, a+m')\}$$

Then $C(\zeta_{j-1}, \zeta_j) = 1$ if $x_j = (1, m+a+1)$, $(2, m+a+1)$, $(1, m'+a)$, $(2, m'+a)$, as well as the first $(1, s)$ and $(2, s)$ in the sequence for $m+a+2 \leqslant s \leqslant a+m'-1$. We have

$$\sum_{j=2m}^{2m'-1} (C(\zeta_j, \zeta_{j+1}) - 1_{\{\zeta_j \in A\}}) \geqslant m' - m + 2 - 1$$

$$\sum_{j=0}^{2m'-1} = \sum_{j=0}^{2m-1} + \sum_{j=2m}^{2m'-1} \geqslant (m+1) + (m'-m+1) = m'+2$$

(iii) Suppose that $\zeta_{2m'}$ is the same case (i),

$$\{x \in \mathbf{T}; \zeta_{2m}(x) = 1\}$$
$$= \{(1, 1), (2, 1), (1, 2), (2, 2), ..., (1, a), (2, a)\}$$
$$\cup \{(1, a+1+m'-m), (2, a+1+m'-m), ..., (1, m'), (2, m')\}$$

and (5.3) holds with $k=2$. Then $C(\zeta_{j-1}, \zeta_j) = 1$ if x_j is the first of $(1, s)$ and $(2, s)$ in the sequence for $a+1 \leqslant s \leqslant a+(m'-m)$. We have

$$\sum_{j=2m}^{2m'-1} (C(\zeta_j, \zeta_{j+1}) - 1_{\{\zeta_j \in A\}}) \geqslant m' - m - 1$$

$$\sum_{j=0}^{2m'-1} = \sum_{j=0}^{2m-1} + \sum_{j=2m}^{2m'-1} \geqslant m+2+(m'-m-1) = m'+1$$

We have shown that (5.3) holds for all ζ_{2m} of form (5.2). Suppose now that ζ_l is the first configuration (see Fig. 5) in the sequence such that

$$\{x \in \mathbf{T}; \zeta_l(x) = 1\} \supset \{(1, 1), (1, 2), ..., (1, N)\}$$
$$\text{or} \quad \{(2, 1), (2, 2), ..., (2, N)\}$$

Suppose that ζ_{2m} is the last level 1 attractor of form (5.2) before ζ_l in the sequence. (In the extreme case $\zeta_{2m} = \mathbf{0}$.) Then ζ_j is not an attractor for

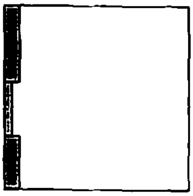

Fig. 5. The first configuration ζ_l in the sequence $\{x \in \mathbf{T}: \zeta_l(x) = 1\} \supset \{(1, 1), (1, 2), ..., (1, N)\}$ or $\{(2, 1), (2, 2), ..., (2, N)\}$.

$2m < j < l$. Note that $C(\zeta_{j-1}, \zeta_j) = 1$ if x_j is the first of $(1, s)$ and $(2, s)$ in the sequence and $\zeta_{2m}((1, s)) = 0$. There are $N - m$ such s's. Hence,

$$\sum_{j=2m}^{l-1} (C(\zeta_j, \zeta_{j+1}) - 1_{\{\zeta_j \in \Lambda\}}) \geqslant N - m - 1$$

$$\sum_{j=0}^{2N-1} \geqslant \sum_{j=0}^{2m-1} + \sum_{j=2m}^{l-1} \geqslant m + 1 + N - m - 1 = N \quad \blacksquare$$

Lemma 16. If the sequence $\{\eta_0, \eta_1, ..., \eta_n\}$ leads $\mathbf{0}$ to the level 2 attractor η_n, then

$$\sum_{j=0}^{n-1} (C(\eta_j, \eta_{j+1}) - 1_{\{\eta_j \in \Lambda\}}) \geqslant N$$

Furthermore, $W(K_2) \geqslant W(K_1) + N$ and $W(B^2(\mathbf{0})) \geqslant N + \sum_{\xi \in B^2(\mathbf{0})} 1_{\{\xi \in \Lambda\}}$.

