HIROSHIMA MATH. J.
30 (2000), 137–165

Application of entropy analysis to discrete-time interacting particle systems on the one-dimensional lattice

Hirotake YAGUCHI

(Received November 2, 1998)
(Revised May 28, 1999)

ABSTRACT. Stationary measures for discrete-time interacting particle systems on the one-dimensional lattice are considered. In our systems infinitely many particles can change their states simultaneously, and the change of each particle state is affected by particles on the surrounding sites. We extensively improve the relative entropy method and make it applicable to such discrete-time particle systems generally. We prove that the stationary measures for Ising models are given by a unique Gibbs state and those for exclusion processes are given by canonical Gibbs states.

1. Introduction

In this paper we aim to establish a general way of analyzing stationary measures for discrete-time interacting particle systems on the one-dimensional lattice. In our systems the particles on sites of $\mathbf{Z}$ change their states at each time according to a given probabilistic rule which satisfies the local equilibrium condition. The number of sites at which changes occur simultaneously is infinite, and the probability of changing a state at each site is affected by the particles in the range of distance R from the focused particle. As such processes we treat discrete-time stochastic Ising models and interactive exclusion processes on the one-dimensional lattice $\mathbf{Z}$. Applying the relative entropy method carefully, we generally discuss a wide class of discrete-time interacting particle systems satisfying the local equilibrium condition. We then determine the structure of stationary measures for the Ising models and the exclusion processes.

Many results have been obtained concerning time evolutions of interacting particle systems (see [6] and the bibliography in [7]). However, in most cases, their time parameters are continuous. We are interested in discrete-time inter-

2000 *Mathematics Subject Classifications*. Primary 60K35; secondary 60J05.
Key words and phrases. Interacting particle system, Stochastic Ising model, Exclusion process, Relative entropy, Markov process, Stationary measure, Discrete time.

acting particle systems *which allow states of particles to change at infinitely many sites simultaneously*. This setting is *extremely different* from the continuous-time cases which allow the change of state at only one site (or finitely many sites) at an instant. This means that, for any interval $[i, j]$ in $\mathbf{Z}$, the particles outside of $[i, j]$ contribute to every term of the equilibrium equation for the particles in $[i, j]$. Hence, differently from the continuous-time cases, we can not isolate the terms which are affected by the particles outside of $[i, j]$. This makes the analysis difficult. Furthermore it unfortunately prevents us from extending the present results to higher dimensional cases.

The well-known tools for the analysis of stationary measures for interacting particle systems are the coupled Markov method ([5, 10]), the relative entropy method ([3, 4, 11]) and recent technologies concerning the entropy of probability density (see, e.g., [8, 9]). The relative entropy method is very useful when one wants to assert that, if a stationary measure for the process is once known, then every stationary measure has the same property as that of the known measure. In this paper we improve and extend the relative entropy method employed in [12], and make it applicable to general finite-range interacting particle systems satisfying the natural conditions (FD1)–(FD5) in the next section. This improved method provides us with parallel treatment of stationary measures for the Ising models and the exclusion processes on the one-dimensional lattice.

In the main body of this paper we restrict our argument to the stochastic Ising models because the notations are rather simple, and in order to avoid confusion with the exclusion processes. We will treat the exclusion processes at the end.

In § 2 we first give some notations and definitions, and introduce discrete-time stochastic Ising models on $\mathbf{Z}$ satisfying the local equilibrium condition together with a simple example. We then state Theorem 1 which determines the structure of stationary measures for the Ising models. The proof of Theorem 1 is given in § 3 by applying a series of lemmas whose proofs are in § 4. In § 5 we give an example of such Ising model. These arguments also work for exclusion processes. An application of our method to discrete-time interactive exclusion processes is discussed in § 6.

2. Definitions and results for stochastic Ising models

Let $\mathscr{X} \equiv \{+1, -1\}^{\mathbf{Z}}$ be the space of spin-configurations on $\mathbf{Z}$. For a given $\eta \equiv (\dots \eta_{-1}\eta_0\eta_1 \dots) \in \mathscr{X}$, we consider that the spin-orientation at site $i \in \mathbf{Z}$ is up [resp. down] if $\eta_i = +1$ [resp. -1]. We endow $\mathscr{X}$ with the product topology of the discrete topology on $\{+1, -1\}$. For $i \leq j$, $i, j \in \mathbf{Z}$, the set of all basic cylinders ${}_i[a_i a_{i+1} \dots a_{j-1} a_j]_j \equiv \{\eta \in \mathscr{X} : \eta_l = a_l, i \leq l \leq j\}$, $a_i \dots a_j \in$

$\{+1,-1\}^{j-i+1}$, is denoted by $\mathscr{C}_{i,j}$. We put $\mathscr{C}=\{\varnothing\}\cup\{\bigcup_{i\le j}\mathscr{C}_{i,j}\}$. As Borel structures on $\mathscr{X}$ we adopt $\mathscr{B}_{i,j}=\sigma(\mathscr{C}_{i,j})$ and $\mathscr{B}=\sigma(\mathscr{C})$, where $\sigma(\mathscr{C}_{i,j})$ [resp. $\sigma(\mathscr{C})$] is the σ-field generated by $\mathscr{C}_{i,j}$ [resp. $\mathscr{C}$]. Given $\boldsymbol{a}\equiv{}_i[a_ia_{i+1}\dots a_{j-1}a_j]_j$ we define $\boldsymbol{b}\langle\boldsymbol{a}\rangle\in\mathscr{C}_{I,J}$, $I\le i,j\le J$, to be an element of $\mathscr{C}_{I,J}$, which is represented such as ${}_I[b_I\dots b_{i-1}a_i\dots a_jb_{j+1}\dots b_J]_J$. A nested notation $\boldsymbol{d}\langle\boldsymbol{c}\langle\boldsymbol{b}\langle\boldsymbol{a}\rangle\rangle\rangle$ is sometimes abbreviated like $\boldsymbol{dcba}$. For an arbitrarily fixed $\boldsymbol{a}\in\mathscr{C}_{i,j}$, which will be clear in each context, we use a notation $\sum^*_{\boldsymbol{b}\langle*\rangle\in\mathscr{C}_{I,J}}$ in order to indicate a summation over all elements $\boldsymbol{b}\langle\boldsymbol{a}\rangle$ in $\mathscr{C}_{I,J}$ with $\boldsymbol{a}$ fixed.

Now let us introduce a random flip of spin-configuration on $\mathbf{Z}$ starting from η. Let $\{P_\eta\equiv P(\eta,\cdot)\mid\eta\in\mathscr{X}\}$ be a set of probability measures on $\mathscr{X}$. We consider that $P(\eta,\boldsymbol{a})$, $\boldsymbol{a}\in\mathscr{B}$, is the probability that a configuration η at time t jumps into a set $\boldsymbol{a}\subset\mathscr{X}$ at time $t+1$. Hence if we assume $P(\cdot,\boldsymbol{a})$ is $\mathscr{B}$-measurable for every $\boldsymbol{a}\in\mathscr{B}$, we have a set of transition probabilities $\{P(\eta,\boldsymbol{a})\}$ and can define a discrete-time Markov process on $\mathscr{X}$ under which each spin on $\mathbf{Z}$ undergoes a random change of spin-orientations. Before describing more precise properties of $P(\eta,\boldsymbol{a})$, it seems to be useful to give here an example of such $\{P(\eta,\boldsymbol{a})\}$.

EXAMPLE. Suppose $0<\alpha_0,\alpha_1,\beta_1<1$ and set $\theta_{+1}=\alpha_0\alpha_1$, $\theta_{-1}=\alpha_0\beta_1$. For $\eta\equiv(\dots\eta_{-1}\eta_0\eta_1\dots)$ let us write $\bar\eta_k=-\eta_k$. Then it is easy to check that the relation

$$P(\eta,{}_i[\bar\eta_i\bar\eta_{i+1}\dots\bar\eta_j]_j)=\alpha_0^{-1}\prod\nolimits_{k=i}^{j+1}\theta_{\eta_{k-1}\eta_k}\qquad\text{and}\tag{2.1}$$

$$P(\eta,A_1\cap A_2\cap\dots\cap A_n)=\prod\nolimits_{s=1}^{n}P(\eta,A_s)\tag{2.2}$$

$$\text{for}\quad A_s={}_{i_s}[\bar\eta_{i_s}\dots\bar\eta_{j_s}]_{j_s}\qquad\text{with }\ i_{s+1}-j_s\ge 2$$

determines a probability measure $P(\eta,\cdot)$ on $\mathscr{X}$ (see (5.1)). This system $\{P(\eta,\cdot)\}$ is a special case of the example given in Section 5 with $R=1$, $J_0=0$ $(\beta_0=\alpha_0)$ and $J_1=\dfrac{1}{2}\log\left(\dfrac{\alpha_1}{\beta_1}\right)$. In this process the change of spin-orientations on the sites i and j with $|i-j|\ge 2$ is mutually independent from (2.2).

Below we describe the precise properties of $P(\eta,\boldsymbol{a})$ which we require for our discrete-time Ising models. We will call the corresponding Markov process (DI) for short.

We introduce a Hamiltonian $\mathscr{H}(\boldsymbol{a})$ as follows: Let R be a positive integer and $\{J_0,J_1,\dots,J_R\}\subset\mathbf{R}$. For $\boldsymbol{a}={}_i[a_i\dots a_j]_j\in\mathscr{C}_{i,j}$ we define

$$\mathscr{H}(\boldsymbol{a})=J_0\sum\nolimits_{x=i}^{j}a_x+\sum_{i\le x<y\le j,\,y-x\le R}\sum\nolimits_x\sum\nolimits_y J_{|x-y|}a_xa_y,$$

which is the energy on the sites $\{i,\dots,j\}$ for $\boldsymbol{a}={}_i[a_i\dots a_j]_j$ w.r.t. the self-potential J_0 and the pair potentials J_r, $r = 1, \dots, R$. We assume that our transition rules P_η, $\eta \in \mathscr{X}$, satisfy the following natural conditions (FD1)–(FD5):

(FD1) $P(\eta, \boldsymbol{a}) \equiv P_\eta(\boldsymbol{a}) > 0$ for every η and $\boldsymbol{a} \in \mathscr{C}_{i,j}$;

(FD2) Given $\boldsymbol{a} \in \mathscr{C}_{i,j}$, $P(\eta, \boldsymbol{a})$ is $\mathscr{B}_{i-R,j+R}$-measurable as a function of η. (Hence we can use a notation $P(\boldsymbol{b}, \boldsymbol{a})$, $\boldsymbol{b} \in \mathscr{C}_{i-R,j+R}$, in the below.)

(FD3) (i) For every $\boldsymbol{a}$, $\tilde{\boldsymbol{a}} \in \mathscr{C}_{i,j}$ and $\boldsymbol{b}\langle * \rangle \in \mathscr{C}_{i-R,j+R}$,

$$\exp\{-\mathscr{H}(\boldsymbol{b}\langle \boldsymbol{a} \rangle)\} P(\boldsymbol{b}\langle \boldsymbol{a} \rangle, \tilde{\boldsymbol{a}}) = \exp\{-\mathscr{H}(\boldsymbol{b}\langle \tilde{\boldsymbol{a}} \rangle)\} P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}} \rangle, \boldsymbol{a}), \tag{2.3}$$

(ii) For every $\boldsymbol{w}$, $\tilde{\boldsymbol{w}} \in \mathscr{C}_{i+R,j-R}$, $\boldsymbol{a}\langle * \rangle, \tilde{\boldsymbol{a}}\langle * \rangle \in \mathscr{C}_{i,j}$ and $\boldsymbol{b}\langle * \rangle \in \mathscr{C}_{i-R,j+R}$, $j - i > 2R$,

$$\begin{aligned} &\exp\{-\mathscr{H}(\boldsymbol{b}\langle \boldsymbol{a}\langle \boldsymbol{w} \rangle\rangle)\} P(\boldsymbol{b}\langle \boldsymbol{a}\langle \boldsymbol{w} \rangle\rangle, \tilde{\boldsymbol{a}}\langle \tilde{\boldsymbol{w}} \rangle) \\ &\qquad = \exp\{-\mathscr{H}(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\langle \boldsymbol{w} \rangle\rangle)\} P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\langle \boldsymbol{w} \rangle\rangle, \boldsymbol{a}\langle \tilde{\boldsymbol{w}} \rangle). \end{aligned} \tag{2.4}$$

(FD4) There exists a positive integer K_1 ($>R$) such that values

$$\frac{P(\boldsymbol{d}\langle \boldsymbol{c}\langle \boldsymbol{b}\langle \boldsymbol{a} \rangle\rangle\rangle, \tilde{\boldsymbol{c}}\langle \boldsymbol{b}\langle \tilde{\boldsymbol{a}} \rangle\rangle)}{P(\boldsymbol{d}\langle \boldsymbol{c}\langle \boldsymbol{b}\langle \boldsymbol{a} \rangle\rangle\rangle, \hat{\boldsymbol{c}}\langle \boldsymbol{b}\langle \tilde{\boldsymbol{a}} \rangle\rangle)},$$

$$\boldsymbol{a}, \tilde{\boldsymbol{a}} \in \mathscr{C}_{i+R,j-R},\ \boldsymbol{b}\langle * \rangle \in \mathscr{C}_{i+R-K_1-L,j-R+K_1+L},$$

$$\boldsymbol{c}\langle * \rangle, \tilde{\boldsymbol{c}}\langle * \rangle, \hat{\boldsymbol{c}}\langle * \rangle \in \mathscr{C}_{i-K_1-L,j+K_1+L},$$

$$\boldsymbol{d}\langle * \rangle \in \mathscr{C}_{i-R-K_1-L,j+R+K_1+L},$$

$$L = 0 \text{ or } R,$$

are independent of $\boldsymbol{a}$ and $\tilde{\boldsymbol{a}}$.

(FD5) There exists a constant $M_1 > 0$ such that

$$\frac{P(\boldsymbol{c}, \boldsymbol{b}\langle \boldsymbol{a} \rangle)}{P(\boldsymbol{c}, \tilde{\boldsymbol{b}}\langle \boldsymbol{a} \rangle)} < M_1$$

for any $\boldsymbol{a} \in \mathscr{C}_{i,j}$, $\boldsymbol{b}\langle * \rangle, \tilde{\boldsymbol{b}}\langle * \rangle \in \mathscr{C}_{i-L,j+L}$, $\boldsymbol{c} \in \mathscr{C}_{i-R-L,j+R+L}$, $0 < L \le K_1 + 3R$.

Let us give some comments on the above conditions. The condition (FD2) means that the change of states on the sites $\{i, \dots, j\}$ is affected by the spins at most on the sites $\{i - R, \dots, j + R\}$. The condition (2.3) in (FD3) is the usual local equilibrium condition, which plays an essential role in our proof. The (FD4) states that changes of spin-orientations near the boundaries of each interval are not affected by spins far from the boundaries. The last (FD5) requires some kind of uniformity of the rate of change of spin-orientations near the boundaries.

We remark that the condition (FD4) is equivalent to the condition

(FD4′) $\dfrac{P(\boldsymbol{d}\langle\boldsymbol{c}\langle\boldsymbol{b}\langle\boldsymbol{a}\rangle\rangle\rangle,\tilde{\boldsymbol{c}}\langle\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle\rangle)}{P(\boldsymbol{c}\langle\boldsymbol{b}\langle\boldsymbol{a}\rangle\rangle,\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle)}$ is independent of $\boldsymbol{a}$ and $\tilde{\boldsymbol{a}}$,

because

$$\frac{P(\boldsymbol{dcba},\tilde{\boldsymbol{c}}\boldsymbol{b}\tilde{\boldsymbol{a}})}{P(\boldsymbol{cba},\boldsymbol{b}\tilde{\boldsymbol{a}})}=\left\{\sum\nolimits^{*}_{\hat{\boldsymbol{c}}\langle*\rangle}\frac{P(\boldsymbol{dcba},\hat{\boldsymbol{c}}\boldsymbol{b}\tilde{\boldsymbol{a}})}{P(\boldsymbol{dcba},\tilde{\boldsymbol{c}}\boldsymbol{b}\tilde{\boldsymbol{a}})}\right\}^{-1}\qquad\text{and}$$

$$\frac{P(\boldsymbol{dcba},\tilde{\boldsymbol{c}}\boldsymbol{b}\tilde{\boldsymbol{a}})}{P(\boldsymbol{dcba},\hat{\boldsymbol{c}}\boldsymbol{b}\tilde{\boldsymbol{a}})}=\frac{P(\boldsymbol{dcba},\tilde{\boldsymbol{c}}\boldsymbol{b}\tilde{\boldsymbol{a}})/P(\boldsymbol{cba},\boldsymbol{b}\tilde{\boldsymbol{a}})}{P(\boldsymbol{dcba},\hat{\boldsymbol{c}}\boldsymbol{b}\tilde{\boldsymbol{a}})/P(\boldsymbol{cba},\boldsymbol{b}\tilde{\boldsymbol{a}})}.$$

We will use (2.4) together with (FD4) and (FD5) for proving the necessity part of Theorem 1. We also remark that if we put $\boldsymbol{w}=\tilde{\boldsymbol{w}}=\varnothing$ formally in (2.4) then it reduces to (2.3) (see Concluding Remark 2). We will call the condition (2.4) the "*F*inite-range *D*ynamic local equilibrium".