Proof. Assume that $\eta_0 = \mathbf{0}$, $\eta_j = \eta_{j-1}^{y_j}$, $j = 1, 2, ..., n$, and η_n is a level 2 attractor. Note that the same y may appear two or more times in the sequence $\{y_j, j = 1, 2, ..., n\}$. Let us call the subset $\{(k, 1), (k, 2), ..., (k, N)\}$ a *column* of $\mathbf{T}$ and the subset $\{(1, k), (2, k), ..., (N, k)\}$ a *row* of $\mathbf{T}$. Because of the symmetry between column and row, we shall only treat columns. Note also that $\{y \in \mathbf{T}; \eta_n(y) = 1\}$ contains at least two points in every row, and $n \geqslant 2N$. Write y_j as $(y_j(1), y_j(2))$. Define a map ϕ: $\{y_j, 1 \leqslant j \leqslant n\} \to \mathbf{T} \cup \{\infty\}$ as follows:

$$\phi(y_j) = \begin{cases} (1, s) & \text{if } y_j \text{ is the first point of row } s \text{ in the sequence} \\ & \quad \text{and } y_j(1) \text{ is odd, } y_j(2) = s \\ & \text{or if } y_j \text{ is the second point of row } s \text{ in the sequence,} \\ & \quad y_j(2) = s,\ y_i \text{ the first point of row } s \text{ in the} \\ & \quad \text{sequence, } y_i(1) \text{ is even, and } y_i \neq y_j \\ (2, s) & \text{if } y_j \text{ is the first of row } s \text{ in the sequence,} \\ & \quad y_j(2) = s, \text{ and } y_j(1) \text{ is even} \\ & \text{or if } y_j \text{ is the second point of row } s \text{ in the sequence,} \\ & \quad y_j(2) = s,\ y_i \text{ the first point of row } s \text{ in the} \\ & \quad \text{sequence, and } y_i(1) \text{ is odd, and } y_i \neq y_j \\ \infty & \text{otherwise} \end{cases}$$

Map ϕ translates every y_j horizontally to the first two columns or the cemetery (∞), with little attention to the parity of the first coordinate of y_j. The fact that $\{y \in \mathbf{T}; \eta_n(y) = 1\}$ contains at least two points in every row guarantees that the subsequent $\phi(y_{j(i)})$, after deleting all ∞'s, is a sequence of (5.1). Rewrite $\phi(y_{j(i)})$ as x_i, $i=1, 2,..., 2N$ and let $\zeta_0 = \mathbf{0}$, $\zeta_i = \zeta_{i-1}^{x_i}$. Since y_j's are more spread and x_i's are more concentrated, x_i is likely to have more neighbors than the corresponding $y_{j(i)}$, and ζ_i is more likely to be a level 1 attractor than the corresponding $\eta_{j(i)}$ is. That is,

$$C(\zeta_i, \zeta_{i+1}) \leqslant C(\eta_{j(i)}, \eta_{j(i)+1}) \qquad \text{and} \qquad 1_{\{\zeta \in \Lambda\}} \geqslant 1_{\{\eta_{j(i)} \in \Lambda\}}$$

We have

$$\begin{aligned}\sum_{j=0}^{n-1} (C(\eta_j, \eta_{j+1}) - 1_{\{\eta_j \in \Lambda\}}) &\geqslant \sum_{i=0}^{2N-1} (C(\eta_{j(i)}, \eta_{j(i)+1}) - 1_{\{\eta_{j(i)} \in \Lambda\}}) \\ &\geqslant \sum_{i=0}^{2N-1} (C(\zeta_i, \zeta_{i+1}) - 1_{\{\zeta_i \in \Lambda\}}) \geqslant N\end{aligned}$$

by Lemma 15. This proves the first part.

Now take an arbitrary $g \in G(K_2)$. Define $\zeta_0 = \mathbf{0}$ and $\zeta_k = g(\zeta_{k-1})$ for $k = 1, 2, 3,....$ Suppose that $l = \min\{n; \zeta_n \text{ is a level 2 attractor}\}$. Let

$$K' = \{\zeta_k, k=0, 1, 2,..., l-1\} \qquad \text{and} \qquad K'' = K_2 \backslash K'$$

Notice that $g(\zeta_l) \in K''$. It is of the same idea as in Lemma 12 to prove that

$$\sum_{\xi \in K''} C(\xi, g(\xi)) \geqslant \sum_{\xi \in K''} 1_{\{\xi \in \Lambda\}} + 2 \times \text{number of level 2 attractors in } K''$$

Furthermore, K'' contains all level 2 attractors. So, by the first part of this lemma and Lemma 12,

$$\begin{aligned}&\sum_{\xi \in S \backslash \{\mathbf{1}\}} C(\xi, g(\xi)) \\ &\quad = \sum_{\xi \in K'} + \sum_{\xi \in K''} \geqslant N + \sum_{\xi \in K'} 1_{\{\xi \in \Lambda\}} + \sum_{\xi \in K''} 1_{\{\xi \in \Lambda\}} \\ &\qquad + 2 \times \text{number of level 2 attractors in } K'' = W(K_1) + N\end{aligned}$$

Taking the minimum over all g's of $G(K_2)$, we get the desired conclusion. The proof of the last inequality is very similar. ∎

Proof of Lemma 14. Let $\{x_j; j=1, 2,..., 2N\}$ be the sequence of $\mathbf{T}$ given as follows:

$$(1, 1), (2, 2), (1, 3), (2, 4), ...; (1, 2), (2, 3), (1, 4),..., (2, 1)$$

Then the corresponding sequence $\{\zeta_0 = \mathbf{0}, \zeta_i = \zeta_{i-1}^{x_i}, i = 0, 1, 2, ..., 2N\}$ leads $\mathbf{0}$ to the level 2 attractor ζ_{2N}. It is easy to see that $C(\zeta_i, \zeta_{i+1}) = 1$ if $0 \leqslant i \leqslant N-1$ or $i = 2N-2$ when N is odd. ζ_i is not a level 1 attractor unless N is odd and $i = 2N-2$. We have

$$\sum_{l=0}^{2N-1} (C(\zeta_l, \zeta_{l+1}) - 1_{\{\zeta_l \in \Lambda\}}) = \begin{cases} N - 0 & \text{if } N \text{ is even} \\ (N+1) - 1 & \text{if } N \text{ is odd} \end{cases} = N$$

Using this sequence and the sequences constructed previously in the proof of Lemma 10, we define g first on the union of these sequences, then extend the domain of definition to K_2, so that $g \in G(K_2)$ and

$$\sum_{\xi \in K_2} (C(\xi, g(\xi)) - 1_{\{\xi \in \Lambda\}}) = N + \text{the number of level 2 attractors}$$

In light of Lemma 16, we conclude $W(K_2) = W(K_1) + N$.

If $\eta \notin B^2(\mathbf{0})$, it follows from Lemmas 13 and 16 that

$$\begin{aligned} W_{\mathbf{0}\eta}(K_2 \backslash \{\eta\}) &\geqslant W(B^2(\mathbf{0})) + W(K_1 \backslash (B^2(\mathbf{0}) \cup \{\eta\})) \\ &\geqslant N + \sum_{\xi \in B^2(\mathbf{0})} 1_{\{\xi \in \Lambda\}} + W(K_1 \backslash (B^2(\mathbf{0}) \cup \{\eta\}) \\ &\geqslant N + W(K_1) - 2 > W(K_1) \end{aligned}$$

If $\eta \in B^2(\mathbf{0})$, $\eta \neq \mathbf{0}$, then

$$\begin{aligned} W_{\mathbf{0}\eta}(K_2 \backslash \{\eta\}) &\geqslant W(B^2(\mathbf{0}) \backslash \{\eta\}) + W(K_1 \backslash B^2(\mathbf{0})) \\ &\geqslant 1 + \sum_{\xi \in B^2(\mathbf{0}) \backslash \{\eta\}} 1_{\{\xi \in \Lambda\}} + W(K_1 \backslash B^2(\mathbf{0})) \geqslant W(K_1) \end{aligned}$$

In either case we conclude that

$$W(K_2) - W_{\mathbf{0}\eta}(K_2 \backslash \{\eta\}) \leqslant W(K_2) - W(K_1) = N$$

and that

$$\begin{aligned} &W(K_2) - W(K_2 \backslash B^2\{\mathbf{0}\}) \wedge \min_{\eta \in K_1} W_{\mathbf{0}\eta}(K_2 \backslash \{\eta\}) \\ &\quad = [W(K_2) - W(K_2 \backslash B^2\{\mathbf{0}\})] \vee \max_{\eta \in K_1} [W(K_2) - W_{\mathbf{0}\eta}(K_2 \backslash \{\eta\})] = N \quad \blacksquare \end{aligned}$$