A probability measure ν on $\mathscr{X}$ is called a *Gibbs state* associated with the self-potential J_0 and the pair potentials J_r, $r=1,\dots,R$, if its conditional probability $\nu\{\boldsymbol{a}|\mathscr{B}^c_{i,j}\}(\eta)$ of $\boldsymbol{a}\in\mathscr{C}_{i,j}$ given $\mathscr{B}^c_{i,j}\equiv\sigma\{\mathscr{C}_{I,J}|I\le J<i \text{ or } j<I\le J\}$ is equal to

$$\Xi_{i,j}(\eta)^{-1}\exp[-\mathscr{H}({}_{i-R}[\eta_{i-R}\cdots\eta_{i-1}a_i\dots a_j\eta_{j+1}\cdots\eta_{j+R}]_{j+R})],$$

where $\Xi_{i,j}(\eta)$ is the normalizing factor which depends on i, j and η (see, e.g., [2, 6]). It is obvious that the above definition is equivalent to the following one: ν is called a Gibbs state if it satisfies

$$\nu(\boldsymbol{b}\langle\boldsymbol{a}\rangle)\exp\{-\mathscr{H}(\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle)\}=\nu(\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle)\exp\{-\mathscr{H}(\boldsymbol{b}\langle\boldsymbol{a}\rangle)\}\tag{2.5}$$

$$\text{for every}\quad \boldsymbol{a},\tilde{\boldsymbol{a}}\in\mathscr{C}_{i,j}\quad\text{and}\quad \boldsymbol{b}\langle*\rangle\in\mathscr{C}_{i-R,j+R}.$$

The set of Gibbs states is written by $\mathscr{G}$. In the one-dimensional case it is known that $\mathscr{G}\ne\varnothing$, and, moreover, $\sharp\mathscr{G}=1$. We remark that if P_η satisfies (2.3), then the condition $\nu\in\mathscr{G}$, namely (2.5), is equivalent to the following equation:

$$\nu(\boldsymbol{b}\langle\boldsymbol{a}\rangle)P(\boldsymbol{b}\langle\boldsymbol{a}\rangle,\tilde{\boldsymbol{a}})=\nu(\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle)P(\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle,\boldsymbol{a})\tag{2.6}$$

$$\text{for every}\quad \boldsymbol{a},\tilde{\boldsymbol{a}}\in\mathscr{C}_{i,j}\quad\text{and}\quad \boldsymbol{b}\langle*\rangle\in\mathscr{C}_{i-R,j+R}.$$

A probability measure ν on $\mathscr{X}$ is said to be *stationary* for the Markov process defined by the transition probabilities $P(\eta,\cdot)$, $\eta\in\mathscr{X}$, if it satisfies

$$\int_{\mathscr{X}}d\nu(\eta)f(\eta)=\int_{\mathscr{X}}d\nu(\eta)\int_{\mathscr{X}}P(\eta,d\xi)f(\xi)$$

for all bounded $\mathscr{B}$-measurable functions f. Let $\mathscr{I}$ be the set of stationary measures for our process (DI).

Now we can state our theorem as follows.

THEOREM 1. *Assume the conditions* (FD1)–(FD5). *Then* $\mathscr{I}=\mathscr{G}$, *that is, a probability measure* ν *on* $\mathscr{X}$ *is stationary for the discrete-time stochastic Ising model (DI) if and only if it is a Gibbs state associated with the potentials* J_r, $r=0,1,\dots,R$.

Thus stationary measures for (DI) is unique and coincides with a unique Gibbs state with potentials J_r. As a corollary we have

COROLLARY 1. *The stationary measure for the discrete-time stochastic Ising model (DI) is reversible, that is,*

$$\int_{\mathscr{X}}\nu(d\eta)f(\eta)\int_{\mathscr{X}}P(\eta,d\xi)g(\xi)=\int_{\mathscr{X}}\nu(d\eta)g(\eta)\int_{\mathscr{X}}P(\eta,d\xi)f(\xi) \tag{2.7}$$

for all continuous functions f *and* g.

The analogous results also hold for discrete-time exclusion processes. We describe them, Theorem 2 and its corollaries, in §6.

3. Proof of Theorem 1

We divide the proof into two parts, that is, a sufficiency part and a necessity part. Most of this paper will be devoted to the proof of the necessity part.

PROOF OF THE SUFFICIENCY PART. The sufficiency is almost obvious. Indeed for $\boldsymbol{a}\in\mathscr{C}_{i,j}$, by (2.6),

$$\begin{aligned}\int_{\mathscr{X}}d\nu(\eta)P(\eta,\boldsymbol{a})&=\sum\nolimits_{\boldsymbol{b}\in\mathscr{C}_{i-R,j+R}}\nu(\boldsymbol{b})P(\boldsymbol{b},\boldsymbol{a})\\&=\sum\nolimits^{*}_{\boldsymbol{b}\langle *\rangle\in\mathscr{C}_{i-R,j+R}}\sum\nolimits_{\tilde{\boldsymbol{a}}\in\mathscr{C}_{i,j}}\nu(\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle)P(\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\rangle,\boldsymbol{a})\\&=\sum\nolimits^{*}_{\boldsymbol{b}\langle *\rangle}\sum\nolimits_{\tilde{\boldsymbol{a}}}\nu(\boldsymbol{b}\langle\boldsymbol{a}\rangle)P(\boldsymbol{b}\langle\boldsymbol{a}\rangle,\tilde{\boldsymbol{a}})\\&=\sum\nolimits^{*}_{\boldsymbol{b}\langle *\rangle}\nu(\boldsymbol{b}\langle\boldsymbol{a}\rangle)=\nu(\boldsymbol{a}).\quad\square\end{aligned}\tag{3.1}$$

PROOF OF COROLLARY 1. For the proof it is sufficient to check (2.7) for the case $f=\chi_{\boldsymbol{a}}$ and $g=\chi_{\boldsymbol{b}}$, $\boldsymbol{a}$, $\boldsymbol{b}\in\mathscr{C}$, under the condition (2.6). But this is easy. In fact, if $\boldsymbol{a}\in\mathscr{C}_{i,j}$, $\boldsymbol{b}\in\mathscr{C}_{I,J}$, $i<I<j<J$, then

$$\begin{aligned}
&\int_{\mathscr{X}} \nu(d\eta)\chi_{\boldsymbol{a}}(\eta) \int_{\mathscr{X}} P(\eta, d\eta')\chi_{\boldsymbol{b}}(\eta') \\
&\quad = \int_{\boldsymbol{a}} \nu(d\eta)P(\eta, \boldsymbol{b}) \\
&\quad = \sum\nolimits^*_{\boldsymbol{u}\langle * \rangle \in \mathscr{C}_{i,J}} \sum\nolimits^*_{\boldsymbol{v}\langle * \rangle \in \mathscr{C}_{i-R,J+R}} \nu(\boldsymbol{v}\langle \boldsymbol{u}\langle \boldsymbol{a}\rangle\rangle)P(\boldsymbol{v}\langle \boldsymbol{u}\langle \boldsymbol{a}\rangle\rangle, \boldsymbol{b}) \\
&\quad = \sum\nolimits^*_{\boldsymbol{u}\langle * \rangle} \sum\nolimits^*_{\boldsymbol{v}\langle * \rangle} \sum\nolimits^*_{\boldsymbol{w}\langle * \rangle \in \mathscr{C}_{i,J}} \nu(\boldsymbol{v}\langle \boldsymbol{u}\langle \boldsymbol{a}\rangle\rangle)P(\boldsymbol{v}\langle \boldsymbol{u}\langle \boldsymbol{a}\rangle\rangle, \boldsymbol{w}\langle \boldsymbol{b}\rangle) \\
&\quad = \sum\nolimits^*_{\boldsymbol{w}\langle * \rangle} \sum\nolimits^*_{\boldsymbol{v}\langle * \rangle} \sum\nolimits^*_{\boldsymbol{u}\langle * \rangle} \nu(\boldsymbol{v}\langle \boldsymbol{w}\langle \boldsymbol{b}\rangle\rangle)P(\boldsymbol{v}\langle \boldsymbol{w}\langle \boldsymbol{b}\rangle\rangle, \boldsymbol{u}\langle \boldsymbol{a}\rangle) \\
&\quad = \int_{\boldsymbol{b}} d\nu(\eta)P(\eta, \boldsymbol{a}) = \int_{\mathscr{X}} \nu(d\eta)\chi_{\boldsymbol{b}}(\eta) \int_{\mathscr{X}} P(\eta, d\eta')\chi_{\boldsymbol{a}}(\eta'). \quad \square
\end{aligned}$$

In the rest of the section we prove the necessity part. In the following discussion by μ we represent the unique Gibbs state with potentials J_r, $r = 0, \dots, R$, and by ν an arbitrary probability measure on $\mathscr{X}$. We remark that for μ it holds that

$$\mu(\boldsymbol{b}\langle \boldsymbol{a}\rangle)P(\boldsymbol{b}\langle \boldsymbol{a}\rangle, \tilde{\boldsymbol{a}}) = \mu(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\rangle)P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\rangle, \boldsymbol{a}), \quad \boldsymbol{a}, \tilde{\boldsymbol{a}} \in \mathscr{C}_{i,j}, \ \boldsymbol{b}\langle \boldsymbol{a}\rangle \in \mathscr{C}_{i-R,j+R} \tag{3.2}$$

from (2.6). We use the notation $\Psi(x)$ for the function $x \log x$, $x \geq 0$, with $\Psi(0) = 0$ as usual. The *relative entropy* of ν with respect to μ on $\{-N, -N+1, \dots, N-1, N\}$, $N \in \mathbf{N}$, is defined by

$$H_N(\nu) = \sum\nolimits_{\boldsymbol{a} \in \mathscr{C}_{-N,N}} \mu(\boldsymbol{a}) \Psi\left(\frac{\nu(\boldsymbol{a})}{\mu(\boldsymbol{a})}\right). \tag{3.3}$$

Suppose that the initial distribution of (DI) at $t = 0$ is ν. Then the distribution of (DI) at $t = 1$ is given by $\tilde{\nu}(\cdot) = \int \nu(d\eta)P(\eta, \cdot)$, and hence by (3.1) for $\boldsymbol{a} \in \mathscr{C}_{-N,N}$

$$\tilde{\nu}(\boldsymbol{a}) = \sum\nolimits_{\boldsymbol{w} \in \mathscr{C}_{-N+R,N-R}} X(\boldsymbol{a}, \boldsymbol{w}), \tag{3.4}$$

where

$$X(\boldsymbol{a}, \boldsymbol{w}) = \sum\nolimits^*_{\tilde{\boldsymbol{a}}\langle * \rangle \in \mathscr{C}_{-N,N}} \sum\nolimits^*_{\boldsymbol{b}\langle * \rangle \in \mathscr{C}_{-N-R,N+R}} \nu(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\langle \boldsymbol{w}\rangle\rangle)P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\langle \boldsymbol{w}\rangle\rangle, \boldsymbol{a}). \tag{3.5}$$

Therefore concerning the relative entropy of $\tilde{\nu}$ we have an inequality

$$\begin{aligned}
H_N(\tilde{\nu}) &= \sum\nolimits_{\boldsymbol{a}} \mu(\boldsymbol{a}) \Psi\left(\frac{\sum_{\boldsymbol{w}} X(\boldsymbol{a}, \boldsymbol{w})}{\mu(\boldsymbol{a})}\right) \\
&\leq \sum\nolimits_{\boldsymbol{a}} \mu(\boldsymbol{a}) \sum\nolimits_{\boldsymbol{w}} P(\boldsymbol{a}, \boldsymbol{w}) \Psi\left(\frac{X(\boldsymbol{a}, \boldsymbol{w})}{P(\boldsymbol{a}, \boldsymbol{w})\mu(\boldsymbol{a})}\right) \\
&\equiv I_N(\nu)
\end{aligned}$$

by the convexity of Ψ. The difference is denoted by Λ_N:

(3.6) $$\Lambda_N(v) = I_N(v) - H_N(\tilde{v}) \geq 0, \qquad N = R+1, R+2, \ldots.$$

The key to the necessity part is to show that $\Lambda_N(v) = 0$ if v is stationary for (DI). We divide the argument into a series of lemmas, whose proofs are given in the next section. The first step is to show that Λ_N's are monotonically increasing w.r.t. N:

LEMMA 1. *Suppose $v(\boldsymbol{a}) > 0$ for every nonempty $\boldsymbol{a} \in \mathscr{C}$. Then we have*

$$\Lambda_N(v) \geq \Lambda_{N-1}(v).$$

The next step is to state that if v is stationary, then $\Lambda_N(v)$'s are bounded above.

LEMMA 2. *Suppose v is stationary for (DI) and $v(\boldsymbol{a}) > 0$ for all nonempty $\boldsymbol{a} \in \mathscr{C}$. Then there exists a positive constant c such that $0 \leq \Lambda_N(v) \leq c$ for all $N(>R)$.*

From Lemmas 1 and 2 we can show that

LEMMA 3. *Suppose v is stationary for (DI) and $v(\boldsymbol{a}) > 0$ for all nonempty $\boldsymbol{a} \in \mathscr{C}$. Define*

$$g(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle) = \log\left(\sum\nolimits^*_{\boldsymbol{y}\langle *\rangle} \frac{X(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)}{\sum^*_{\hat{\boldsymbol{y}}\langle *\rangle} X(\boldsymbol{c}, \hat{\boldsymbol{y}}\langle \boldsymbol{v}\rangle)} \cdot \frac{X(\boldsymbol{c}, \check{\boldsymbol{y}}, \langle \boldsymbol{v}\rangle)}{X(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)} \cdot \frac{P(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)}{P(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle)}\right)$$
$$- \sum\nolimits^*_{\boldsymbol{y}\langle *\rangle} \frac{X(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)}{\sum^*_{\hat{\boldsymbol{y}}\langle *\rangle} X(\boldsymbol{c}, \hat{\boldsymbol{y}}\langle \boldsymbol{v}\rangle)} \log\left(\frac{X(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle)}{X(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)} \cdot \frac{P(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)}{P(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle)}\right)$$

where $\boldsymbol{c} \in \mathscr{C}_{-N-2R-K_1, N+2R+K_1}$, $\boldsymbol{v} \in \mathscr{C}_{-N+2R, N-2R}$ and $\check{\boldsymbol{y}}\langle v\rangle \in \mathscr{C}_{-N-R-K_1, N+R+K_1}$. Let $\delta > 0$ and $\gamma > 0$. Then for an arbitrarily fixed $\check{\boldsymbol{y}}\langle\cdot\rangle$, if N is sufficiently large, it holds that

(3.7) $$0 \leq \sum_{\substack{\boldsymbol{c}\ \ \boldsymbol{v} \\ g(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle) \geq \delta}}\!\!\sum \left\{\sum\nolimits^*_{\boldsymbol{y}\langle *\rangle} X(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)\right\} < \gamma.$$

Here $\sum_{\boldsymbol{c}} \sum_{\boldsymbol{v}}$ with $g(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle) \geq \delta$ is the summation over $\boldsymbol{c}$ and $\boldsymbol{v}$ satisfying $g(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle) \geq \delta$.

This lemma states that if v is stationary then the sum of $X(\boldsymbol{c}, \boldsymbol{y}\langle \boldsymbol{v}\rangle)$'s over $\boldsymbol{c}$, $\boldsymbol{y}$ and $\boldsymbol{v}$ which satisfy $g(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle) \geq \delta$ becomes very small as N goes to infinity. (Note that the function g is nonnegative.)

The following lemma reflects the strict concavity of the function log.

LEMMA 4. *Suppose $v(\boldsymbol{a}) > 0$ for all nonempty $\boldsymbol{a} \in \mathscr{C}$. Define $g(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v}\rangle)$ as in Lemma 3. Then for each sufficiently small $\gamma > 0$ we can choose $\delta > 0$ such that*

(3.8) $$\text{if } g(\boldsymbol{c},\check{\boldsymbol{y}}\langle\boldsymbol{v}\rangle)<\delta \text{ then } \left|1-\frac{X(\boldsymbol{c},\check{\boldsymbol{y}}\langle\boldsymbol{v}\rangle)}{X(\boldsymbol{c},\boldsymbol{y}\langle\boldsymbol{v}\rangle)}\cdot\frac{P(\boldsymbol{c},\boldsymbol{y}\langle\boldsymbol{v}\rangle)}{P(\boldsymbol{c},\check{\boldsymbol{y}}\langle\boldsymbol{v}\rangle)}\right|<\gamma$$

for every $\boldsymbol{y}\langle\boldsymbol{v}\rangle$ *with* $\boldsymbol{v}$ *fixed. Further we can take* δ *to be independent of* $\boldsymbol{c}$, $\boldsymbol{v}$ *and* N.

Using Lemmas 3 and 4 we can prove the next essential lemma:

LEMMA 5. *Suppose* ν *is stationary for (DI) and* $\nu(\boldsymbol{a})>0$ *for all nonempty* $\boldsymbol{a}\in\mathscr{C}$. *Then* $\lim_{N\to\infty}\Lambda_N(\nu)=0$, *and hence* $\Lambda_N(\nu)=0$ *for all* $N(>R)$.

We finally verify the assumption in the previous lemmas.

LEMMA 6. *If* ν *is stationary, then* $\nu(\boldsymbol{a})>0$ *for every nonempty* $\boldsymbol{a}\in\mathscr{C}$.

PROOF OF THE NECESSITY PART OF THEOREM 1. By Lemma 6 we have $\nu(\boldsymbol{a})>0$ for $\boldsymbol{a}\neq\varnothing$, and hence $\Lambda_N=0$ for all N by Lemma 5. Since Ψ is strictly convex, this means that the values $\dfrac{X(\boldsymbol{a},\boldsymbol{w})}{P(\boldsymbol{a},\boldsymbol{w})\mu(\boldsymbol{a})}$ are the same for all $\boldsymbol{w}$, that is, they are independent of $\boldsymbol{w}\in\mathscr{C}_{-N+R,N-R}$. From this and (3.4) and the stationarity of ν we see that $\nu(\boldsymbol{a})=\dfrac{X(\boldsymbol{a},\boldsymbol{w})}{P(\boldsymbol{a},\boldsymbol{w})}$ for any $\boldsymbol{w}$. Let $\boldsymbol{u},\hat{\boldsymbol{u}}\in\mathscr{C}_{-N+2R+K_1,N-2R-K_1}$, $\boldsymbol{v}\langle*\rangle\in\mathscr{C}_{-N+2R,N-2R}$, $\boldsymbol{w}\langle*\rangle\in\mathscr{C}_{-N+R,N-R}$ and $\boldsymbol{a}\langle*\rangle\in\mathscr{C}_{-N,N}$. Then

$$\begin{aligned}
(3.9)\quad \nu(\boldsymbol{awvu}) &= \frac{X(\boldsymbol{awvu},\boldsymbol{wv\hat{u}})}{P(\boldsymbol{awvu},\boldsymbol{wv\hat{u}})}\\
&= \frac{P(\boldsymbol{awv\hat{u}},\boldsymbol{wvu})}{P(\boldsymbol{awvu},\boldsymbol{wv\hat{u}})}\sum\nolimits^*_{\tilde{\boldsymbol{b}}\langle*\rangle\in\mathscr{C}_{-N-R,N+R}}\sum\nolimits^*_{\tilde{\boldsymbol{a}}\langle*\rangle}\frac{P(\tilde{\boldsymbol{b}}\tilde{\boldsymbol{a}}\boldsymbol{wv\hat{u}},\boldsymbol{awvu})}{P(\boldsymbol{awv\hat{u}},\boldsymbol{wvu})}\nu(\tilde{\boldsymbol{b}}\tilde{\boldsymbol{a}}\boldsymbol{wv\hat{u}})
\end{aligned}$$

by (3.5). Especially if we take $\boldsymbol{u}=\hat{\boldsymbol{u}}$ in the above, we have

$$\nu(\boldsymbol{awv\hat{u}})=\sum\nolimits^*_{\tilde{\boldsymbol{b}}\langle*\rangle}\sum\nolimits^*_{\tilde{\boldsymbol{a}}\langle*\rangle}\frac{P(\tilde{\boldsymbol{b}}\tilde{\boldsymbol{a}}\boldsymbol{wv\hat{u}},\boldsymbol{awv\hat{u}})}{P(\boldsymbol{awv\hat{u}},\boldsymbol{wv\hat{u}})}\nu(\tilde{\boldsymbol{b}}\tilde{\boldsymbol{a}}\boldsymbol{wv\hat{u}}).$$

Therefore applying the equalities

$$\frac{P(\boldsymbol{awv\hat{u}},\boldsymbol{wvu})}{P(\boldsymbol{awvu},\boldsymbol{wv\hat{u}})}=\frac{\exp\{-\mathscr{H}(\boldsymbol{awvu})\}}{\exp\{-\mathscr{H}(\boldsymbol{awv\hat{u}})\}}=\frac{\exp\{-\mathscr{H}(\boldsymbol{vu})\}}{\exp\{-\mathscr{H}(\boldsymbol{v\hat{u}})\}}\qquad\text{and}$$

$$\begin{aligned}
\frac{P(\tilde{\boldsymbol{b}}\tilde{\boldsymbol{a}}\boldsymbol{wv}\langle\hat{\boldsymbol{u}}\rangle,\boldsymbol{awv}\langle\boldsymbol{u}\rangle)}{P(\boldsymbol{awv}\langle\hat{\boldsymbol{u}}\rangle,\boldsymbol{wv}\langle\boldsymbol{u}\rangle)} &= P(\tilde{\boldsymbol{b}}\tilde{\boldsymbol{a}}\boldsymbol{wv}\langle\hat{\boldsymbol{u}}\rangle,\boldsymbol{awv}\langle\boldsymbol{u}\rangle)/P(\tilde{\boldsymbol{a}}\boldsymbol{wv}\langle\hat{\boldsymbol{u}}\rangle,\boldsymbol{wv}\langle\boldsymbol{u}\rangle)\\
&\qquad\times\frac{P(\tilde{\boldsymbol{a}}\boldsymbol{wv}\langle\hat{\boldsymbol{u}}\rangle,\boldsymbol{wv}\langle\boldsymbol{u}\rangle)/P(\boldsymbol{wv\hat{u}},\boldsymbol{v}\langle\boldsymbol{u}\rangle)}{P(\boldsymbol{awv}\langle\hat{\boldsymbol{u}}\rangle,\boldsymbol{wv}\langle\boldsymbol{u}\rangle)/P(\boldsymbol{wv\hat{u}},\boldsymbol{v}\langle\boldsymbol{u}\rangle)}\\
&= \frac{P(\tilde{\boldsymbol{b}}\tilde{\boldsymbol{a}}\boldsymbol{wv\hat{u}},\boldsymbol{awv}\langle\hat{\boldsymbol{u}}\rangle)}{P(\boldsymbol{awv\hat{u}},\boldsymbol{wv}\langle\hat{\boldsymbol{u}}\rangle)},
\end{aligned}$$

which follow from (2.3) and (FD4′), to (3.9), we obtain

$$\nu(\boldsymbol{awvu}) = \frac{\exp\{-\mathscr{H}(\boldsymbol{vu})\}}{\exp\{-\mathscr{H}(\boldsymbol{v\hat{u}})\}}\nu(\boldsymbol{awv\hat{u}}).$$

This yields $\nu(\boldsymbol{vu})\exp\{-\mathscr{H}(\boldsymbol{v\hat{u}})\} = \nu(\boldsymbol{v\hat{u}})\exp\{-\mathscr{H}(\boldsymbol{vu})\}$. Hence (2.5) is proved. □

4. Proofs of lemmas

In this section we give the proofs of lemmas which are used in the preceding section. We frequently use notations such as $\mathscr{C}_N$, $\boldsymbol{aw}$, $\sum_{\boldsymbol{a}}^{*}$, $\sum_{\boldsymbol{w},\tilde{\boldsymbol{w}}}^{*},\dots$, instead of $\mathscr{C}_{-N,N}$, $\boldsymbol{a}\langle\boldsymbol{w}\rangle$, $\sum_{\boldsymbol{a}\langle *\rangle}^{*}$, $\sum_{\boldsymbol{w}\langle *\rangle}^{*}\sum_{\tilde{\boldsymbol{w}}\langle *\rangle}^{*},\dots$, respectively for brevity.

PROOF OF LEMMA 1. From (3.6) we have

$$\Lambda_N(\nu) = \sum_{\boldsymbol{a}}\sum_{\boldsymbol{v}\in\mathscr{C}_{N-R-1}}\sum_{\boldsymbol{w}\in\mathscr{C}_{N-R}}^{*} X(\boldsymbol{a},\boldsymbol{wv})\log\frac{X(\boldsymbol{a},\boldsymbol{wv})}{P(\boldsymbol{a},\boldsymbol{wv})}$$
$$-\sum_{\boldsymbol{a}}\Big(\sum_{\boldsymbol{w}} X(\boldsymbol{a},\boldsymbol{w})\Big)\log\Big(\sum_{\tilde{\boldsymbol{w}}} X(\boldsymbol{a},\tilde{\boldsymbol{w}})\Big).$$

Then by the concavity

(4.1) $$a\cdot\log(b/a) + c\cdot\log(d/c)$$
$$\le (a+c)\log[(b+d)/(a+c)] \quad (a,b,c,d>0)$$

of log-function, it holds that

(4.2) $$\sum_{\boldsymbol{a}}\sum_{\boldsymbol{v}}\sum_{\boldsymbol{w}}^{*} X(\boldsymbol{a},\boldsymbol{wv})\log\frac{X(\boldsymbol{a},\boldsymbol{wv})}{P(\boldsymbol{a},\boldsymbol{wv})}$$
$$\ge \sum_{\boldsymbol{a}}\sum_{\boldsymbol{v}}\Big(\sum_{\boldsymbol{w}}^{*} X(\boldsymbol{a},\boldsymbol{wv})\Big)\log\frac{\sum_{\check{\boldsymbol{w}}}^{*} X(\boldsymbol{a},\check{\boldsymbol{w}}\boldsymbol{v})}{\sum_{\check{\boldsymbol{w}}}^{*} P(\boldsymbol{a},\check{\boldsymbol{w}}\boldsymbol{v})},$$

and hence

(4.3) $$\Lambda_N(\nu) \ge \sum_{\boldsymbol{a}}\sum_{\boldsymbol{v}}\Big(\sum_{\boldsymbol{w}}^{*} X(\boldsymbol{a},\boldsymbol{wv})\Big)\log\frac{\sum_{\check{\boldsymbol{w}}}^{*} X(\boldsymbol{a},\check{\boldsymbol{w}}\boldsymbol{v})}{\sum_{\tilde{\boldsymbol{w}}} X(\boldsymbol{a},\tilde{\boldsymbol{w}})}$$
$$-\sum_{\boldsymbol{a}}\sum_{\boldsymbol{v}}\Big(\sum_{\boldsymbol{w}}^{*} X(\boldsymbol{a},\boldsymbol{wv})\Big)\log P({}_{-N+1}[\boldsymbol{a}]_{N-1},\boldsymbol{v}),$$

where ${}_{-N+1}[\boldsymbol{a}]_{N-1} = {}_{-N+1}[a_{-N+1}\dots a_{N-1}]_{N-1}$. Rewriting the sum $\sum_{\boldsymbol{a}}$ into the form $\sum_{\boldsymbol{c}\in\mathscr{C}_{N-1}}\sum_{\boldsymbol{a}\in\mathscr{C}_N}^{*}$ and applying (4.1) again to $\sum_{\boldsymbol{a}}^{*}$, we finally get

$$
\begin{aligned}
\Lambda_N(\nu) &\geq \sum_c \sum_v \Big(\sum_a^* \sum_w^* X(ac, wv)\Big) \log \frac{\sum_{\tilde{a}}^* \sum_{\tilde{w}}^* X(\tilde{a}c, \tilde{w}v)}{\sum_{\tilde{a}}^* \sum_{\tilde{w}} X(\tilde{a}c, \tilde{w})} \\
&\quad - \sum_c \sum_v \Big(\sum_a^* \sum_w^* X(ac, wv)\Big) \log P(c, v) \\
&= \sum_c \sum_v X(c, v) \log \frac{X(c, v)}{\sum_{\tilde{v}} X(c, \tilde{v})} - \sum_c \sum_v X(c, v) \log P(c, v) \\
&\equiv \Lambda_{N-1}(\nu),
\end{aligned}
$$

here we have used $\sum_a^* \sum_w^* X(ac, wv) = X(c, v)$ and $\sum_{\tilde{a}}^* \sum_{\tilde{w}} X(\tilde{a}c, \tilde{w}) = \sum_{\tilde{v}} \sum_{\tilde{a}}^* \sum_{\tilde{w}}^* X(\tilde{a}c, \tilde{w}\tilde{v}) = \sum_{\tilde{v}} X(c, \tilde{v})$. □

Proof of Lemma 2. From the stationarity of ν and $\sum_{\tilde{a}} P(b, \tilde{a}) = 1$ we have

$$
\begin{aligned}
H_N(\tilde{\nu}) &= H_N(\nu) \\
&= \sum_a \sum_{b \in \mathscr{C}_{N+R}}^* \sum_{\tilde{a} \in \mathscr{C}_N} \nu(ba) P(ba, \tilde{a}) \log \frac{\nu(a)}{\mu(a)} \\
&= \sum_a \sum_{\tilde{a}} \sum_b^* \nu(b\tilde{a}) P(b\tilde{a}, a) \log \frac{\nu(\tilde{a})}{\mu(\tilde{a})}.
\end{aligned}
$$

Since

$$
I_N(\nu) = \sum_a \sum_w \sum_{\tilde{a}}^* \sum_b^* \nu(b\tilde{a}w) P(b\tilde{a}w, a) \log \frac{X(a, w)}{P(a, w)\mu(a)}
$$

by (3.5), (3.6) is reduced to

$$
\begin{aligned}
\Lambda_N(\nu) &= -\sum_a \sum_w \sum_{\tilde{a}}^* \sum_b^* \nu(b\tilde{a}w) P(b\tilde{a}w, a) \log \left(\frac{P(a, w)\mu(a)}{X(a, w)} \cdot \frac{\nu(\tilde{a}w)}{\mu(\tilde{a}w)} \right) \\
&= -\sum_a \sum_w \sum_{\tilde{a}}^* \sum_b^* \frac{\nu(b\tilde{a}w)}{\nu(\tilde{a}w)} \times \frac{\mu(\tilde{a}w)}{\mu(b\tilde{a}w)} \times \frac{P(ba, \tilde{a}w)}{P(a, w)} \\
&\qquad \times \frac{\mu(ba)}{\mu(a)} \times X(a, w) \times \Psi \left(\frac{P(a, w)\mu(a)}{X(a, w)} \cdot \frac{\nu(\tilde{a}w)}{\mu(\tilde{a}w)} \right) \\
&\qquad\qquad [\text{by } P(b\tilde{a}, a) = P(ba, \tilde{a})(\mu(ba)/\mu(b\tilde{a})) \text{ from } (3.2)] \\
&\equiv \sum_a \sum_w \sum_{\tilde{a}}^* \sum_b^* T_1 \times T_2 \times T_3 \times T_4 \times X(a, w) \times (-T_5).
\end{aligned}
$$

Then the lemma follows from the facts (i) $0 < T_1$, T_3, $T_4 \leq 1$ (because each measurable set in the numerators is a subset of the corresponding denominator's), (ii) T_2 is bounded above uniformly (because μ is a Gibbs state), (iii)

$-T_5 \le e^{-1}$ (by $-\Psi(u) \le e^{-1}$), (iv) $\sum_{\tilde{a}}^{*} \sum_{\tilde{b}}^{*} 1 = 2^{4R}$, and (v) $\sum_{a} \sum_{w} X(a, w) = \sum_{a} v(a) = 1$ (since v is stationary). $\square$

PROOF OF LEMMA 3. Just as (4.2) and the following in the proof of Lemma 1 we have

$$\begin{aligned}
\Lambda_{N+2R+K_1} &= \sum_{c} \sum_{v} \sum_{y}^{*} X(c, yv) \log \frac{X(c, yv)}{P(c, yv)} \\
&\quad - \sum_{c} \Big(\sum_{y} X(c, y) \Big) \log \Big(\sum_{\tilde{y}} X(c, \tilde{y}) \Big) \\
&\ge \sum_{c} \sum_{v} \Big(\sum_{y}^{*} X(c, yv) \Big) \log \frac{\sum_{\tilde{y}}^{*} X(c, \tilde{y}v)}{\sum_{\tilde{y}}^{*} P(c, \tilde{y}v)} \\
&\quad - \sum_{c} \Big(\sum_{y} X(c, y) \Big) \log \Big(\sum_{\tilde{y}} X(c, \tilde{y}) \Big) \\
&\ge \Lambda_{N+R}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\Lambda_{N+2R+K_1} - \Lambda_{N+R} \\
&\quad \ge \sum_{c} \sum_{v} \sum_{y}^{*} X(c, yv) \log \frac{X(c, yv)}{P(c, yv)} \\
&\qquad - \sum_{c} \sum_{v} \Big(\sum_{y}^{*} X(c, yv) \Big) \log \frac{\sum_{\tilde{y}}^{*} X(c, \tilde{y}v)}{\sum_{\tilde{y}}^{*} P(c, \tilde{y}v)} \\
&\quad = -\sum_{c} \sum_{v} \sum_{y}^{*} X(c, yv) \log \frac{P(c, yv)/X(c, yv)}{P(c, \check{y}v)/X(c, \check{y}v)} \\
&\qquad + \sum_{c} \sum_{v} \Big(\sum_{y}^{*} X(c, yv) \Big) \log \Big(\sum_{\tilde{y}}^{*} X(c, \tilde{y}v) \frac{P(c, \tilde{y}v)/X(c, \tilde{y}v)}{P(c, \check{y}v)/X(c, \check{y}v)} \Big) \\
&\qquad - \sum_{c} \sum_{v} \Big(\sum_{y}^{*} X(c, yv) \Big) \log \Big(\sum_{\tilde{y}}^{*} X(c, \tilde{y}v) \Big) \\
&\quad = \sum_{c} \sum_{v} \Big(\sum_{\hat{y}}^{*} X(c, \hat{y}v) \Big) g(c, \check{y}v) \\
&\quad \ge \sum_{c} \sum_{\substack{v \\ g(c, \check{y}v) \ge \delta}} \Big\{ \sum_{\hat{y}}^{*} X(c, \hat{y}v) g(c, \check{y}v) \Big\} \\
&\quad \ge \sum_{c} \sum_{\substack{v \\ g(c, \check{y}v) \ge \delta}} \Big\{ \sum_{y}^{*} X(c, yv) \delta \Big\}
\end{aligned}$$

for arbitrarily fixed $\check{y}\langle \cdot \rangle$. On the other hand, from Lemmas 1 and 2 it holds that

$$0 \le \Lambda_{N+2R+K_1} - \Lambda_{N+R} < \gamma\delta$$

for all sufficiently large N. Hence the lemma follows. $\square$

Let G be a constant defined by $G = 2^{2(3R+K_1)}$.