6. PROOF OF THEOREMS

Proof of Theorem 1. Applying (2.7) to $K_1 = S\backslash\{\mathbf{0},\mathbf{1}\}$ and by Lemma 11, we get

$$\lim_{\delta\to 0}\frac{\log E_\xi T_1}{-\log\delta}$$

$$= W(K_1) - W(K_1\backslash\{\xi\}) \wedge \min_{\zeta\in K_1} W_{\xi\zeta}(K_1\backslash\{\zeta\}) = \begin{cases} 0 & \text{if } \xi\in S_0 \\ 1 & \text{if } \xi\in S_1 \\ 2 & \text{if } \xi\in S_2 \end{cases}$$

$E_\xi T_1$ is a ratio of two huge polynomials of δ, by (2.2). Every $\pi(g)$ is either 0 or of form $\delta^a(1-\delta)^b$ for some very large a and b. The coefficients of the polynomials are integers. The above limit is the difference of the minimum exponents of the numerator and denominator. Hence, for $\xi\in S_i$, $i=0$, 1, or 2, $\delta^i E_\xi T_1$ converges to the ratio of the coefficients of the minimum exponents of δ, which is a rational number (depending on ξ). ∎

Proof of Theorem 2. By (2.7) and Lemma 14, with $K_2 = S\backslash\{\mathbf{1}\}$,

$$\lim_{\delta\to 0}\frac{\log E_{\mathbf{0}}\tau(S\backslash\{\mathbf{1}\})}{-\log\delta} = W(K_2) - W(K_2\backslash\{\mathbf{0}\}) \wedge \min_{\zeta\in K_2} W_{\mathbf{0}\zeta}(K_2\backslash\{\zeta\}) = N$$

Namely, $E_{\mathbf{0}}\tau(S\backslash\{\mathbf{1}\})$ is of the order $(1/\delta)^N$. By the strong Markov property and the symmetry between $\mathbf{0}$ and $\mathbf{1}$,

$$E_\xi T_2 = E_\xi T_1 + E_{\mathbf{0}}\tau(S\backslash\{\mathbf{1}\}), \qquad \forall\xi\in S \tag{6.1}$$

$E_\xi T_1$ is at most of order $(1/\delta)^2$, by Theorem 1. Therefore $E_\xi T_2$ is of order $(1/\delta)^N$. Again $E_\xi T_2$ is a ratio of two huge polynomials of δ, by (2.2). All coefficients are integers. N is the difference of the minimum exponents of the numerator and denominator. Hence $\delta^N E_\xi T_2$ converges to a rational number (independent of ξ).

Starting at $\mathbf{0}$, with very small probability, $\{\xi_t\}$ will visit $\mathbf{1}$ before returning to $\mathbf{0}$. Thus $\{\xi_t\}$ returns to $\mathbf{0}$ many times before hitting $\mathbf{1}$. By the strong Markovian property, each excursion is independent of others. Thus the occurence time T_2 of the rare event, scaled appropriately, is exponentially distributed. The following is a rigorous proof of this observation. It follows from (6.1) that $\forall\xi\in S$

$$\lim_{\delta\to 0}\frac{E_\xi^\delta T_2}{E_{\mathbf{0}}^\delta T_2} = 1 \tag{6.2}$$

Let $A_\delta = E_{\mathbf{0}}^\delta T_2$ and $\mu_\xi^\delta(\cdot)$ be the probability distribution of T_2/A_δ starting at ξ. Then

$$\mu_\xi^\delta\left(\left[0, \frac{2}{\varepsilon}\right]\right) = 1 - P_\xi^\delta\left(\frac{T_2}{A_\delta} > \frac{2}{\varepsilon}\right) \geqslant 1 - \frac{\varepsilon}{2}\frac{E_\xi^\delta T_2}{A_\delta}$$