Proof of Lemma 4. *Step 1.* Let $\{\boldsymbol{y}_1\boldsymbol{v}, \boldsymbol{y}_2\boldsymbol{v}, \dots, \boldsymbol{y}_G\boldsymbol{v}\}$ be the set of all $\boldsymbol{y}\langle\boldsymbol{v}\rangle$ for a given $\boldsymbol{v}$. For simplicity we use a notation X_k [resp. P_k] instead of $X(\boldsymbol{c}, \boldsymbol{y}_k\boldsymbol{v})$ [resp. $P(\boldsymbol{c}, \boldsymbol{y}_k\boldsymbol{v})$], and assume that $X_1 = X(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v})$. For given $\boldsymbol{c}$ and $\boldsymbol{v}$ we put $\ell_{\boldsymbol{c},\boldsymbol{v}} = \min\{X_k/X_1 \mid k = 1, \dots, G\}$ and $L_{\boldsymbol{c},\boldsymbol{v}} = \max\{X_k/X_1 \mid k = 1, \dots, G\}$. We first show that if $\delta < 2^{-1}\log(1 + M_1^{-1}(1 + GM_1)^{-1})$ (M_1 is the one given in (FD5)) then $0 < \inf\{\ell_{\boldsymbol{c},\boldsymbol{v}} \mid g(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v}) < \delta\} \le \sup\{L_{\boldsymbol{c},\boldsymbol{v}} \mid g(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v}) < \delta\} < \infty$, that is, for all $\boldsymbol{c}$ and $\boldsymbol{v}$ satisfying $g(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v}) < \delta$ (X_k/X_1)'s are bounded below and above uniformly by some positive constants. If $\inf\{\ell_{\boldsymbol{c},\boldsymbol{v}} \mid g(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v}) < \delta\} = 0$, then we can choose a sequence $\ell^{(n)}_{\boldsymbol{c}_n,\boldsymbol{v}_n} = X^{(n)}_{k_n}/X^{(n)}_1$, $n = 1, 2, \dots$, such that $\lim_{n\to\infty} \ell^{(n)}_{\boldsymbol{c}_n,\boldsymbol{v}_n} = 0$. By renumbering the indices if necessary we can assume that $\lim_{n\to\infty} X^{(n)}_2/X^{(n)}_1 = 0$. In the following we omit the superfix (n) in $X^{(n)}_k$ and $P^{(n)}_k$, and put $\boldsymbol{X} = \sum_{k=1}^G X^{(n)}_k$, $\check{\boldsymbol{X}} = \boldsymbol{X} - X^{(n)}_2$. Then we have

$$\begin{aligned}
g(\boldsymbol{c}_n, \check{\boldsymbol{y}}\boldsymbol{v}_n) &= \log\left(\frac{X_1}{\boldsymbol{X}} + \sum_{k=2}^G \frac{X_1 P_k}{\boldsymbol{X} P_1}\right) - \sum_{k=1}^G \frac{X_k}{\boldsymbol{X}} \log\left(\frac{X_1 P_k}{X_k P_1}\right) \\
&= \log\left(1 + \left(\frac{P_2}{P_1}\right)\left(1 + \sum_{k=3}^G \frac{P_k}{P_1}\right)^{-1}\right) + \log\left(\frac{\check{\boldsymbol{X}}}{\boldsymbol{X}}\right) \\
&\quad + \frac{\check{\boldsymbol{X}}}{\boldsymbol{X}}\left[\log\left(\frac{X_1}{\check{\boldsymbol{X}}} \cdot 1 + \sum_{k=3}^G \frac{X_k}{\check{\boldsymbol{X}}} \cdot \frac{X_1 P_k}{X_k P_1}\right)\right. \\
&\qquad\qquad \left. - \left\{\frac{X_1}{\check{\boldsymbol{X}}} \log(1) + \sum_{k=3}^G \frac{X_k}{\check{\boldsymbol{X}}} \log\left(\frac{X_1 P_k}{X_k P_1}\right)\right\}\right] \\
&\quad + \left(1 - \frac{\check{\boldsymbol{X}}}{\boldsymbol{X}}\right) \log\left(\frac{X_1}{\check{\boldsymbol{X}}} \cdot 1 + \sum_{k=3}^G \frac{X_1 P_k}{\check{\boldsymbol{X}} P_1}\right) - \frac{X_2}{\boldsymbol{X}} \log\left(\frac{X_1 P_2}{X_2 P_1}\right) \\
&\ge \log\left(1 + \frac{1}{M_1}(1 + GM_1)^{-1}\right) + \log\left(\frac{\check{\boldsymbol{X}}}{\boldsymbol{X}}\right) \\
&\quad + \frac{X_2}{\boldsymbol{X}} \log\left(\frac{X_1}{\check{\boldsymbol{X}}} \cdot 1 + \sum_{k=3}^G \frac{X_1 P_k}{\check{\boldsymbol{X}} P_1}\right) - \frac{X_2}{\boldsymbol{X}} \log\left(\frac{X_1 P_2}{X_2 P_1}\right) \\
&\equiv Z_1 + Z_2 + Z_3 + Z_4
\end{aligned}$$

and

$$\lim_{n\to\infty} Z_2 = \lim_{n\to\infty} \log\left(1 - \frac{X_2}{X}\right) = 0,$$

$$\lim_{n\to\infty} Z_3 = \lim_{n\to\infty}\left\{\frac{X_2\check{X}}{X_1 X}\Psi\left(\frac{X_1}{\check{X}}\right) + \frac{X_2}{X}\log\left(1 + \sum_{k=3}^{G}\frac{P_k}{P_1}\right)\right\} = 0,$$

$$\lim_{n\to\infty} Z_4 = \lim_{n\to\infty}\left\{\frac{X_1}{X}\Psi\left(\frac{X_2}{X_1}\right) + \frac{X_2}{X}\log\left(\frac{P_1}{P_2}\right)\right\} = 0,$$

which lead to a contradiction such as $g(\boldsymbol{c}_n, \check{\boldsymbol{y}}\boldsymbol{v}_n) > \delta$ for sufficiently large n. So $\inf\{\ell_{c,v} | g(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v}) < \delta\} = 0$ does not happen. Now suppose that $\sup\{L_{c,v} | g(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v}) < \delta\} = \infty$ does happen, and that we can choose a sequence $L^{(n)}_{c_n,v_n} = X^{(n)}_2 / X^{(n)}_1$, $n = 1, 2, \ldots$, such that $\lim_{n\to\infty} X^{(n)}_2 / X^{(n)}_1 = \infty$. Then, putting $\hat{X} = X - X^{(n)}_1$,

$$\begin{aligned}
g(\boldsymbol{c}_n, \check{\boldsymbol{y}}\boldsymbol{v}_n) &= \log\left(1 + \sum_{k=2}^{G}\frac{P_k}{P_1}\right) - \log\left(\sum_{k=2}^{G}\frac{P_k}{P_1}\right) + \log\left(\frac{\hat{X}}{X}\right) \\
&\quad + \frac{\hat{X}}{X}\left[\log\left(\sum_{k=2}^{G}\frac{X_k}{\hat{X}}\cdot\frac{X_1 P_k}{X_k P_1}\right) - \sum_{k=2}^{G}\frac{X_k}{\hat{X}}\log\left(\frac{X_1 P_k}{X_k P_1}\right)\right] \\
&\quad + \left(1 - \frac{\hat{X}}{X}\right)\log\left(\sum_{k=2}^{G}\frac{X_k}{\hat{X}}\cdot\frac{X_1 P_k}{X_k P_1}\right) \\
&\geq \log\left(1 + \left(\sum_{k=2}^{G}\frac{P_k}{P_1}\right)^{-1}\right) + \log\left(\frac{\hat{X}}{X}\right) + \frac{X_1}{X}\log\left(\sum_{k=2}^{G}\frac{X_1 P_k}{\hat{X} P_1}\right) \\
&\geq \log(1 + (GM_1)^{-1}) + \log\left(\frac{\hat{X}}{X}\right) + \frac{X_1}{X}\log\left(\sum_{k=2}^{G}\frac{X_1 P_k}{\hat{X} P_1}\right) \\
&\equiv Z_5 + Z_6 + Z_7
\end{aligned}$$

and

$$\lim_{n\to\infty} Z_6 = \lim_{n\to\infty}\log\left(1 - \frac{X_1}{X}\right) = 0,$$

$$\lim_{n\to\infty} Z_7 = \lim_{n\to\infty}\left\{\frac{\hat{X}}{X}\Psi\left(\frac{X_1}{\hat{X}}\right) + \frac{X_1}{X}\log\left(\sum_{k=2}^{G}\frac{P_k}{P_1}\right)\right\} = 0,$$

which lead to a contradiction such as $g(\boldsymbol{c}_n, \check{\boldsymbol{y}}\boldsymbol{v}_n) > \delta$ for sufficiently large n. Hence $\sup\{L_{c,v} | g(\boldsymbol{c}, \check{\boldsymbol{y}}\boldsymbol{v}) < \delta\} = \infty$ also does not happen.

Step 2. The rest of proof is easy. Let (δ_i) be a sequence satisfying $\delta_1 > \delta_2 > \delta_3 > \cdots \to 0$. To complete the proof it is sufficient to show that

$$\sup_{\substack{c, v \\ g(c, \check{y}v) < \delta_i}} \max\left\{\left|1 - \frac{X_1 P_k}{X_k P_1}\right| \,\middle|\, k = 1, 2, \ldots, G\right\} \to 0$$

as $i \to \infty$. Suppose it is not true, and so, there exist $\boldsymbol{c}_i$ and $\boldsymbol{v}_i$, $i = 1, 2, \ldots,$ such that

$$g(\boldsymbol{c}_i, \check{\boldsymbol{y}}\boldsymbol{v}_i) < \delta_i \quad \text{and} \quad \max\left\{\left|1 - \frac{X_1^{(i)} P_k^{(i)}}{X_k^{(i)} P_1^{(i)}}\right| \,\middle|\, k = 1, 2, \ldots, G\right\} \to {}^{\exists}\gamma_0 > 0.$$

Without loss of generality we can assume that $\lim_{i\to\infty}\left|1 - \dfrac{X_1 P_2}{X_2 P_1}\right| = \gamma_0$. However this implies that

$$\begin{aligned} g(\boldsymbol{c}_i, \check{\boldsymbol{y}}\boldsymbol{v}_i) &= \log\left(\frac{X_1}{\boldsymbol{X}} \cdot 1 + \frac{X_2}{\boldsymbol{X}} \cdot \frac{X_1 P_2}{X_2 P_1} + \cdots + \frac{X_G}{\boldsymbol{X}} \cdot \frac{X_1 P_G}{X_G P_1}\right) \\ &\quad - \left\{\frac{X_1}{\boldsymbol{X}} \log(1) + \frac{X_2}{\boldsymbol{X}} \log\left(\frac{X_1 P_2}{X_2 P_1}\right) + \cdots \frac{X_G}{\boldsymbol{X}} \log\left(\frac{X_1 P_G}{X_G P_1}\right)\right\} \end{aligned}$$

does not converge to zero because of the strict concavity of log-function and the boundedness of (X_k/X_1)'s. This is a contradiction. □

Proof of Lemma 5. *Step 1.* Since

$$\begin{aligned} H_N(\tilde{v}) &= H_N(v) \\ &= \sum\nolimits^*_{\boldsymbol{a}\in\mathscr{C}_N} \sum\nolimits_{\boldsymbol{w}\in\mathscr{C}_{N-R}} \sum\nolimits_{\tilde{\boldsymbol{w}}\in\mathscr{C}_{N-R}} \mu(\boldsymbol{aw}) P(\boldsymbol{aw}, \tilde{\boldsymbol{w}}) \Psi\left(\frac{v(\boldsymbol{aw}) P(\boldsymbol{aw}, \tilde{\boldsymbol{w}})}{\mu(\boldsymbol{aw}) P(\boldsymbol{aw}, \tilde{\boldsymbol{w}})}\right) \\ &= \sum\nolimits^*_{\boldsymbol{a}} \sum\nolimits_{\boldsymbol{w}, \tilde{\boldsymbol{w}}} \mu(\boldsymbol{a}\tilde{\boldsymbol{w}}) P(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w}) \Psi\left(\frac{Y(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w})}{\mu(\boldsymbol{a}\tilde{\boldsymbol{w}}) P(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w})}\right) \qquad \text{[by (3.2)]}, \\ &\qquad \text{where } Y(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w}) \equiv v(\boldsymbol{aw}) P(\boldsymbol{aw}, \tilde{\boldsymbol{w}}), \end{aligned}$$

the difference (3.6) is reduced to

$$\begin{aligned} \text{(4.4)} \quad \Lambda_N(v) &= \sum\nolimits^*_{\boldsymbol{a}} \sum\nolimits_{\boldsymbol{w}, \tilde{\boldsymbol{w}}} \{\Psi(X(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w})) - \Psi(Y(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w}))\} \\ &\quad + \sum\nolimits^*_{\boldsymbol{a}} \sum\nolimits_{\boldsymbol{w}, \tilde{\boldsymbol{w}}} \{X(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w}) - Y(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w})\} \log \frac{1}{\mu(\boldsymbol{a}\tilde{\boldsymbol{w}}) P(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w})} \\ &\equiv \Lambda_N^{(1)}(v) + \Lambda_N^{(2)}(v). \end{aligned}$$

Let us show first

$$\text{(4.5)} \qquad \lim_{N\to\infty} \sum\nolimits^*_{\boldsymbol{a}} \sum\nolimits_{\boldsymbol{w}, \tilde{\boldsymbol{w}}} |X(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w}) - Y(\boldsymbol{a}\tilde{\boldsymbol{w}}, \boldsymbol{w})| = 0.$$

We have

$$\sum_a^* \sum_{w,\tilde{w}} |X(a\tilde{w}, w) - Y(a\tilde{w}, w)|$$

$$= \sum_a^* \sum_{w,\tilde{w}} \left| \left(\sum_{\tilde{a}}^* \sum_{b \in \mathscr{C}_{N+R}}^* v(b\tilde{a}w) P(b\tilde{a}w, a\tilde{w}) \right) - v(aw) P(aw, \tilde{w}) \right|$$

$$\leq \sum_b^* \sum_{a,\tilde{a}}^* \sum_{w,\tilde{w}}^* \sum_{v,\tilde{v} \in \mathscr{C}_{N-2R}} \tag{4.6}$$

$$|v(b\tilde{a}wv) P(b\tilde{a}wv, a\tilde{w}\tilde{v}) - v(bawv) P(bawv, \tilde{a}\tilde{w}\tilde{v})|$$

and, using K_1 in (FD4),

$$v(b\breve{a}wv) = \sum_{c \in \mathscr{C}_{N+2R+K_1}}^* \sum_{y \in \mathscr{C}_{N+R+K_1}}^* \tilde{v}(cyb\breve{a}wv)$$

$$= \sum_{c,y}^* \sum_{\check{y} \in \mathscr{C}_{N+R+K_1}}^* \sum_{\check{v} \in \mathscr{C}_{N-2R}} X(cyb\breve{a}wv, \check{y}\check{v}) \qquad \text{[by (3.4)]}$$

$$= \sum_{c,y}^* \sum_{\check{y}}^* \sum_{\check{v}} P(cyb\breve{a}wv, \check{y}\check{v}) \left\{ \frac{X(cyb\breve{a}wv, \check{y}\check{v})}{P(cyb\breve{a}wv, \check{y}\check{v})} - \frac{X(cyb\breve{a}wv, ybaw\check{v})}{P(cyb\breve{a}wv, ybaw\check{v})} \right\}$$

$$+ \sum_{c,y}^* \sum_{\check{v}} P(wv, \check{v}) \frac{X(cyb\breve{a}wv, ybaw\check{v})}{P(cyb\breve{a}wv, ybaw\check{v})}$$

$$\equiv T(b\breve{a}wv) + S(b\breve{a}wv).$$

Therefore setting $\breve{a}\langle * \rangle = \tilde{a}\langle * \rangle$ and $\breve{a}\langle * \rangle = a\langle * \rangle$ respectively in the above yields

$$(4.6) \leq \sum_b^* \sum_{a,\tilde{a}}^* \sum_{w,\tilde{w}}^* \sum_{v,\tilde{v}}$$

$$\{|T(b\tilde{a}wv)| P(b\tilde{a}wv, a\tilde{w}\tilde{v}) + |T(bawv)| P(bawv, \tilde{a}\tilde{w}\tilde{v})$$

$$+ |S(b\tilde{a}wv) P(b\tilde{a}wv, a\tilde{w}\tilde{v}) - S(bawv) P(bawv, \tilde{a}\tilde{w}\tilde{v})|\}$$

$$\equiv T_1 + T_2 + \sum_b^* \sum_{a,\tilde{a}}^* \sum_{w,\tilde{w}}^* \sum_{v,\tilde{v}} |S_1 - S_2|.$$

Step 2. Let us show that $S_1 = S_2$. Because

$$\frac{X(cyb\breve{a}wv, ybaw\check{v})}{P(cyb\breve{a}wv, ybaw\check{v})}$$

$$= \sum_{\check{d} \in \mathscr{C}_{N+3R+K_1}}^* \sum_{\check{c}}^* \frac{v(\check{d}\check{c}ybaw\check{v}) P(\check{d}\check{c}ybaw\check{v}, cyb\breve{a}wv)}{P(cyb\breve{a}wv, ybaw\check{v})}$$

$$= \sum_{\check{d},\check{c}}^* v(\check{d}\check{c}ybaw\check{v}) \times \frac{P(\check{d}cyb\breve{a}wv, \check{c}ybaw\check{v})}{P(cyb\breve{a}wv, ybaw\check{v})} \times \frac{\exp\{-\mathscr{H}(\check{d}cyb\breve{a}wv)\}}{\exp\{-\mathscr{H}(\check{d}\check{c}ybaw\check{v})\}}$$

$$\equiv \sum_{\check{d},\check{c}}^* v(\check{d}\check{c}ybaw\check{v}) \times U_p(\breve{a}) \times U_e(\breve{a})$$

by (2.3), we have

$$S_1 = P(\boldsymbol{b\tilde{a}wv}, \boldsymbol{a\tilde{w}\tilde{v}}) \sum\nolimits_{c,y}^{*} \sum\nolimits_{\check{v}} P(\boldsymbol{wv}, \check{\boldsymbol{v}}) \sum\nolimits_{\check{d},\check{c}}^{*} \nu(\check{\boldsymbol{d}}\check{\boldsymbol{c}}\boldsymbol{ybaw}\check{\boldsymbol{v}}) U_p(\tilde{\boldsymbol{a}}) U_e(\tilde{\boldsymbol{a}})$$

and

$$S_2 = P(\boldsymbol{bawv}, \boldsymbol{\tilde{a}\tilde{w}\tilde{v}}) \sum\nolimits_{c,y}^{*} \sum\nolimits_{\check{v}} P(\boldsymbol{wv}, \check{\boldsymbol{v}}) \sum\nolimits_{\check{d},\check{c}}^{*} \nu(\check{\boldsymbol{d}}\check{\boldsymbol{c}}\boldsymbol{ybaw}\check{\boldsymbol{v}}) U_p(\boldsymbol{a}) U_e(\boldsymbol{a}).$$

From

$$\frac{P(\check{\boldsymbol{d}}\boldsymbol{cy}\langle\boldsymbol{bawv}\rangle, \check{\boldsymbol{c}}\boldsymbol{y}\langle\boldsymbol{baw}\check{\boldsymbol{v}}\rangle)}{P(\boldsymbol{cy}\langle\boldsymbol{bawv}\rangle, \boldsymbol{y}\langle\boldsymbol{baw}\check{\boldsymbol{v}}\rangle)} = \frac{P(\check{\boldsymbol{d}}\boldsymbol{cy}\langle\boldsymbol{b\tilde{a}wv}\rangle, \check{\boldsymbol{c}}\boldsymbol{y}\langle\boldsymbol{baw}\check{\boldsymbol{v}}\rangle)}{P(\boldsymbol{cy}\langle\boldsymbol{b\tilde{a}wv}\rangle, \boldsymbol{y}\langle\boldsymbol{baw}\check{\boldsymbol{v}}\rangle)} \qquad \text{[by (FD4}')],$$

$$\frac{\exp\{-\mathscr{H}(\check{\boldsymbol{d}}\boldsymbol{cybawv})\}}{\exp\{-\mathscr{H}(\check{\boldsymbol{d}}\check{\boldsymbol{c}}\boldsymbol{ybaw}\check{\boldsymbol{v}})\}} = \frac{\exp\{-\mathscr{H}(\check{\boldsymbol{d}}\boldsymbol{cyb\tilde{a}wv})\}}{\exp\{-\mathscr{H}(\check{\boldsymbol{d}}\check{\boldsymbol{c}}\boldsymbol{ybaw}\check{\boldsymbol{v}})\}} \cdot \frac{\exp\{-\mathscr{H}(\boldsymbol{bawv})\}}{\exp\{-\mathscr{H}(\boldsymbol{b\tilde{a}wv})\}},$$

it holds that

$$U_p(\boldsymbol{a}) = U_p(\tilde{a}) \qquad \text{and} \qquad U_e(\boldsymbol{a}) = U_e(\tilde{\boldsymbol{a}}) \cdot \frac{\exp\{-\mathscr{H}(\boldsymbol{bawv})\}}{\exp\{-\mathscr{H}(\boldsymbol{b\tilde{a}wv})\}}.$$

Since

$$P(\boldsymbol{ba}\langle\boldsymbol{wv}\rangle, \tilde{\boldsymbol{a}}\langle\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}}\rangle) = P(\boldsymbol{b\tilde{a}}\langle\boldsymbol{wv}\rangle, \boldsymbol{a}\langle\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}}\rangle) \frac{\exp\{-\mathscr{H}(\boldsymbol{b\tilde{a}}\langle\boldsymbol{wv}\rangle)\}}{\exp\{-\mathscr{H}(\boldsymbol{ba}\langle\boldsymbol{wv}\rangle)\}}$$

by (2.4), we finally have $S_1 = S_2$.

Step 3. Let us show $\lim_{N\to\infty} T_1 = \lim_{N\to\infty} T_2 = 0$. We split the term T_1 into three parts:

$$T_1 = \sum\nolimits_{b}^{*} \sum\nolimits_{a,\tilde{a}}^{*} \sum\nolimits_{w,\tilde{w}}^{*} \sum\nolimits_{v,\tilde{v}} P(\boldsymbol{b\tilde{a}wv}, \boldsymbol{a\tilde{w}\tilde{v}})\{V_1 + V_2 + V_3\},$$

where

$$V_1 = \sum\nolimits_{\substack{c,y \\ g(\boldsymbol{cyb\tilde{a}wv}, \boldsymbol{ybaw}\langle\check{\boldsymbol{v}}\rangle)<\delta}}^{*} \sum\nolimits_{\check{y}}^{*} \sum\nolimits_{\check{y}} X(\boldsymbol{cyb\tilde{a}wv}, \check{\boldsymbol{y}}\langle\check{\boldsymbol{v}}\rangle)$$
$$\times \left|1 - \frac{X(\boldsymbol{cyb\tilde{a}wv}, \boldsymbol{ybaw}\langle\check{\boldsymbol{v}}\rangle)}{X(\boldsymbol{cyb\tilde{a}wv}, \check{\boldsymbol{y}}\langle\check{\boldsymbol{v}}\rangle)} \cdot \frac{P(\boldsymbol{cyb\tilde{a}wv}, \check{\boldsymbol{y}}\langle\check{\boldsymbol{v}}\rangle)}{P(\boldsymbol{cyb\tilde{a}wv}, \boldsymbol{ybaw}\langle\check{\boldsymbol{v}}\rangle)}\right|,$$

$$V_2 = \sum\nolimits_{\substack{c,y \\ g(\boldsymbol{cyb\tilde{a}wv}, \boldsymbol{ybaw}\langle\check{\boldsymbol{v}}\rangle)\geq\delta}}^{*} \sum\nolimits_{\check{y}}^{*} \sum\nolimits_{\check{v}} X(\boldsymbol{cyb\tilde{a}wv}, \check{\boldsymbol{y}}\langle\check{\boldsymbol{v}}\rangle),$$

$$V_3 = \sum\nolimits_{\substack{c,y \\ g(\boldsymbol{cyb\tilde{a}wv}, \boldsymbol{ybaw}\langle\check{\boldsymbol{v}}\rangle)\geq\delta}}^{*} \sum\nolimits_{\check{y}}^{*} \sum\nolimits_{\check{v}} \frac{P(\boldsymbol{cyb\tilde{a}wv}, \check{\boldsymbol{y}}\langle\check{\boldsymbol{v}}\rangle)}{P(\boldsymbol{cyb\tilde{a}wv}, \boldsymbol{ybaw}\langle\check{\boldsymbol{v}}\rangle)} X(\boldsymbol{cyb\tilde{a}wv}, \boldsymbol{ybaw}\langle\check{\boldsymbol{v}}\rangle).$$

As to V_1 we have

$$V_1 \leq \sum\nolimits^*_{\substack{c,y \\ g(cyb\tilde{a}wv,\, ybaw\langle\check{v}\rangle)<\delta}} \sum\nolimits^*_{\check{y}} \sum\nolimits_{\check{v}} X(cyb\tilde{a}wv, \check{y}\langle\check{v}\rangle)\gamma$$

by (3.8), and so, by (3.4)

$$\begin{aligned}
&\sum\nolimits^*_b \sum\nolimits^*_{a,\tilde{a}} \sum\nolimits^*_{w,\tilde{w}} \sum\nolimits_{v,\tilde{v}} P(b\tilde{a}wv, a\tilde{w}\tilde{v}) V_1 \\
&\quad \leq \sum\nolimits^*_b \sum\nolimits^*_{a,\tilde{a}} \sum\nolimits^*_{w,\tilde{w}} \sum\nolimits_{v,\tilde{v}} P(b\tilde{a}wv, a\tilde{w}\tilde{v}) \sum\nolimits^*_{c,y} \nu(cyb\tilde{a}wv)\gamma \\
&\quad \leq \gamma.
\end{aligned}$$

Since $P(b\tilde{a}wv, a\tilde{w}\tilde{v}) \leq P(\tilde{a}wv, \tilde{w}\tilde{v})$ and $\sum^*_{\tilde{w}} \sum_{\tilde{v}} P(\tilde{a}wv, \tilde{w}\tilde{v}) = 1$,

$$\begin{aligned}
&\sum\nolimits^*_b \sum\nolimits^*_{a,\tilde{a}} \sum\nolimits^*_{w,\tilde{w}} \sum\nolimits_{v,\tilde{v}} P(b\tilde{a}wv, a\tilde{w}\tilde{v}) V_2 \\
&\quad \leq \sum\nolimits^*_b \sum\nolimits^*_{a,\tilde{a}} \sum\nolimits^*_w \sum\nolimits_v \left\{ \sum\nolimits^*_{\substack{c,y \\ g(cyb\tilde{a}wv,\, ybaw\langle\check{v}\rangle)\geq\delta}} \sum\nolimits^*_{\check{y}} \sum\nolimits_{\check{v}} X(cyb\tilde{a}wv, \check{y}\langle\check{v}\rangle) \right\} \\
&\quad \leq \sum\nolimits^*_{y,b,a,w} \left[\sum\nolimits^*_{\substack{c,\hat{y},\hat{b},\tilde{a},\hat{w} \\ g(c\hat{y}\hat{b}\tilde{a}\hat{w}v,\, ybaw\langle\check{v}\rangle)\geq\delta}} \sum\nolimits_v \sum\nolimits_{\check{v}} \left\{ \sum\nolimits^*_{\check{y}} X(c\hat{y}\hat{b}\tilde{a}\hat{w}v, \check{y}\langle\check{v}\rangle) \right\} \right] \\
&\quad \leq 2^{2(3R+K_1)}\gamma
\end{aligned}$$

by (3.7) for sufficiently large N. Analogously, using (FD5),

$$\begin{aligned}
&\sum\nolimits^*_b \sum\nolimits^*_{a,\tilde{a}} \sum\nolimits^*_{w,\tilde{w}} \sum\nolimits_{v,\tilde{v}} P(b\tilde{a}wv, a\tilde{w}\tilde{v}) V_3 \\
&\quad \leq \sum\nolimits^*_b \sum\nolimits^*_{a,\tilde{a}} \sum\nolimits^*_w \sum\nolimits_v \left\{ \sum\nolimits^*_{\substack{c,y \\ g(cyb\tilde{a}wv,\, ybaw\langle\check{v}\rangle)\geq\delta}} \sum\nolimits^*_{\check{y}} \sum\nolimits_{\check{v}} M_1 X(cyb\tilde{a}wv, ybaw\langle\check{v}\rangle) \right\} \\
&\quad \leq M_1 \sum\nolimits^*_{y,b,a,w} \sum\nolimits^*_{\check{y}} \left[\sum\nolimits^*_{\substack{c,\tilde{y},\tilde{b},\tilde{a},\tilde{w} \\ g(c\tilde{y}\tilde{b}\tilde{a}\tilde{w}v,\, ybaw\langle\check{v}\rangle)\geq\delta}} \sum\nolimits_v \sum\nolimits_{\check{v}} \left\{ \sum\nolimits^*_{\hat{y}} X(c\tilde{y}\tilde{b}\tilde{a}\tilde{w}v, \hat{y}\langle\check{v}\rangle) \right\} \right] \\
&\quad \leq M_1 2^{4(3R+K_1)}\gamma
\end{aligned}$$

by (3.7). Summing up these estimates gives us

$$T_1 \leq \gamma + 2^{2(3R+K_1)}\gamma + M_1 2^{4(3R+K_1)}\gamma.$$

As we can take γ arbitrarily small, it follows that $\lim_{N\to\infty} T_1 = 0$. Similarly we can show that $\lim_{N\to\infty} T_2 = 0$. Thus (4.5) is verified.