By (6.2) there exists δ_0 such that as $\delta \leqslant \delta_0, E_\xi^\delta T_2/A_\delta \leqslant 2$. Hence $\mu_\xi^\delta([0, 2/\varepsilon]) \geqslant 1-\varepsilon$ and $\{\mu_\xi^\delta(\cdot); \delta_0 \geqslant \delta > 0\}$ is tight. Consequently we can choose a sequence $\{\delta_n\}$ such that $\mu_\xi^{\delta_n}(\cdot)$ converges weakly to a distribution $\mu_\xi(\cdot)$ for $\forall \xi \in S$ as $n \to +\infty$. Let

$$F = \{\gamma \geqslant 0 \mid \exists \delta_n \text{ and } \xi \text{ such that } \mu_\xi^{\delta_n}(\{\gamma\}) > 0\}$$

Suppose that $0 < \varepsilon < 1, 0 < \alpha$, and $\alpha \notin F$. By the strong Markovian property, the asymptotic estimate of T_1 and the symmetry between $\mathbf{0}$ and $\mathbf{1}$,

$$\begin{aligned}
\mu_\xi([\alpha, +\infty)) &= \lim_{n\to\infty} \mu_\xi^{\delta_n}([\alpha, +\infty)) \\
&= \lim_{n\to\infty} P_\xi^{\delta_n}(T_1 \geqslant A_{\delta_n}\alpha) \\
&= \lim_{n\to\infty} P_\xi^{\delta_n}(T_2 \geqslant A_{\delta_n}\alpha > (1/\delta_n)^{N-\varepsilon} \geqslant T_1) \\
&= \lim_{n\to\infty} P_\xi^{\delta_n}(\xi_{T_1} = \mathbf{0})\, P_{\mathbf{0}}^{\delta_n}(T_2 \geqslant A_{\delta_n}\alpha - (1/\delta_n)^{N-\varepsilon}) \\
&\quad + \lim_{n\to\infty} P_\xi^{\delta_n}(\xi_{T_1} = \mathbf{1})\, P_{\mathbf{1}}^{\delta_n}(T_2 \geqslant A_{\delta_n}\alpha - (1/\delta_n)^{N-\varepsilon}) \\
&= \mu_{\mathbf{0}}([\alpha, +\infty))
\end{aligned}$$

Recall $\sigma(\mathbf{1})$ defined in (1.2). Similarly, if $\beta \notin F$,

$$\begin{aligned}
&\lim_{n\to\infty} P_\eta^{\delta_n}(\sigma(\mathbf{1}) \geqslant \beta A_{\delta_n}) \\
&\quad = \lim_{n\to\infty} P_\eta^{\delta_n}(\sigma(\mathbf{1}) \geqslant \beta A_{\delta_n} > (1/\delta_n)^{2+\varepsilon} > T_1) \\
&\quad = \lim_{n\to\infty} P_{\mathbf{0}}^{\delta_n}(\sigma(\mathbf{1}) \geqslant \beta A_{\delta_n} - (1/\delta_n)^{2+\varepsilon}) \\
&\quad = \mu_{\mathbf{0}}([\beta, +\infty))
\end{aligned}$$

Now simply write $\mu_0([\alpha, +\infty)) = \mu(\alpha)$. For $\alpha, \beta, \alpha+\beta \notin F$,

$$\begin{aligned}
\mu(\alpha+\beta) &= \lim_{n\to\infty} P_{\mathbf{0}}^{\delta_n}(T_2 \geq A_{\delta_n}(\alpha+\beta)) \\
&= \lim_{n\to\infty} \sum_{\eta\neq \mathbf{1}} P_{\mathbf{0}}^{\delta_n}(T_2 \geq \alpha A_{\delta_n}, \xi_{\alpha A_{\delta n}} = \eta) \lim_{n\to\infty} P_{\eta}^{\delta_n}(\sigma(\mathbf{1}) \geq \beta A_{\delta_n}) \\
&= \lim_{n\to\infty} \sum_{\eta\neq \mathbf{1}} P_{\mathbf{0}}^{\delta_n}(T_2 \geq \alpha A_{\delta_n}, \xi_{\alpha A_{\delta n}} = \eta)\, \mu(\beta) \\
&= \lim_{n\to\infty} P_{\mathbf{0}}^{\delta_n}(T_2 \geq \alpha A_{\delta_n})\, \mu(\beta) = \mu(\alpha)\, \mu(\beta) \qquad (6.3)
\end{aligned}$$