Step 4. Let us complete the proof. From the definition of $X(\boldsymbol{a},\boldsymbol{w})$ and $Y(\boldsymbol{a},\boldsymbol{w})$, for each $\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}}$ and $\boldsymbol{v}$ it holds that

$$\sum\nolimits_{\boldsymbol{a}}^{*}\sum\nolimits_{\boldsymbol{w}}^{*} X(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v}) = \sum\nolimits_{\boldsymbol{a}}^{*}\sum\nolimits_{\boldsymbol{w}}^{*} Y(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v}), \tag{4.8}$$

and so

$$\sum\nolimits_{\boldsymbol{a}}^{*}\sum\nolimits_{\boldsymbol{w},\tilde{\boldsymbol{w}}}^{*}\sum\nolimits_{\boldsymbol{v},\tilde{\boldsymbol{v}}}\{X(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v}) - Y(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v})\}\log\{\mu(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}})P(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})\} = 0.$$

Hence $\Lambda_N^{(2)}(v)$ in (4.4) is equal to

$$\sum\nolimits_{\boldsymbol{a}}^{*}\sum\nolimits_{\boldsymbol{w},\tilde{\boldsymbol{w}}}^{*}\sum\nolimits_{\boldsymbol{v},\tilde{\boldsymbol{v}}}\{X(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v}) - Y(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v})\}\log\frac{\mu(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}})P(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})}{\mu(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}})P(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v})},$$

whose log-part it bounded uniformly from (FD5) and (2.5) with $v=\mu$. Therefore $\lim_{N\to\infty}\Lambda_N^{(2)}(v)=0$ follows from (4.5). Finally let us show that $\lim_{N\to\infty}\Lambda_N^{(1)}(v)=0$. If we put the value of (4.8) as $k(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})$, we have

$$\Lambda_N^{(1)} = \sum\nolimits_{\boldsymbol{a}}^{*}\sum\nolimits_{\boldsymbol{w},\tilde{\boldsymbol{w}}}^{*}\sum\nolimits_{\boldsymbol{v},\tilde{\boldsymbol{v}}} k(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})\left\{\Psi\left(\frac{X(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v})}{k(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})}\right) - \Psi\left(\frac{Y(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v})}{k(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})}\right)\right\},$$

and each argument of Ψ is less than or equal to one. It is elementary to check that for a given $\varepsilon>0$ there exists a constant $M_\varepsilon>0$ such that $|\Psi(u)-\Psi(v)|<\varepsilon+M_\varepsilon|u-v|$ for every $0\le u,\ v\le 1$. Hence for any fixed $\varepsilon>0$ we have

$$|\Lambda_N^{(1)}| \le \sum\nolimits_{\boldsymbol{a}}^{*}\sum\nolimits_{\boldsymbol{w},\tilde{\boldsymbol{w}}}^{*}\sum\nolimits_{\boldsymbol{v},\tilde{\boldsymbol{v}}}(k(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})\varepsilon + M_\varepsilon|X(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v}) - Y(\boldsymbol{a}\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{w}\boldsymbol{v})|).$$

Since $\sum_{\tilde{\boldsymbol{w}}}^{*}\sum_{\tilde{\boldsymbol{v}},\boldsymbol{v}} k(\tilde{\boldsymbol{w}}\tilde{\boldsymbol{v}},\boldsymbol{v})=1$ and $\sum_{\boldsymbol{a}}^{*}\sum_{\boldsymbol{w}}^{*}1=2^{4R}$, we have the result. □

Proof of Lemma 6. This is almost obvious. Let $\boldsymbol{a}$ be in $\mathscr{C}_{i,j}$. Since $\sum_{\boldsymbol{b}\in\mathscr{C}_{i-R,j+R}} v(\boldsymbol{b})=1$, there exists an element $\tilde{\boldsymbol{b}}$ such that $v(\tilde{\boldsymbol{b}})>0$. Then $v(\boldsymbol{a})=\tilde{v}(\boldsymbol{a})>0$ by (3.5) and (FD1) with $\boldsymbol{b}\langle\tilde{\boldsymbol{a}}\langle w\rangle\rangle=\tilde{\boldsymbol{b}}$. □

5. An example of the discrete-time Ising model

We give here an example of the discrete-time stochastic Ising models satisfying the conditions (FD1)–(FD5). Its intuitive interpretation is given at the end of this section.

Let $\mathscr{X}=\{(x,y)\in\mathbf{Z}^2\,|\,x\le y, y-x\le R\}$, and set $\mathscr{E}=\{\circ,\times\}^{\mathscr{X}}$. We endow $\mathscr{E}$ with the topology given by the product of the discrete topology and consider a Borel structure on $\mathscr{E}$ as usual. Let α_i and β_i, $i=0,1,\dots,R$, be numbers in $(0,1)$. Suppose that the configuration of spins on $\mathbf{Z}$ at time t is $\eta \in \mathscr{X}$. Then we attach the state $\circ$(permission) or $\times$(prohibition) to each $(x,y) \in \mathscr{Z}$ as follows:

i) if $\eta_x = +1$, attach $\circ$ [resp. $\times$] to (x,x) with probability α_0 [resp. $1-\alpha_0$];

ii) if $\eta_x = -1$, attach $\circ$ [resp. $\times$] to (x,x) with probability β_0 [resp. $1-\beta_0$];

iii) if $\eta_x = \eta_y$ for $x \neq y$, attach $\circ$ [resp. $\times$] to (x,y) with probability α_{y-x} [resp. $1-\alpha_{y-x}$];

iv) if $\eta_x \neq \eta_y$ for $x \neq y$, attach $\circ$ [resp. $\times$] to (x,y) with probability β_{y-x} [resp. $1-\beta_{y-x}$];

v) the random choices of $\circ$ and $\times$ for $(x,y) \in \mathscr{Z}$ are independent.

For each η the above rule defines a probability measure Q_η on $\mathscr{E}$. As a time evolution of spin-configurations on $\mathbf{Z}$ as time goes to $t+1$, we reverse the spin-orientation on the site i if and only if the state $\circ$ is attached to all $(x,y) \in \mathscr{Z}$ satisfying $\{x,y\} \ni i$, i.e., to every $(i-R,i),\ldots,(i,i),\ldots,(i,i+R)$. Thus the transition probabilities $P_\eta, \eta \in \mathscr{X}$, are determined through Q_η. It is easy to check that each P_η satisfies (FD1)–(FD5) except (FD3). Let us check (FD3). By the definition of $P(\eta,\cdot) \equiv P_\eta(\cdot)$, we have for $\boldsymbol{a}, \tilde{\boldsymbol{a}} \in \mathscr{C}_{i,j}$ and $\boldsymbol{b}\langle * \rangle \in \mathscr{C}_{i-R,j+R}$

$$\frac{P(\boldsymbol{b}\langle \boldsymbol{a} \rangle, \tilde{\boldsymbol{a}})}{P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}} \rangle, \boldsymbol{a})} = \left(\frac{\alpha_0}{\beta_0}\right)^{\frac{1}{2}\sum_{k=i}^{j}(a_k-\tilde{a}_k)} \prod_{r=1}^{R} \left(\frac{\alpha_r}{\beta_r}\right)^{\frac{1}{2}\sum_{k=i}^{j+r}(a_{k-r}a_k-\tilde{a}_{k-r}\tilde{a}_k)},$$

here a_k and $\tilde{a}_k$ for $k<i$ or $k>j$ should be read as b_k. Therefore if we define

$$J_r = \frac{1}{2}\log\left(\frac{\alpha_r}{\beta_r}\right), \qquad r = 0,1,\ldots,R,$$

the above equals

$$\exp\left[J_0\left\{\sum_{k=i}^{j}(a_i-\tilde{a}_k)\right\}\right] \times \prod_{r=1}^{R} \exp\left[J_r \sum_{k=i}^{j+r}(a_{k-r}a_k - \tilde{a}_{k-r}\tilde{a}_k)\right]$$
$$= \frac{\exp\{-\mathscr{H}(\boldsymbol{b}\langle \tilde{\boldsymbol{a}} \rangle)\}}{\exp\{-\mathscr{H}(\boldsymbol{b}\langle \boldsymbol{a} \rangle)\}},$$

which yields (2.3). Let $\boldsymbol{w}, \tilde{\boldsymbol{w}}, \boldsymbol{a}\langle * \rangle, \tilde{\boldsymbol{a}}\langle * \rangle$ and $\boldsymbol{b}\langle * \rangle$ be as in (FD3)–(ii). Since the set of sites at which reversal of spins occurs in $P(\boldsymbol{b}\langle \boldsymbol{a}\langle \boldsymbol{w} \rangle \rangle, \tilde{\boldsymbol{a}}\langle \tilde{\boldsymbol{w}} \rangle)$ and $P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\langle \boldsymbol{w} \rangle \rangle, \boldsymbol{a}\langle \tilde{\boldsymbol{w}} \rangle)$ is the same, and since the reversal of spins for $\boldsymbol{w}$ is the same, it holds that

$$\frac{P(\boldsymbol{b}\langle \boldsymbol{a}\langle \boldsymbol{w}\rangle\rangle, \tilde{\boldsymbol{a}}\langle \tilde{\boldsymbol{w}}\rangle)}{P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\langle \boldsymbol{w}\rangle\rangle, \boldsymbol{a}\langle \tilde{\boldsymbol{w}}\rangle)}$$

$$= \exp\Big[J_0\Big\{\Big(\sum_{k=i}^{i+R-1} + \sum_{k=j-R+1}^{j}\Big)(a_k - \tilde{a}_k)\Big\}\Big]$$

$$\times \prod_{r=1}^{R} \exp\Big[J_r\Big\{\Big(\sum_{k=i}^{i+R-1} + \sum_{k=j-R+1+r}^{j+r}\Big)(a_{k-r}a_k - \tilde{a}_{k-r}\tilde{a}_k)$$

$$+ \sum_{k=i+R}^{i+R-1+r}(a_{k-r}w_k - \tilde{a}_{k-r}w_k)$$

$$+ \sum_{k=j-R+1}^{j-R+r}(w_{k-r}a_k - w_{k-r}\tilde{a}_k)\Big\}\Big]$$

$$= \frac{\exp\{-\mathscr{H}(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\langle \boldsymbol{w}\rangle\rangle)\}}{\exp\{-\mathscr{H}(\boldsymbol{b}\langle \boldsymbol{a}\langle \boldsymbol{w}\rangle\rangle)\}},$$

which is (2.4). Therefore by Theorem 1 and the succeeding remark we know that the stationary measure for the corresponding discrete-time stochastic Ising model is unique and is given by the Gibbs state associated with the potentials $J_r, r = 0, \dots, R$.

We can interpret this example as an extended version of (2.1) and (2.2). Indeed, letting $\delta_x(y) = 1$ if $x = y$ and $=0$ otherwise, define

$$P(\eta, {}_i[\bar{\eta}_i\bar{\eta}_{i+1}\cdots\bar{\eta}_j]_j) = \Big[\prod_{k=i}^{j}\{\delta_{+1}(\eta_k)\alpha_0 + \delta_{-1}(\eta_k)\beta_0\}\Big]$$

$$\times \prod_{r=1}^{R}\Big[\prod_{k=i}^{j+r}\{\delta_{+1}(\eta_{k-r}\eta_k)\alpha_r + \delta_{-1}(\eta_{k-r}\eta_k)\beta_r\}\Big],$$

and

$$P(\eta, A_1 \cap A_2 \cap \dots \cap A_n)$$

$$= \Big[\prod_{s=1}^{n} P(\eta, A_s)\Big] \Big/ \prod_{r=2}^{R}\Big[\widetilde{\prod_{k=i_1}^{j_n+r}}\{\delta_{+1}(\eta_{k-r}\eta_k)\alpha_r + \delta_{-1}(\eta_{k-r}\eta_k)\beta_r\}\Big]$$

$$\text{for } A_s = {}_{i_s}[\bar{\eta}_{i_s}\cdots\bar{\eta}_{j_s}]_{j_s} \text{ with } i_s - j_{s-1} \geq 2,$$

here $\widetilde{\prod}$ means that the product over $\eta_{k-r}\eta_k$ should be taken only for the pairs (η_{k-r}, η_k) such that $\bar{\eta}_{k-r}$ and $\bar{\eta}_k$ belong to different A_s. Then it is elementary to check that this inductively determines a probability measure $P(\eta, \cdot)$ on $\mathscr{X}$, which coincides with the measure determined by Q_η. In fact, denoting ${}_i[\bar{\eta}_i\cdots\bar{\eta}_{k-1}]_{k-1} \cap {}_{k+1}[\bar{\eta}_{k+1}\cdots\bar{\eta}_j]_j$ by ${}_i[\bar{\eta}_i\cdots\bar{\eta}_{k-1} * \bar{\eta}_{k+1}\cdots\bar{\eta}_j]_j$, we have

$$(5.1)\quad P(\eta, {}_i[\bar{\eta}_i\cdots\bar{\eta}_{k-1}\eta_k\bar{\eta}_{k+1}\cdots\bar{\eta}_j]_j)$$

$$= P(\eta, {}_i[\bar{\eta}_i\cdots\bar{\eta}_{k-1} * \bar{\eta}_{k+1}\cdots\bar{\eta}_j]_j) - P(\eta, {}_i[\bar{\eta}_i\cdots\bar{\eta}_{k-1}\bar{\eta}_k\bar{\eta}_{k+1}\cdots\bar{\eta}_j]_j),$$

$$P(\eta, {}_i[\bar{\eta}_i \cdots \bar{\eta}_{k-1}\eta_k\bar{\eta}_{k+1} \cdots \bar{\eta}_{l-1}\eta_l\bar{\eta}_{l+1} \cdots \bar{\eta}_j]_j)$$
$$= P(\eta, {}_i[\bar{\eta}_i \cdots \bar{\eta}_{k-1} * \bar{\eta}_{k+1} \cdots \bar{\eta}_{l-1} * \bar{\eta}_{l+1} \cdots \bar{\eta}_j]_j)$$
$$- \sum\nolimits_{(\zeta_k,\zeta_l)\neq(\eta_k,\eta_l)} P(\eta, {}_i[\bar{\eta}_i \cdots \bar{\eta}_{k-1}\zeta_k\bar{\eta}_{k+1} \cdots \bar{\eta}_{l-1}\zeta_l\bar{\eta}_{l+1} \cdots \bar{\eta}_j]_j)$$

and so on. We remark that $P(\eta, A_s \cap A_{s+1}) = P(\eta, A_s)P(\eta, A_{s+1})$ if $i_{s+1} - j_s \geq R + 1$.

6. Application to exclusion processes

In this section we apply the preceding argument to discrete-time interactive exclusion processes on the one-dimensional lattice. In the processes particles are located at most one on each site of $\mathbf{Z}$, and each particle executes mutually exclusive random walk to an unoccupied site situated within the range R. Infinitely many particles can move simultaneously at time $t = 0, 1, \ldots$. The movement of a particle is affected by the particles which are located within the distance R from both of two related sites concerning the jump of particle. As $\mathscr{X}$ we take $\{0,1\}^{\mathbf{Z}}$ and consider that there exists a particle at site i iff $\eta_i = 1$ for $(\ldots \eta_{i-1}\eta_i\eta_{i+1} \ldots) \in \mathscr{X}$. We use the notations given in the previous sections without any comments.

Let $\mathscr{Z} = \{(x, y) \in \mathbf{Z}^2 \mid 0 < y - x \leq R\}$. Each element of $\mathscr{Z}$ indicates a pair (x, y) of sites x and y at which the values η_x and η_y of $(\ldots \eta_{-1}\eta_0\eta_1 \ldots)$ might be exchanged. We put $\mathscr{E} = \{0,1\}^{\mathscr{Z}}$, and let $\omega(x, y)$ denote the value of $\omega \in \mathscr{E}$ at $(x, y) \in \mathscr{Z}$. (For convenience sake we set $\omega(x, y) = 0$ if $(x, y) \notin \mathscr{Z}$.) Because we want to exchange values η_x and η_y for all (x, y)'s satisfying $\omega(x, y) = 1$, we extract an essential part $\mathscr{E}^*$ from $\mathscr{E}$ by

$$\mathscr{E}^* = \{\omega \in \mathscr{E} : \text{if } \omega(x, y) = 1 \text{ then } \omega(\tilde{x}, \tilde{y}) = 0 \text{ for all } (\tilde{x}, \tilde{y}) \in \mathscr{Z}$$
$$\text{satisfying } (\tilde{x}, \tilde{y}) \neq (x, y) \text{ and } \{\tilde{x}, \tilde{y}\} \cap \{x, y\} \neq \emptyset\}.$$

Then for $\eta \equiv (\ldots \eta_{-1}\eta_0\eta_1 \ldots) \in \mathscr{X}$, by exchanging the values η_x and η_y iff $\omega(x, y) = 1$, we can associate each element ω of $\mathscr{E}^*$ with a movement of infinitely many particles starting from η. (We identify a jump of particle from site x to y with an exchange of values $\eta_x = 1$ and $\eta_y = 0$.) Let $V_\omega, \omega \in \mathscr{E}^*$, denote the map from $\mathscr{X}$ to $\mathscr{X}$ defined by $V_\omega(\eta) \equiv (\ldots \eta'_{-1}\eta'_0\eta'_1 \ldots)$ where $(\eta'_x, \eta'_y) = (\eta_y, \eta_x)$ iff $\omega(x, y) = 1$, and $\eta'_i = \eta_i$ otherwise.

A random movement of infinitely many particles starting from η is introduced by indicating a map $V_\omega, \omega \in \mathscr{E}^*$, randomly. Let $\{\Theta_\eta | \eta \in \mathscr{X}\}$ be a set of probability measures on $\mathscr{E}$ satisfying $\Theta_\eta(\mathscr{E}^*) = 1$. Suppose $\boldsymbol{e}_\eta$, $\eta \in \mathscr{X}$, is a random element which takes its value on $\mathscr{E}$ and of which distribution is Θ_η. (As such $\boldsymbol{e}_\eta$ we can take an identity map on $\mathscr{E}$.) Then by considering a random map V_{e_η}, we have a random configuration $V_{e_\eta}(\eta) \in \mathscr{X}$ starting from η whose distribution is controlled by Θ_η. For $\eta \in \mathscr{X}$ and $A \in \mathscr{B}$, we define

$$(6.1) \qquad P(\eta, A) = \mathrm{Prob}\{V_{e_\eta}(\eta) \in A\} = \Theta_\eta\{\omega \in \mathscr{E} : V_\omega(\eta) \in A\},$$

which is the probability that η jumps into a set A. In this way we can define a set of transition probabilities $\{P(\eta, A)\}$ and have a discrete time Markov process on $\mathscr{X}$ under which each particle undergoes an interactive exclusive random walk on $\mathbf{Z}$. In the following we will refer to this process by (DX).

Let $\mathscr{E}_{i,j}, i \le j, i, j \in \mathbf{Z}$, be the set of all basic cylinders $\boldsymbol{E} \subset \mathscr{E}$ given by

$$\boldsymbol{E} \equiv \{\omega \in \mathscr{E} : \omega(x, y) = e_{xy} \text{ for } (x, y) \in \mathscr{Z}, i \le x < y \le j\}, \quad e_{xy} = 0 \text{ or } 1,$$

and endow $\mathscr{E}$ with the Borel structure generated by $\sigma(\mathscr{E}_{i,j}), i \le j$. (It is easy to see that $\mathscr{E}^*$ is a measurable set.) We define $\boldsymbol{E}(x, y)$ just as $\omega(x, y)$ and put $\boldsymbol{E}(x, y) = 0$ if $\boldsymbol{E}(x, y)$ is not defined. We also set

$$\mathscr{E}^*_{i,j} = \{\boldsymbol{E} \in \mathscr{E}_{i,j} \,|\, \text{if } \boldsymbol{E}(x, y) = 1 \text{ then } \boldsymbol{E}(\tilde{x}, \tilde{y}) = 0 \text{ for all } (\tilde{x}, \tilde{y}) \in \mathscr{Z}$$
$$\text{satisfying } (\tilde{x}, \tilde{y}) \cap (x, y) \ne \varnothing \text{ and } (\tilde{x}, \tilde{y}) \ne (x, y)\}.$$

For $\boldsymbol{E} \in \mathscr{E}^*_{i,j}$ let $\boldsymbol{F}\langle \boldsymbol{E}\rangle$ denote an element of $\mathscr{E}^*_{I,J}, I \le i, j \le J$, such that $\boldsymbol{F}\langle \boldsymbol{E}\rangle(x, y) = \boldsymbol{E}(x, y)$ for every $(x, y) \in \mathscr{Z}$ with $i \le x < y \le j$. We also define $\sum^*_{\boldsymbol{F}\langle * \rangle}$ analogously to $\sum^*_{\boldsymbol{b}\langle * \rangle}$. Given $\boldsymbol{E} \in \mathscr{E}^*_{i+R,j-R}$ let $V_{\boldsymbol{E}} : \mathscr{C}_{i,j} \to \mathscr{C}_{i,j}$ be the map defined by $V_{\boldsymbol{E}}({}_i[a_i \dots a_j]_j) ={}_i [a'_i \dots a'_j]_j$ where $(a'_x, a'_y) = (a_y, a_x)$ iff $\boldsymbol{E}(x, y) = 1$, and $a'_x = a_x$ otherwise. For $\boldsymbol{a} \in \mathscr{C}_{i,j}$, $\boldsymbol{b}\langle \boldsymbol{a}\rangle \in \mathscr{C}_{I,J}$ and $\boldsymbol{E} \in \mathscr{E}^*_{i+R,j-R}$, we set $V_{\boldsymbol{E}}(\boldsymbol{b}\langle \boldsymbol{a}\rangle) = \boldsymbol{b}\langle V_{\boldsymbol{E}}(\boldsymbol{a})\rangle$. The definition of energy $\mathscr{H}(\boldsymbol{a})$ for $\boldsymbol{a}$ is the same as before except for $a_x = 0$ or 1.

We assume that our transition rule Θ_η, $\eta \in \mathscr{X}$, whose example will be given at the end, satisfy the following ($\mathrm{FD1_x}$)–($\mathrm{FD5_x}$):

($\mathrm{FD1_x}$) $\Theta_\eta(\mathscr{E}^*) = 1$ and $\Theta(\eta, \boldsymbol{E}) \equiv \Theta_\eta(\boldsymbol{E}) > 0$ for every η and $\boldsymbol{E} \in \mathscr{E}^*_{i,j}$.

($\mathrm{FD2_x}$) Given $\boldsymbol{E} \in \mathscr{E}^*_{i,j}$, $\Theta(\eta, \boldsymbol{E})$ is $\mathscr{B}_{i-R,j+R}$-measurable as a function of η.

($\mathrm{FD3_x}$) (i) For every $\boldsymbol{b} \in \mathscr{C}_{i-R,j+R}$ and $\boldsymbol{E} \in \mathscr{E}^*_{i,j}$,

$$(2.3_\mathrm{x}) \qquad \Theta(V_{\boldsymbol{E}}(\boldsymbol{b}), \boldsymbol{E}) \exp\{-\mathscr{H}(V_{\boldsymbol{E}}(\boldsymbol{b}))\} = \Theta(\boldsymbol{b}, \boldsymbol{E}) \exp\{-\mathscr{H}(\boldsymbol{b})\},$$

(ii) For every $\boldsymbol{b} \in \mathscr{C}_{i-R,j+R}$ and $\boldsymbol{D} \in \mathscr{E}^*_{i+2R,j-2R}, \boldsymbol{E}\langle \boldsymbol{D}\rangle \in \mathscr{E}^*_{i,j}$,

$$(2.4_\mathrm{x}) \qquad \Theta(V_{\boldsymbol{E}\langle \mathbf{O}\rangle}(\boldsymbol{b}), \boldsymbol{E}\langle \boldsymbol{D}\rangle) \exp\{-\mathscr{H}(V_{\boldsymbol{E}\langle \mathbf{O}\rangle}(\boldsymbol{b}))\}$$
$$= \Theta(\boldsymbol{b}, \boldsymbol{E}\langle \boldsymbol{D}\rangle) \exp\{-\mathscr{H}(\boldsymbol{b})\},$$

where

$$\boldsymbol{E}\langle \mathbf{O}\rangle(x, y) = \begin{cases} 0 & \text{if } \boldsymbol{D}(x, y) = 1 \\ \boldsymbol{E}\langle \boldsymbol{D}\rangle(x, y) & \text{otherwise.} \end{cases}$$

(FD4$_x$) There exists a positive integer $K_1(>2R)$ such that values

$$\frac{\Theta(\boldsymbol{d}\langle \boldsymbol{c}\langle \boldsymbol{b}\rangle\rangle, \boldsymbol{G}\langle \mathbf{O}\langle \boldsymbol{E}\rangle\rangle)}{\Theta(\boldsymbol{d}\langle \boldsymbol{c}\langle \boldsymbol{b}\rangle\rangle, \tilde{\boldsymbol{G}}\langle \mathbf{O}\langle \boldsymbol{E}\rangle\rangle)}$$

$$\boldsymbol{b} \in \mathscr{C}_{i,j}, \quad \boldsymbol{c}\langle \boldsymbol{b}\rangle \in \mathscr{C}_{i-K_1-L,j+K_1+L},$$

$$\boldsymbol{d}\langle \boldsymbol{c}\langle \boldsymbol{b}\rangle\rangle \in \mathscr{C}_{i-2R-K_1-L,j+2R+K_1+L},$$

$$\boldsymbol{E} \in \mathscr{E}^*_{i+R,j-R}, \quad \mathbf{O}\langle \boldsymbol{E}\rangle \in \mathscr{E}^*_{i+R-K_1-L,j-R+K_1+L},$$

$$\boldsymbol{G}\langle \mathbf{O}\langle \boldsymbol{E}\rangle\rangle, \ \tilde{\boldsymbol{G}}\langle \mathbf{O}\langle \boldsymbol{E}\rangle\rangle \in \mathscr{E}^*_{i-R-K_1-L,j+R+K_1+L},$$

$$L = 0 \text{ or } 2R,$$

are independent of $\boldsymbol{b}$ and $\boldsymbol{E}$.

(FD5$_x$) There exists a constant $M_1 > 0$ such that

$$\frac{\Theta(\boldsymbol{b}, \boldsymbol{E}\langle \boldsymbol{D}\rangle)}{\Theta(\boldsymbol{b}, \tilde{\boldsymbol{E}}\langle \boldsymbol{D}\rangle)} < M_1$$

for any $\boldsymbol{b} \in \mathscr{C}_{i-R-L,j+R+L}$, $\boldsymbol{D} \in \mathscr{E}^*_{i,j}$, $\boldsymbol{E}\langle \boldsymbol{D}\rangle, \tilde{\boldsymbol{E}}\langle \boldsymbol{D}\rangle \in \mathscr{E}^*_{i-L,j+L}$ and $0 < L \leq K_1 + 4R$.

Notice that the exchange of particles on the sites $\{i, \dots, j\}$ is influenced by the particles on the sites at most $\{i - R, \dots, j + R\}$ from (FD2$_x$). Therefore if $\boldsymbol{b}$ is in $\mathscr{B}_{i,j}$, then $P(\cdot, \boldsymbol{b})$ becomes $\mathscr{B}_{i-2R,j+2R}$-measurable.

A probability measure ν on $\mathscr{X}$ is called a *canonical Gibbs state* associated with the potentials $J_r, r = 1, \dots, R$, if it satisfies

$$\nu(\boldsymbol{b}\langle \boldsymbol{a}\rangle)\exp\{-\mathscr{H}(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\rangle)\} = \nu(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}\rangle)\exp\{-\mathscr{H}(\boldsymbol{b}\langle \boldsymbol{a}\rangle)\} \tag{2.5$_x$}$$

for every $\boldsymbol{a}, \tilde{\boldsymbol{a}} \in \mathscr{C}_{i,j}$ and $\boldsymbol{b}\langle * \rangle \in \mathscr{C}_{i-R,j+R}$, satisfying

$$\sum_{\ell=i}^{j} a_\ell = \sum_{\ell=i}^{j} \tilde{a}_\ell. \tag{6.2}$$

(Notice that in (2.5$_x$) the value of J_0 in $\mathscr{H}(\cdot)$ has no meaning because of (6.2).) The set of such canonical Gibbs states is written by $\mathscr{G}_c$. We remark that if $\Theta(\cdot, \cdot)$ satisfies (2.3$_x$), then $\nu \in \mathscr{G}_c$ is equivalent to the following equation:

$$\Theta(V_{\boldsymbol{E}}(\boldsymbol{b}), \boldsymbol{E})\nu(V_{\boldsymbol{E}}(\boldsymbol{b})) = \Theta(\boldsymbol{b}, \boldsymbol{E})\nu(\boldsymbol{b}), \quad \boldsymbol{E} \in \mathscr{E}^*_{i,j}, \quad \boldsymbol{b} \in \mathscr{C}_{i-R,j+R}. \tag{2.6$_x$}$$

The set of stationary measures for (DX) is denoted by $\mathscr{I}$. Now we can state the theorem as follows:

THEOREM 2. *Assume the conditions* (FD1$_x$)–(FD5$_x$). *Then* $\mathscr{I} = \mathscr{G}_c$, *that is, a probability measure* ν *on* $\mathscr{X}$ *is stationary for the exclusion process* (*DX*) *if* *and only if it is a canonical Gibbs state associated with the potentials $J_r, r = 1, \ldots, R$.*

It is known (see, e.g., [1, 2]) that the set of extremal points of the closed convex set $\mathscr{G}_c$ in the topology of weak convergence is equal to the set $\{\mu_\rho | -\infty < \rho < \infty\} \cup \{\delta_0, \delta_1\}$, where μ_ρ is the unique Gibbs state associated with the potentials $J_r, r = 1, \ldots, R$, with $J_0 = \rho$; and δ_0 [resp. δ_1] is a Dirac measure concentrated at $\mathbf{0} \equiv (\ldots 0000 \ldots)$ [resp. $\mathbf{1} \equiv (\ldots 1111 \ldots)] \in \mathscr{X}$. Therefore we get the complete description of the stationary measures for (DX).

COROLLARY 2. $\operatorname{ext} \mathscr{J} = \{\mu_\rho | -\infty < \rho < \infty\} \cup \{\delta_0, \delta_1\}$, *where* $\operatorname{ext} \mathscr{J}$ *denotes the totality of extremal points of* $\mathscr{J}$.

We can also prove that

COROLLARY 3. *Every stationary measure for (DX) is reversible.*

Proofs for the sufficiency part of the theorem and the corollary 3 are just the same as for (DI) by virtue of (6.1) and (2.6_{x}). For the proof of necessity part we have only to set

$$(3.4_{\mathrm{x}}) \qquad \tilde{\nu}(\boldsymbol{a}) = \sum\nolimits^*_{\boldsymbol{b}\langle * \rangle \in \mathscr{C}_{-N-2R, N+2R}} \sum\nolimits_{\boldsymbol{E} \in \mathscr{E}^*_{-N-R, N+R}} \nu(V_{\boldsymbol{E}}(\boldsymbol{b}\langle \boldsymbol{a} \rangle)) \Theta(V_{\boldsymbol{E}}(\boldsymbol{b}\langle \boldsymbol{a} \rangle), \boldsymbol{E})$$

$$= \sum\nolimits_{\boldsymbol{D} \in \mathscr{E}^*_{-N+R, N-R}} X(\boldsymbol{a}, \boldsymbol{D}), \qquad \boldsymbol{a} \in \mathscr{C}_{-N,N},$$

where

$$(3.5_{\mathrm{x}}) \qquad X(\boldsymbol{a}, \boldsymbol{D}) = \sum\nolimits^*_{\boldsymbol{b}\langle * \rangle} \sum\nolimits^*_{\boldsymbol{E}\langle * \rangle} \nu(V_{\boldsymbol{E}\langle \boldsymbol{D} \rangle}(\boldsymbol{b}\langle \boldsymbol{a} \rangle)) \Theta(V_{\boldsymbol{E}\langle \boldsymbol{D} \rangle}(\boldsymbol{b}\langle \boldsymbol{a} \rangle), \boldsymbol{E}\langle \boldsymbol{D} \rangle),$$

and apply the preceding argument under the assumption $\nu(\{\mathbf{0}, \mathbf{1}\}) = 0$. (It is easy to see that $\{\mathbf{0}, \mathbf{1}\} \in \mathscr{J} \cap \mathscr{G}_c$.) In this case $I_N(\nu)$ in (3.6) becomes

$$I_N(\nu) \equiv \sum\nolimits_{\boldsymbol{a}} \mu(a) \sum\nolimits_{\boldsymbol{D}} \Theta(\boldsymbol{a}, \boldsymbol{D}) \Psi\left(\frac{X(\boldsymbol{a}, \boldsymbol{D})}{\Theta(\boldsymbol{a}, \boldsymbol{D})\mu(\boldsymbol{a})}\right),$$

and $g(\boldsymbol{c}, \check{\boldsymbol{y}}\langle \boldsymbol{v} \rangle)$ in Lemmas 3 and 4 should be replaced with

$$g(\boldsymbol{c}, \check{\boldsymbol{F}}\langle \boldsymbol{C} \rangle) = \log\left(\sum\nolimits^*_{\boldsymbol{F}\langle * \rangle} \frac{X(\boldsymbol{c}, \boldsymbol{F}\langle \boldsymbol{C} \rangle)}{\sum^*_{\hat{\boldsymbol{F}}\langle * \rangle} X(\boldsymbol{c}, \hat{\boldsymbol{F}}\langle \boldsymbol{C} \rangle)} \cdot \frac{X(\boldsymbol{c}, \check{\boldsymbol{F}}\langle \boldsymbol{C} \rangle)}{X(\boldsymbol{c}, \boldsymbol{F}\langle \boldsymbol{C} \rangle)} \cdot \frac{\Theta(\boldsymbol{c}, \boldsymbol{F}\langle \boldsymbol{C} \rangle)}{\Theta(\boldsymbol{c}, \check{\boldsymbol{F}}\langle \boldsymbol{C} \rangle)}\right)$$

$$- \sum\nolimits^*_{\boldsymbol{F}\langle * \rangle} \frac{X(\boldsymbol{c}, \boldsymbol{F}\langle \boldsymbol{C} \rangle)}{\sum^*_{\hat{\boldsymbol{F}}\langle * \rangle} X(\boldsymbol{c}, \hat{\boldsymbol{F}}\langle \boldsymbol{C} \rangle)} \log\left(\frac{X(\boldsymbol{c}, \check{\boldsymbol{F}}\langle \boldsymbol{C} \rangle)}{X(\boldsymbol{c}, \boldsymbol{F}\langle \boldsymbol{C} \rangle)} \cdot \frac{\Theta(\boldsymbol{c}, \boldsymbol{F}\langle \boldsymbol{C} \rangle)}{\Theta(\boldsymbol{c}, \check{\boldsymbol{F}}\langle \boldsymbol{C} \rangle)}\right),$$

where $\boldsymbol{c} \in \mathscr{G}_{-N-2R-K_1, N+2R+K_{-1}}$, $\boldsymbol{C} \in \mathscr{E}^*_{-N+3R, N-3R}$ and $\check{\boldsymbol{F}}\langle * \rangle \in \mathscr{E}^*_{-N-R-K_1, N+R+K_1}$. Then the proofs given in Sections 3 and 4 go through parallelly after few modifications: (1) The constant G which will appear in the proof of Lemma 4 such as $\{\boldsymbol{F}_1\langle \boldsymbol{C}\rangle, \boldsymbol{F}_2\langle \boldsymbol{C}\rangle, \dots, \boldsymbol{F}_G\langle \boldsymbol{C}\rangle\}$ varies depending on $\boldsymbol{C}$, but they are uniformly bounded from above. Hence we can apply the proof of Lemma 4 for each class of $\boldsymbol{C}$ whose G is the same. (2) In exclusion processes, differently from stochastic Ising models, particles can not be born or disappear; and so a configuration of particles on an interval can not change to another configuration directly at one time. Hence the proof of Lemma 6 needs another discussion under the assumption $\nu(\{\mathbf{0},\mathbf{1}\}) = 0$. However the proof is elementary.