To show that (6.3) holds for any $\alpha, \beta > 0$, notice that F is at most a countable set. Take sequences $\alpha_n \nearrow \alpha$ and $\beta_n \nearrow \beta$ such that $\alpha_n, \beta_n, \alpha_n + \beta_n \notin F$ and (6.3) holds for all α_n, β_n. Since $\mu(\cdot)$ is left continuous, we obtain that

$$\mu(\alpha+\beta) = \lim_{n\to\infty} \mu(\alpha_n+\beta_n) = \lim_{n\to\infty} \mu(\alpha_n)\, \mu(\beta_n) = \mu(\alpha)\, \mu(\beta)$$

So $\mu(\alpha) = Ce^{-a\alpha}$. That $\mu(0) = 1$ and $\int_0^\infty \alpha\mu(d\alpha) = 1$ implies $\mu(\alpha) = e^{-\alpha}$.

Finally, by the uniqueness of the weak limit of $\{\mu_{\xi}^{\delta_n}(\cdot)\}$, $T_2/E_{\xi}^{\delta} T_2$ converges in law to the exponential distribution with mean 1. ∎

Proof of Corollary 3. We have

$$\lim_{\delta\to 0} \frac{\log T_2}{-\log\delta} = \lim_{\delta\to 0} \frac{\log(T_2/ET_2)}{-\log\delta} + \lim_{\delta\to 0} \frac{\log ET_2}{-\log\delta} = 0 + N = N \quad ∎$$

Remark on (1.3). We claim that (1.3) holds if $\xi \in S\backslash(B^2(\mathbf{0}) \cup B^2(\mathbf{1}))$. By the Chebyshev inequality and Theorem 1,

$$\lim_{\delta\to 0} P_{\xi}\left(\frac{\log T_1}{-\log\delta} < 2 + \varepsilon\right) = 1 \qquad \text{if} \quad \xi \in S_2$$

To obtain the other half, let $\mathbf{A}_2 = \{$all level 2 attractors and $\mathbf{0}, \mathbf{1}\}$. If $\xi \in S\backslash(B^2(\mathbf{0}) \cup B^2(\mathbf{1}))$, then $P_{\xi}(\xi_{\sigma(\mathbf{A}_2)} = \mathbf{0}$ or $\mathbf{1}) \to 0$ by (2.6), and $T_1 = \sigma(\mathbf{A}_2) + T'$, where T' is the first exit time of K_1 by $\{\xi_t\}$ starting at $\xi_{\sigma(\mathbf{A}_2)}$. Let $\zeta = \xi_{\sigma(\mathbf{A}_2)}$ and $K_3 = B^2(\zeta)\backslash \bigcup \{B^2(\theta) | \theta \neq \zeta, \theta \in \mathbf{A}_2\}$. Then

$$\lim_{\delta\to 0} \frac{E_{\xi}^{\delta}\tau(K_3)}{-\log\delta} = 2, \qquad \lim_{\delta\to 0} \frac{E_{\xi}^{\delta}\tau(K_3)}{E_{\eta}^{\delta}\tau(K_3)} = 1 \qquad \text{for any} \quad \xi, \eta \in K_3$$

It is similar to prove that $\tau(K_3)/E^{\delta}_{\zeta}(K_3)$ converges in law to the exponential distribution with mean 1. Since $\zeta \in K_3 \subset K_1$, $T' \geqslant \tau(K_3)$ and

$$\lim_{\delta \to 0} \frac{\log T'}{-\log \delta} \geqslant \lim_{\delta \to 0} \frac{\log \tau(K_3)}{-\log \delta}$$

$$= \lim_{\delta \to 0} \frac{\log(\tau(K_3)/E_\tau(K_3))}{-\log \delta} + \lim_{\delta \to 0} \frac{E_\tau(K_3)}{-\log \delta} = 2$$

Thus we have shown that (1.3) holds if $\xi \in S\backslash(B^2(\mathbf{0}) \cup B^2(\mathbf{1}))$. By the same argument we prove that

$$\lim_{\delta \to 0} P\left(\left|\frac{\log T_1}{-\log \delta} - 1\right| \leqslant \varepsilon\right) = 1 \qquad \text{if} \quad \xi \in (B^2(\mathbf{0}) \cup B^2(\mathbf{1}))\backslash(B^1(\mathbf{0}) \cup B^1(\mathbf{1}))$$