PROOF OF LEMMA 6 FOR (DX) UNDER $\nu(\{\mathbf{0},\mathbf{1}\}) = 0$. It is easy to see that if $\boldsymbol{a} \in \mathscr{C}_{i,j}$ is of the form $\boldsymbol{a} =_i [V_{\boldsymbol{E}_n} V_{\boldsymbol{E}_{n-1}} \dots V_{\boldsymbol{E}_1}(\boldsymbol{b})]_j$ for some $\boldsymbol{b} \in \mathscr{C}_{I,J}, \nu(\boldsymbol{b}) > 0$, and $\{\boldsymbol{E}_k\}_{k=1}^n \subset \mathscr{E}^*_{I+R,J-R}$, then $\nu(\boldsymbol{a}) > 0$. This implies that if $\nu(\boldsymbol{a}) = 0$, then $\nu(\boldsymbol{c}) = 0$ for every $\boldsymbol{c} \in \mathscr{C}$ satisfying $\sharp_1(\boldsymbol{a}) \le \sharp_1(\boldsymbol{c})$ and $\sharp_0(\boldsymbol{a}) \le \sharp_0(\boldsymbol{c})$, where $\sharp_1(\boldsymbol{a}) = \sum_{k=i}^j a_k$, $\sharp_0(\boldsymbol{a}) = \sum_{k=i}^j (1-a_k)$ and so on.

Now suppose $\nu(\boldsymbol{a}) = 0$ for some $\boldsymbol{a} \neq \varnothing$. From the above fact we have $\nu(\{\eta \in \mathscr{X} \mid \sum_{k=-\infty}^{\infty} \eta_k \ge \sharp_1(\boldsymbol{a})$ and $\sum_{k=-\infty}^{\infty}(1-\eta_k) \ge \sharp_0(\boldsymbol{a})\}) = 0$. This implies $\nu(\{\eta \in \mathscr{X} \mid \sum_{k=-\infty}^{\infty} \eta_k < \sharp_1(\boldsymbol{a})$ or $\sum_{k=-\infty}^{\infty}(1-\eta_k) < \sharp_0(\boldsymbol{a})\}) = 1$, and hence

$$\sum_{q=1}^{\sharp_1(\boldsymbol{a})-1} \nu(Z_{1,q}) + \sum_{q=1}^{\sharp_0(\boldsymbol{a})-1} \nu(Z_{0,q}) = 1, \tag{6.3}$$

where $Z_{1,q} \equiv \{\eta \in \mathscr{X} \mid \sum_{k=-\infty}^{\infty} \eta_k = q\}$ and $Z_{0,q} \equiv \{\eta \in \mathscr{X} \mid \sum_{k=-\infty}^{\infty}(1-\eta_k) = q\}$. (Note that $\nu(Z_{1,0}) = \nu(Z_{0,0}) = 0$ from the assumption $\nu(\{\mathbf{0},\mathbf{1}\}) = 0$).

It is straightforward to show that if $\nu(Z_{1,q+d}) = 0$ for $d = 1, \dots, R$, then $\nu(Z_{1,q}) = 0$. Indeed for each $\eta \in Z_{1,q}$ choose N_η sufficiently large so that $\sum_{k=-N_\eta+R}^{N_\eta-R} \eta_k = q$ and put $:\eta: =_{-N_\eta} [\eta_{-N_\eta} \dots \eta_{N_\eta}]_{N_\eta}$. Then define

$$M_{1,q} = \max\{\nu(:\eta:)\exp\{\mathscr{H}(:\eta:)\} \mid \eta \in \boldsymbol{Z}_{1,q}\},$$

which is well-defined since $\nu(\mathscr{X}) = 1$. Suppose $\nu(:\hat{\eta}:)\exp\{\mathscr{H}(:\hat{\eta}:)\} = M_{1,q}$. Taking $\boldsymbol{a}$ to be $:\hat{\eta}: \equiv_{-N_{\hat{\eta}}} [\hat{\eta}_{-N_{\hat{\eta}}} \dots \hat{\eta}_{N_{\hat{\eta}}}]_{n_{\hat{\eta}}}$ in (3.4_{x}) we have

$$\begin{aligned}
\nu(:\hat{\eta}:) &= \sum\nolimits^*_{\boldsymbol{c}\langle * \rangle \in \mathscr{C}_{-N_{\hat{\eta}}-2R, N_{\hat{\eta}}+2R}} \sum\nolimits_{\boldsymbol{E}} \Theta(V_{\boldsymbol{E}}(\boldsymbol{c}\langle :\hat{\eta}:\rangle), \boldsymbol{E})\nu(V_{\boldsymbol{E}}(\boldsymbol{c}\langle :\hat{\eta}:\rangle)) \\
&= \sum\nolimits^*_{\boldsymbol{c}\langle * \rangle} \sum\nolimits_{\boldsymbol{E}} \Theta(\boldsymbol{c}\langle :\hat{\eta}:\rangle, \boldsymbol{E}) \frac{\Theta(V_{\boldsymbol{E}}(\boldsymbol{c}\langle :\hat{\eta}:\rangle), \boldsymbol{E})}{\Theta(\boldsymbol{c}\langle :\hat{\eta}:\rangle, \boldsymbol{E})} \nu(V_{\boldsymbol{E}}(\boldsymbol{c}\langle :\hat{\eta}:\rangle)) \\
&= \sum\nolimits^*_{\boldsymbol{c}\langle * \rangle} \sum\nolimits_{\boldsymbol{E}} \Theta(\boldsymbol{c}\langle :\hat{\eta}:\rangle, \boldsymbol{E}) \frac{\exp\{-\mathscr{H}(\boldsymbol{c}\langle :\eta:\rangle)\}}{\exp\{-\mathscr{H}(V_{\boldsymbol{E}}(\boldsymbol{c}\langle :\eta:\rangle))\}} \nu(V_{\boldsymbol{E}}(\boldsymbol{c}\langle :\hat{\eta}:\rangle)) \\
&= \sum\nolimits^*_{\boldsymbol{D} \in \mathscr{E}^*_{-N_{\hat{\eta}}+R, N_{\hat{\eta}}-R}} \Theta(:\hat{\eta}:, \boldsymbol{D}) \frac{\exp\{-\mathscr{H}(:\hat{\eta}:)\}}{\exp\{-\mathscr{H}(V_{\boldsymbol{D}}(:\hat{\eta}:))\}} \nu(V_{\boldsymbol{D}}(:\hat{\eta}:))
\end{aligned}$$

by (2.3_{x}) and $\nu(Z_{1,q+d}) = 0$, that is,

$$M_{1,q} = \sum\nolimits_{\boldsymbol{D}} \Theta(:\hat{\eta}:, \boldsymbol{D})\nu(V_{\boldsymbol{D}}(:\hat{\eta}:))\exp\{\mathscr{H}(V_{\boldsymbol{D}}(:\hat{\eta}:))\},$$

which implies $\nu(V_{\boldsymbol{D}}(:\hat{\eta}:))\exp\{\mathscr{H}(V_{\boldsymbol{D}}(:\hat{\eta}:))\} = M_{1,q}$ for every $\boldsymbol{D}$ since $\sum_{\boldsymbol{D}} \Theta(:\hat{\eta}:, \boldsymbol{D}) = 1$. Thus we know that $\nu(:\eta:)\exp\{\mathscr{H}(:\eta:)\} = M_{1,q}$ if $:\eta:$ is related to $:\hat{\eta}:$ by $:\eta: = V_{\boldsymbol{D}}(:\hat{\eta}:)$ for some $\boldsymbol{D}$. By repeating this argument we have $\nu(:\eta:)\exp\{\mathscr{H}(:\eta:)\} = M_{1,q}$ for all $\eta \in Z_{1,q}$. But this is impossible if $M_{1,q} > 0$ since $\nu(\mathscr{X}) = 1$; and so $M_{1,q}$ must be zero. Hence $\nu(Z_{1,q}) = 0$ if $\nu(Z_{1,q+d}) = 0$ for $d = 1, \dots, R$.

This argument gives us $\nu(Z_{1,q}) = 0$ for $q = \sharp_1(\boldsymbol{a}) - 1, \dots, 2, 1$ because $\nu(Z_{1,q+d}) = 0$ for $q = \sharp_1(\boldsymbol{a}) - 1, d = 1, \dots, R$ by (6.3). So we have $\sum_{q=1}^{\sharp_1(\boldsymbol{a})-1} \nu(Z_{1,q}) = 0$ in (6.3). Analogously we can show that $\sum_{q=1}^{\sharp_0(\boldsymbol{a})-1} \nu(Z_{0,q}) = 0$ by choosing N_η sufficiently large so that $\sum_{k=N_\eta+R}^{N_\eta-R}(1-\eta_k) = q$ and putting $M_{0,q} = \max\{\nu(:\eta:)\exp\{\mathscr{H}(:\eta:) - \mathscr{H}({}_{-N_\eta}[\mathbf{1}]_{N_\eta})\} \mid \eta \in Z_{0,q}\}$. Consequently the l.h.s. of (6.3) reduces to zero, a contradiction. Therefore $\nu(\boldsymbol{a}) > 0$ for any $\boldsymbol{a} \neq \varnothing$ if ν is stationary and $\nu(\{\mathbf{0},\mathbf{1}\}) = 0$. $\square$

Finally we give an example of (DX) by defining $\{\Theta(\eta,\cdot) | \eta \in \mathscr{X}\}$, of which idea and interpretation are essentially the same as in §5 (see also (2.1) and (2.2)). Let α_i's and β_i's, $i = 1, \dots, R$, be numbers in $(0,1)$. Suppose that the configuration of particles on $\mathbf{Z}$ at time t is $\eta \in \mathscr{X}$. Then we attach the state $\circ$(permission) or $\times$(prohibition) to each $(x,y) \in \mathscr{Z}$ as follows:

i) if $\eta_x \neq \eta_y$, attach $\circ$ [resp. $\times$] to (x,y) with probability α_{y-x} [resp. $1-\alpha_{y-x}$];

ii) if $\eta_x = \eta_y$, attach $\circ$ [resp. $\times$] to (x,y) with probability β_{y-x} [resp. $1-\beta_{y-x}$];

iii) The random choices of $\circ$ and $\times$ are independent.

This rule defines a probability measure $\hat{Q}_\eta$ on $\{\circ, \times\}^{\mathscr{Z}}$ which depends on η. As time goes to $t+1$, we exchange the state η_x and η_y if the state $\circ$ is attached to (x,y) and the state $\times$ is attached to all $(\tilde{x},\tilde{y}) \in \mathscr{Z}$ satisfying $\{\tilde{x},\tilde{y}\} \cap \{x,y\} \neq \varnothing$ and $(\tilde{x},\tilde{y}) \neq (x,y)$. This defines a probability measure Θ_η on $\mathscr{E}$ satisfying $\Theta_\eta(\mathscr{E}^*) = 1$. If we define

$$J_r = 2\log\left(\frac{1-\beta_r}{1-\alpha_r}\right), \qquad r = 1, \dots, R,$$

we can show just as in §5 that Θ_η satisfies $(\mathrm{FD1_x})$–$(\mathrm{FD5_x})$. Thus, by Theorem 2, we know that the set $\mathscr{J}$ of stationary measures for the corresponding (DX) is equal to the set $\mathscr{G}_c$ of canonical Gibbs states with pair potentials $J_r, r = 1, \dots, R$. Especially an extremal point of $\mathscr{J}$ which is different from $\delta_{\mathbf{0}}$ and $\delta_{\mathbf{1}}$ is a Gibbs state with some self-potential J_0.

Concluding Remarks. **1**. The proof of our theorem is making use of the one-dimensionality of the configuration space $\mathscr{X}$ in the sense that the number of sites within the distance R from the boundary of each interval is bounded uniformly. (This is used in the proofs of Lemmas 2, 4 and 5 as estimates $\sharp\{+1,-1\}^R(=2^R)$'s.) So our argument here does not go through directly in higher dimensional cases.

2. We think that the conditions (2.3) and (2.4) in (FD3) should be grasped as a special case of the following general "dynamic local equilibrium": Let $\boldsymbol{a}, \tilde{\boldsymbol{a}} \in \mathscr{C}_{i,j}$ and $\boldsymbol{b}\langle * \rangle \in \mathscr{C}_{i-R,j+R}$. Then for every $A \subset \{i, i+1, \ldots, j\}$

$$\begin{aligned} &\exp\{-\mathscr{H}(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}|_A \cdot \boldsymbol{a}|_{\bar{A}}\rangle\} P(\boldsymbol{b}\langle \tilde{\boldsymbol{a}}|_A \cdot \boldsymbol{a}|_{\bar{A}}\rangle, \boldsymbol{a}|_A \cdot \tilde{\boldsymbol{a}}|_{\bar{A}}) \\ &\qquad = \exp\{-\mathscr{H}(\boldsymbol{b}\langle \boldsymbol{a}|_A \cdot \boldsymbol{a}|_{\bar{A}}\rangle\} P(\boldsymbol{b}\langle \boldsymbol{a}|_A \cdot \boldsymbol{a}|_{\bar{A}}\rangle, \tilde{\boldsymbol{a}}|_A \cdot \tilde{\boldsymbol{a}}|_{\bar{A}}) \\ &\qquad (\equiv \exp\{-\mathscr{H}(\boldsymbol{b}\langle \boldsymbol{a}\rangle\} P(\boldsymbol{b}\langle \boldsymbol{a}\rangle, \tilde{\boldsymbol{a}})), \end{aligned}$$

where $\tilde{\boldsymbol{a}}|_A \cdot \boldsymbol{a}|_{\bar{A}}$ is the element of $\mathscr{C}_{i,j}$ such that spin-orientations on A are the same as $\tilde{\boldsymbol{a}}$ and on $\bar{A} \equiv \{i, \ldots, j\} \backslash A$ the same as $\boldsymbol{a}$, and so on. If $A = \{i, \ldots, j\}$ [resp. $= \{i, \ldots, i+R\} \cup \{j-R, \ldots, j\}$], the above condition reduces to (2.3) [resp. (2.4)].

Acknowledgment

The author would like to express his gratitude to Professors H. Murata, Y. Higuchi and I. Kubo for their advices and interests in these models during the preparation of this paper. He also thanks his PC9801 computer without which he would not have reached the proof of Lemma 1.

References

[1] H. O. Georgii, Canonical Gibbs Measures, Lecture Notes in Math. Vol. 760, Springer-Verlag, Berlin · Heidelberg · New York, 1979.
[2] H. O. Georgii, Gibbs measures and phase transitions, Walter de Gruyter, Berlin · New York, 1988.
[3] R. Holley, Free energy in a Markovian model of a lattice spin system, Comm. Math. Phys. **23** (1971), 87–99.
[4] R. Holley and D. Stroock, In one and two dimensions, every stationary measure for a stochastic Ising model is a Gibbs state, Comm. Math. Phys. **55** (1977), 37–45.
[5] T. M. Liggett, Coupling the simple exclusion process, Ann. Probab. **4** (1976), 339–356.
[6] T. M. Liggett, Interacting particle systems, Springer-Verlag, Berlin · Heidelberg · New York, 1985.
[7] T. M. Liggett, Survival and coexistence in interacting particle systems, Probability and phase transition (G. Grimmett, ed.), 209–226, Kluwer Acad. Publ, Dordrecht, 1994.
[8] F. Martinelli and E. Olivieri, Approach to equilibrium of Glauber dynamics in the one phase region I, II, Comm. Math. Phys. **161** (1994), 447–486, 487–514.
[9] D. W. Stroock and B. Zegarlinski, The logarithmic Sobolev inequality for discrete spin systems on a lattice, Comm. Math. Phys. **149** (1992), 175–193.

[10] H. Yaguchi, Stationary measures for an exclusion process on one-dimensional lattices with infinitely many hopping sites, Hiroshima Math. J. **16** (1986), 449–475.

[11] H. Yaguchi, Entropy analysis of a nearest neighbor attractive/repulsive exclusion process on one-dimensional lattices, Ann. Probab. **18** (1990), 556–580.

[12] H. Yaguchi, A discrete time interactive exclusive random walk of infinitely many particles on one-dimensional lattices, Hiroshima Math. J. **21** (1991), 267–283.

Department of Mathematics
Faculty of Education
Mie University
Tsu, Mie
514-8507 *Japan*