Remark on (1.4). A similar (and simpler) discussion can be carried out if the torus $\{1, 2, 3,..., N\} \times \{1, 2, 3,..., N\}$ is replaced by the circle $\{1, 2, 3,..., N\}$. In the one-dimensional case, configuration ξ is an attractor if and only if

$$|\xi(x-1) - \xi(x)| + |\xi(x+1) - \xi(x)| \leqslant 1 \qquad \text{for} \quad x = 1, 2,..., N$$

Except for $\xi = \mathbf{0}$ or $\mathbf{1}$, $\xi \xrightarrow{(1)} \eta$ holds for every pair of attractors (ξ, η). It is easy to find a sequence $\{\zeta_0, \zeta_1,..., \zeta_n\}$ leading $\mathbf{0}$ to an attractor that $\sum_{i=0}^{n-1} C(\zeta_i, \zeta_{i+1}) = 2$. These facts together imply (1.4). Comparing the two-dimensional torus with the one-dimensional circle, one can also sense the complexity in dealing with the higher dimensional case. We conjecture that $\lim_{\delta \to 0}(ET_2)/(-\log \delta) = N^2$ in the three-dimensional case.

ACKNOWLEDGMENTS

This work is done while the author was visiting UCLA. The author wishes to thank Prof. Tom Liggett for his hospitality and for his insightful comments on an earlier draft of this paper. The author also thanks a referee whose critical comments helped to improve the presentation. This work was supported in part by a grant from the NSF of China.

REFERENCES

1. R. Bott and J. P. Mayberry, *Matrices and trees*, (Wiley, New York, 1954).
2. M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares, Metastable behavior of stochastic dynamics: A pathwise approach, *J. Stat. Phys.* **35**:603–634 (1984).
3. D. Chen, J. Feng, and M. P. Qian, The metastability of exponentially perturbed Markov chains, *Sci. China A Eng. Ed.* **39**: 7–28 (1996).

4. D. Chen, J. Feng, and M. P. Qian, The metastable behavior of two-dimensional Ising model, in *Proceedings of the International Conference on Dirichlet Form and Stochastic Processes*, Z. Ma, M. Röckner, and J. Yan, eds. (de Gruyter, Berlin, 1995), pp. 73–86.
5. D. Chen, J. Feng, and M. P. Qian, The metastable behavior of the three dimensional Ising model, Preprint (1993).
6. T. Cox, Coalescing random walks and voter model consensus times on the torus in Z^d, *Ann. Prob.* **17**:1333–1366 (1989).
7. J. D. Deuschel and C. Mazza, L^2 convergence of time nonhomogeneous Markov process: I. Spectral estimates, *Ann. Appl. Prob.* **4**:1012–1056 (1994).
8. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer-Verlag, New York, 1984).
9. A. Galves, E. Olivieri, and M. E. Vares, Metastability for a class of dynamical systems subject to small random perturnations, *Ann. Prob.* **15**:1288–1305 (1987).
10. T. M. Liggett, *Interacting Particle Systems* (Springer-Verlag, New York, 1985).
11. E. J. Neves and R. H. Schonmann, Critical droplets and metastability for a Glauber dynamics at very low temperature, *Commun. Math. Phys.* **137**:209–230 (1991).
12. E. J. Neves and R. H. Schonmann, Behavior of droplets for a class of Glauber dynamics at very low temperature, *Prob. Theory Related Fields* **91**:331–354 (1992).
13. O. Penrose and J. L. Lebowitz, Towards a rigorous molecular theory of metastability, in *Fluctuation Phenomena*, 2nd ed., E. W. Montroll and J. L. Lebowitz, eds. (North-Holland, Amsterdam, 1987).
14. E. Scopolla, Renormalization group for Markov chains and application to metastability, *J. Stat. Phys.* **73**:83–122 (1993).
15. R. H. Schonmann, The pattern of escape from metastability of a stochastic Ising model, *Commun. Math. Phys.* **147**:231–240 (1992).
16. R. H. Schonmann, An approach to characterize metastability and critical droplets in stochastic Ising model, *Ann. Inst. H. Poincaré Phys. Theorique* **55**:591–600 (1991).